LENGTH OF THE OBSERVATION PERIOD AS A FACTOR AFFECTING THE CONTRACEPTIVE FAILURE RATE

R. G. POTTER, JR.[1]

INTRODUCTION

THE protection which a group of couples receive from contraception is conventionally measured by the number of pregnancies experienced per 100 years of contraceptive exposure. To compute this rate one must first determine for each couple their number of contraceptive failures as well as the number of months they practiced contraception when there existed a risk of pregnancy. These contraceptive failures and months of contraceptive exposure are then summed for the total group and a pregnancy rate computed by the following formula:

$$\frac{\text{total number of contraceptive failures} \times 1200}{\text{total months of contraceptive exposure}}$$

Naturally one may use this failure rate to compare the contraceptive performances of two groups. But a frequent error is to infer more than is justified. Suppose that contraceptive failure rates have been computed for two independent random samples and that Sample A exhibits a lower rate than Sample B. If the difference is large enough to dismiss sampling variability as a plausible explanation, then one is entitled to infer that probably the couples of Population A are enjoying a lower pregnancy rate during contraception than members of Population B. But unless certain controls have been built into the comparison of Samples A and B, one is not justified in going further to attribute this lower pregnancy rate to more efficient contraception.

To justify this additional inference, several controls are cru-

[1] Office of Population Research, Princeton University. The writer is indebted to Doctors F. W. Notestein, P. C. Sagi, and C. Tietze for many helpful comments and to Mrs. E. Harm for her computational work.

cial. For example, it is essential that the two samples be similar in fecundity, that contraceptive failures be uniformly defined for both groups, and that postpartum amenorrhea and anovulatory cycles be as successfully eliminated from the contraceptive exposures of one sample as the other. An especially difficult problem of control is minimizing the number of couples lost before the end of the study for such reasons as changed address or disinterest in the investigation. The importance of controlling these several factors is widely recognized. What is not generally appreciated, however, is the need for standardizing the length of the observation period. As will be shown, the same sample observed for a longer period exhibits a lower failure rate.

The objective of this paper is to demonstrate that the contraceptive failure rate is sufficiently sensitive to the length of the observation period so that this factor deserves high priority as a control in any comparative work.

A Model

Before illustrating empirically how sensitive the contraceptive failure rate is to differing lengths of observation period, it is worthwhile to explore, by means of a simplified model, the conditions under which this sensitivity is greater or less. A nonprobabilistic model will be used which draws upon, but also generalizes, certain features of a model recently published by C. Tietze.[2]

In the present model, it is assumed that a large group of couples, starting contraception at the same time, are observed for "t" months. During this observation period, all couples continue practicing contraception unless interrupted by an unplanned pregnancy. Hence their contraceptive exposures either equal "t" months, if they succeed in avoiding pregnancy during the observation period, or else equal the smaller number of months preceding contraceptive failure. It is assumed that contraceptive exposures do not include any months of post-

[2] Tietze, C.: Differential Fecundity and Effectiveness of Contraception. *The Eugenics Review,* January, 1959, 50, No. 4.

partum amenorrhea. Also, none of the couples is lost to the study before the end of the observation period for such reasons as changed address or ceasing to cooperate.

Other simplifying assumptions pertain to the monthly risk of contraceptive failure. Specifically it is assumed that this monthly risk of pregnancy during contraceptive exposure varies among couples but is constant for a single couple. Moreover, a couple's constant risk of unplanned pregnancy is interpreted as a product of two values. First is the couple's fecundability, or monthly likelihood of pregnancy in the absence of contraception. Second is their "contraceptive efficiency," or the percentage reduction they are effecting in their monthly likelihood of pregnancy by practice of contraception. For example, if their contraception is .9 efficient, they are lessening their chance of pregnancy to .1 of its original value in the absence of contraception. Hence, if their fecundability is .5 say, their monthly risk of pregnancy with contraception becomes .1(.5) or .05.

For purposes of the model it is very important that the distribution of fecundabilities be made as realistic as possible. Considerable pains have been taken to derive a distribution of fecundabilities that is plausible for urban United States. The distribution of fecundabilities is graphed in Figure 1. On the assumption that no contraception is being practiced, this distribution yields a set of pregnancy delays which closely match those of a criterion sample of successful contraceptors after deliberately stopping contraception in order to become pregnant. This criterion sample is based on successful contraceptors from the Indianapolis Study and from a large, unselected group of obstetric patients of Baltimore. Together these two series afford perhaps the best data of this type available for urban United States.[3] The fit between model and empirical standard is given in Table 1.

[3] Tietze uses the same empirical standard and for details the reader is referred to his article, *ibid.*, p. 231. Using this standard, Tietze derives a distribution which allocates couples 60:38:2 among fecundabilities of .50, .10, and .01. This 3-point distribution of fecundabilities has computational advantages over the continuous distribution adopted in the present article, but yields a poorer fit with the empirical standard.

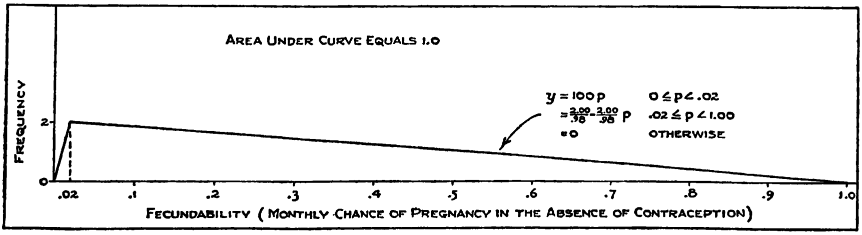
Source of Pregnancy Delays	Percentages of Pregnancies Occurring after Delays of:						
	One Month	2–3	4–6	7–12	13–24	Over 24	Total
Criterion Sample	34.0	27.7	16.3	10.7	5.4	5.7	99.8
Hypothetical Distribution	34.0	27.3	15.1	11.8	5.7	6.1	100.0

Table 1. Comparison of pregnancy delays generated by hypothetical distribution of fecundabilities with those observed in criterion sample.

It is seen from Figure 1 that the distribution of fecundabilities has the form of an asymmetrical triangle. Thus the distribution is continuous except for one point of discontinuity at .02. The fecundabilities of this hypothetical distribution have a standard deviation of .23 and a mean value of .34, this mean corresponding to the proportion of pregnancies occurring in the criterion sample the first month after stopping contraception. From a peak density at .02, the frequency of fecundabilities decreases progressively toward the limits of zero and unity where the frequency becomes zero. Zero frequency at these limits appears reasonable. Zero fecundability is incompatible with pregnancy so that in a fecund group such a fecundability should be lacking. Likewise, a fecundability of 1.0, or certainty of conception during the first month of exposure, is implausible because of such hazards as occasional anovulatory cycles, sickness, and temporary separations.

To complete the model, it is necessary to assume something about contraceptive efficiency and its relationship to fecundability. The simplest assumption is that all couples practice

Fig. 1. A hypothetical distribution of fecundabilities relevant for the urban United States.



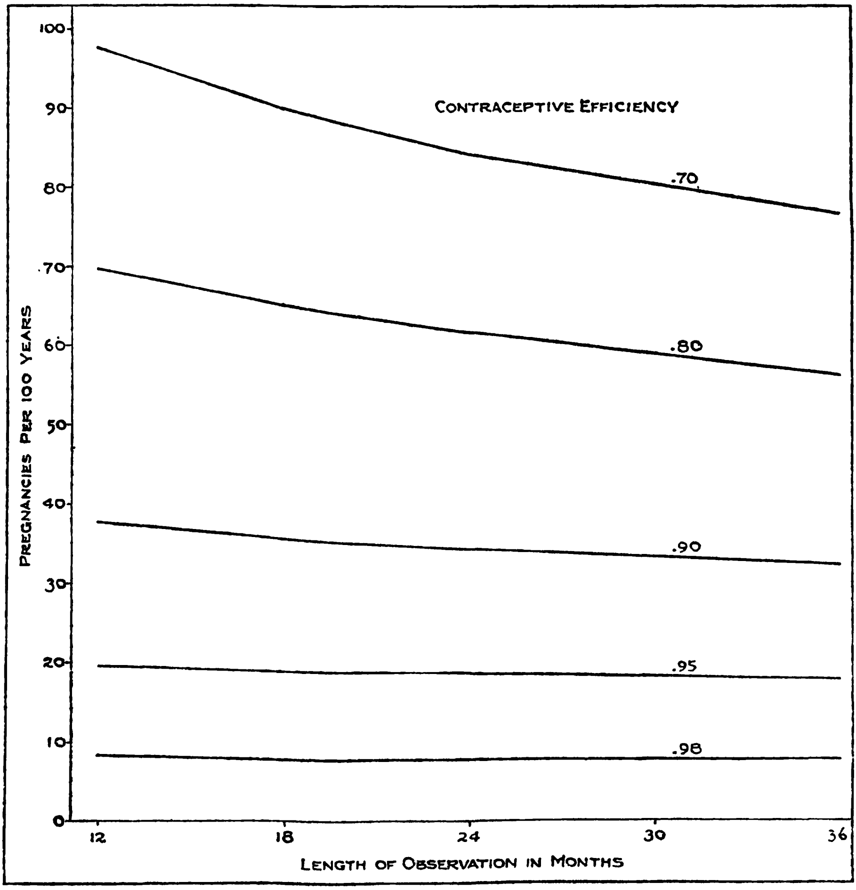
contraception with the same efficiency "e"—i.e., all couples reduce their fecundabilities by a factor of "e," thereby reducing their monthly risk of pregnancy to (1-e)p, where "p" signifies their fecundability. Given this assumption about contraceptive efficiency, together with all the other simplifying assumptions, it becomes possible, with the help of calculus, to estimate P, the proportion who become pregnant during the "t" months of the observation period, as well as E, the average length of contraceptive exposure per couple during this observation span. The mathematical details are given in Appendix A. The contraceptive failure rate is then easily derived as (P/E)1200. Since "t," the length of the observation period, and "e," contraceptive efficiency, are being treated as variables in the model, it is possible to compute failure rates for any combination of contraceptive efficiency and observation length.

To assume that all couples practice the same efficiency of contraception is not very realistic and it turns out that this assumption minimizes the responsiveness of the failure rate to the length of the observation period. Fortunately, there is an alternative assumption which is also easy to apply and which maximizes the dependence of the failure rate upon length of observation period. This is to posit maximum variability of individual efficiencies around the group mean, instead of no variability at all. For example, maximum variability around a mean efficiency of .9 is given when 90 per cent of the group are treated as practicing perfect contraception (efficiency of 1.0) and 10 per cent are treated as practicing perfectly hopeless contraception (efficiency of zero). More generally, if a mean efficiency of "e" is stipulated, posit a proportion of "e" as practicing perfect contraception and a proportion of 1-e as practicing hopeless contraception. Now the typical degrees to which individual efficiencies of contraception vary around group averages are unknown, but obviously these variations lie between the extremes of no variability and algebraically maximal variability. Accordingly, one hopes that two sets of results based on these extreme assumptions will bracket the correct values; of course, still contingent upon the other simplifying assumptions.

Results

Results are summarized in Figures 2 and 3. Figure 2 assumes that all couples are practicing the same efficiency of contraception. Figure 3 assumes that individual efficiencies vary maximally around the group mean. In both charts the declines of failure rate with increases in length of observation are graphed for contraceptive efficiencies of .70, .80, .90, .95, and .98. (For each level of contraceptive efficiency, a curve connects failure rates computed for observation lengths of 12, 18, 24, and 36

Fig. 2. Pregnancies per 100 years of contraceptive exposure as related to length of observation period and mean contraceptive efficiency, assuming that the efficiencies of individual couples equal the group's mean efficiency.



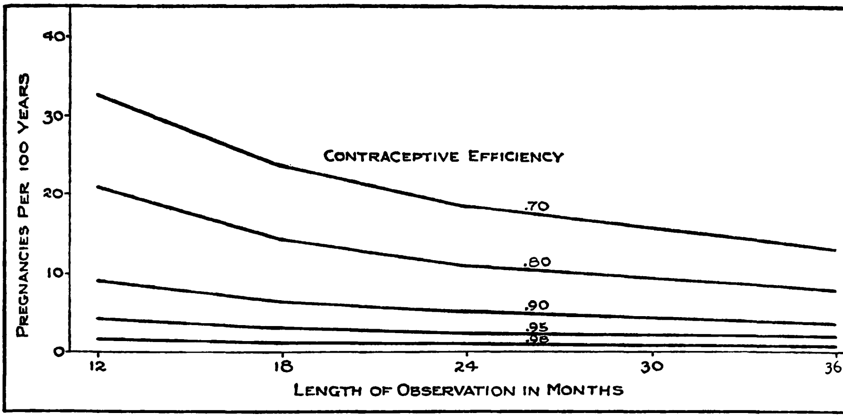


Fig. 3. Pregnancies per 100 years of contraceptive exposure as related to length of observation period and mean contraceptive efficiency, assuming that efficiencies of individual couples are maximally variable around the group's mean efficiency.

months.) It seems hardly useful to consider lower efficiencies since they generate failure rates so much higher than those published.

Several tendencies are worth noting. In both graphs, when mean efficiency of contraception is held constant, the failure rate is consistently lower when the observation period is longer. Moreover the differentials by length of observation period rapidly increase in size as one moves from high to low efficiency of contraception.

Comparing the two graphs, one sees that when individual efficiencies are maximally variable around the group mean, rather than all identical with it, the failure rates are much lower and much more differentiated by length of observation period. Doubtless Figure 3 overestimates the degree to which the failure rate varies as a function of observation length, but just as surely Figure 2 underestimates this responsiveness. Potentially, then, the effects registered upon the failure rate by different lengths of observation period are appreciable even when mean contraceptive efficiency is as high as .95.

Parenthetically it is worth noting, from the contrast between Figures 2 and 3, that there is no simple relationship between the contraceptive failure rate and mean efficiency of contraception as it is being defined in this paper. Given the same mean efficiency, the contraceptive failure rate will be higher or lower depending on whether individual efficiencies are more or less variable around the group mean.

The results just reviewed are based on a number of simplifying assumptions. One assumption is a successful exclusion of postpartum amenorrhea and anovulation from contraceptive exposures. In practice, when dealing with contraception after a birth, it is conventional to allow one month for postpartum sterility. Even in the contemporary urban United States, where few infants are nursed long durations, periods of postpartum amenorrhea and anovulatory cycles probably average three or four months, perhaps longer.[4] Thus, as contraceptive exposure is defined in many studies, it includes at least 2 or 3 months of postpartum sterility. Obviously the inclusion of such sterility lowers contraceptive failure rates and the bias is proportionately greater the shorter the observation period. Therefore the inclusion of postpartum sterility tends to work against the tendency for a failure rate to be higher when the observation period is shorter. But in each instance the net balance between the two biases is problematical.

Another simplifying assumption is that no couple deliberately stops contraception before the end of the observation period. Such deliberate cessations shorten contraceptive exposures and their effect upon the failure rate is similar to that from shortening the observation period. But quantitative estimates are difficult. The pregnancy postponements intended by couples are so highly variable from one sample to another that a general formulation is virtually impossible.

Some Empirical Examples

The distribution of intended pregnancy postponements is especially important when one is dealing with contraceptive histories that encompass entire interpregnancy intervals or at

[4] See Sharman, A.: Ovulation After Pregnancy. *Fertility and Sterility*, 1951, 2, No. 5, pp. 371–393; Guttmacher, A. D.: Fertility of Man. *Fertility and Sterility*, 1952, 3, No. 4, pp. 281–289; and Tietze, C.: *op. cit.*, pp. 235, 236.

least long portions thereof. In this case, more contraceptive exposures may be deliberately terminated than involuntarily interrupted by pregnancy. Typically too, contraceptive exposures are highly variable and average well in excess of 12 months.

To investigate empirically the effect of observation length in the case of these histories, one may first compute a failure rate using entire pregnancy intervals and then compute a second failure rate using only the first twelve months of these intervals. This latter failure rate simulates the consequence of a 12-month observation period. Thus, if a particular couple experiences an unplanned pregnancy during their 15th month of contraception, their experience is defined as 15 months of contraceptive exposure with one unplanned pregnancy in the former failure rate, based on entire intervals, but as 12 months of contraceptive exposure without a pregnancy in the latter failure rate, which considers only the first 12 months of any interval. These pairs of failure rates have been computed for data from the Indianapolis Study and from the Family Growth in Metropolitan America Study.[5]

All couples of this latter study have two children with few of the parents seeking long postponements of either birth. As a result, contraceptive exposures average a little under 18 months before first pregnancy and barely under 24 months following the first birth. A failure rate of 25.5 pregnancies per 100 years of contraceptive exposure is observed during the initial interval and a rate of 20.4 pregnancies during the interval between first birth and next pregnancy. When

[5] Dr. C. V. Kiser kindly made available the contraceptive records of the Indianapolis couples. A thorough report of these data is contained in Westoff, C. F.; Herrera, L. F.; and Whelpton, P. K.: The Use, Effectiveness, and Acceptability of Methods of Fertility Control, in Whelpton, P. K. and Kiser, C. V. (ed.): SOCIAL AND PSYCHOLOGICAL FACTORS AFFECTING FERTILITY, V. 4. New York: Milbank Memorial Fund, 1954, pp. 885–952. A much more specialized analysis of contraceptive materials from the Family Growth In Metropolitan America Study is given in Potter, R. G.: Contraceptive Practice and Birth Intervals Among Two-Child White Couples in Metropolitan America. THIRTY YEARS OF RESEARCH IN HUMAN FERTILITY. New York: Milbank Memorial Fund, 1959, pp. 74–92. A more comprehensive report of these data will be furnished in Westhoff, Charles F.; Potter, Robert G., Jr.; Sagi, Philip C.; and Mishler, Elliot G.: FAMILY GROWTH IN METROPOLITAN AMERICA. To be published by Princeton University Press.

contraceptive exposures are truncated after the twelfth month, to simulate an observation period of that length, the failure rate of 25.5 jumps to 35.7, while the other failure rate jumps from 20.4 to 27.2. These represent increases of 40 and 35 per cent.

These percentage increases would be less similar except for a balancing of factors. Since contraceptive exposures average longer after first birth than before first pregnancy—24 months instead of 18—one would expect that truncating at 12 months would increase the failure rate more in the later period. But this expectation is not borne out because only one month is allowed for postpartum amenorrhea. As a result, the contraceptive exposures of this later period include appreciable amounts of postpartum sterility and of course such sterility reduces failure rates more when intervals are truncated than when they are left unrestricted.

The initial pregnancy interval of the Indianapolis Study offers an even more spectacular example of what happens when contraceptive histories are truncated at 12 months. The contraceptive histories of this study span marriage durations of 12 to 15 years, with some of the couples practicing contraception this entire time without a pregnancy. Partly because of these couples, contraceptive exposures prior to first pregnancy average 45 months. The failure rate for this first pregnancy interval is 14.4 pregnancies per 100 years of contraceptive exposure.[6] But when exposures are truncated at 12 months, this failure rate almost trebles to 39.7. Clearly a failure rate based on contraceptive histories is not always comparable with a failure rate based on a limited span of observation, such as 12 months.

Conclusion

The contraceptive failure rate, defined as the number of

[6] This rate of 14.4, based on the uninflated sample, differs barely from the rate of 15 published in Westoff, *et al*, *op. cit.*, p. 928, which was based on the inflated sample. For details of this inflation, see Whelpton, P. K. and Kiser, C. V.: The Sampling Plan, Selection, and the Representativeness of Couples in the Inflated Sample. In Social and Psychological Factors Affecting Fertility, V. 2, New York: Milbank Memorial Fund, 1950, pp. 190–200.

pregnancies per 100 years of contraceptive exposure, is sufficiently inflated by a short observation period so that this factor of observation length deserves high priority as a control in any comparative work. Other things equal, sensitivity to this factor is greatest when contraceptive effectiveness is low, and members of the sample vary greatly in their individual efficiencies. So sensitive is the failure rate to observation length, that it is virtually meaningless to compare a failure rate based on a limited period of observation, such as 12 months, with a failure rate based on entire interpregnancy intervals.

Appendix A. Formulas Used in Deriving Contraceptive Failure Rates for a Hypothetical Population

The density of fecundabilities "p," graphed in Figure 1, is defined by

$$
\begin{aligned}
f(p) &= 100\,p && 0 \leqq p \leqq .02 \\
&= 2.00/.98 - (2.00/.98)p && .02 \leqq p \leqq 1 \\
&= 0 && \text{otherwise.}
\end{aligned}
$$

In the absence of contraception the pregnancy rate per 100 years of exposure during an observation period of "t" months is

$$1200\,P(t)/E(t),$$

where $P(t)$ is the proportion of couples becoming pregnant during the observation period and $E(t)$ is the average number of exposure months per couple. To obtain $P(t)$, we set it equal to $1 - \overline{P}(t)$ and solve for $\overline{P}(t)$, the proportion of couples not becoming pregnant during the observation span. Evidently,

$$
\begin{aligned}
\overline{P}(t) &= \int_0^1 (1-p)^t f(p)\,dp \\
&= 100 \left\{ \left(\frac{1-.98^{t+1}}{t+1} \right) - \left(\frac{1-.98^{t+2}}{t+2} \right) \right\} \\
&\qquad + \frac{2.00}{.98} \left(\frac{.98^{t+2}}{t+2} \right)
\end{aligned}
$$

Regarding E(t), the average number of exposure months among couples sharing a fecundability of "p" is

$$p + 2 \cdot qp + 3 \cdot q^2p + \ldots + tq^{t-1}p + tq^t$$

$$= p(1 + 2q + 3q^2 + \ldots + tq^{t-1}) + tq^t$$

$$= p \left\{ \frac{1 + q + q^2 + \ldots + q^{t-1} - tq^t}{1-q} \right\} + tq^t$$

$$= 1 + q + q^2 + \ldots + q^{t-1}$$

$$= (1 - q^t)/(1 - q)$$

$$= \frac{1 - (1-p)^t}{p}.$$

Taking the full range of fecundabilities, we have then

$$E(t) = \int_0^1 \left(\frac{1 - (1-p)^t}{p} \right) f(p)\, dp$$

$$= 2 - \frac{100}{t+1}(1 - .98^{t+1}) + \frac{2.00}{.98}\left(\frac{.98^2}{2} + \frac{.98^3}{3} + \ldots + \frac{.98^{t+1}}{t+1} \right)$$

When contraception of "e" efficiency is practiced, fecundabilities of "p" are replaced by monthly risks of contraceptive failure of

$$p' = (1 - e)p.$$

The density of "p′" is defined by

$$g(p') = 100\, p'/(1-e)^2 \qquad 0 \leq p' \leq .02(1-e)$$

$$= \frac{2.00}{.98}\left(\frac{1}{1-e}\right) - \frac{2.00}{.98}\left(\frac{1}{1-e}\right)^2 p' \qquad .02(1-e) \leq p' \leq (1-e)$$

$$= 0 \qquad \text{otherwise.}$$

In accord with its definition as the number of pregnancies per 100 years of contraceptive exposure, the contraceptive failure rate is given by

$$1200\, P'(t)/E'(t).$$

Here $P'(t) = 1 - \overline{P}(t)$, with

$$\overline{P}'(t) = \int_0^{(1-e)} (1-p')^t g(p')dp'$$

$$= \frac{100}{(1-e)^2}\left\{\left(\frac{1-v^{t+1}}{t+1}\right) - \left(\frac{1-v^{t+2}}{t+2}\right)\right\}$$

$$+ \frac{2.00}{.98}\left(\frac{1}{1-e}\right)\left(\frac{v^{t+1}-e^{t+1}}{t+1}\right)$$

$$- \frac{2.00}{.98}\left(\frac{1}{1-e}\right)^2 \left\{\left(\frac{v^{t+1}-e^{t+1}}{t+1}\right) - \left(\frac{v^{t+2}-e^{t+2}}{t+2}\right)\right\}$$

when $v = .98 + .02e.$

$$E'(t) = \int_0^{1-e} \left(\frac{1-(1-p')^t}{p'}\right) g(p')dp'$$

$$= \frac{100}{(1-e)^2}\left\{(1-v) - \left(\frac{1-v^{t+1}}{t+1}\right)\right\}$$

$$- \left(\frac{2.00}{.98}\right)\left(\frac{1}{1-e}\right)^2 \left\{(v-e) - \left(\frac{v^{t+1}-e^{t+1}}{t+1}\right)\right\}$$

$$+ \left(\frac{2.00}{.98}\right)\left(\frac{1}{1-e}\right)\left\{(v-e) + \left(\frac{v^2-e^2}{2}\right) + \cdots + \left(\frac{v^t-e^t}{t}\right)\right\}$$

The above formulas are useful only when it is assumed that all couples are practicing contraception with the same efficiency "e." Alternatively one might assume that a proportion "m" of the couples are practicing perfect contraception while a proportion $1 - m$ are practicing contraception of zero efficiency. For the former group, $P'(t) = 0$ and $E'(t) = t$; for the latter, $P'(t) = P(t)$ and $E'(t) = E(t)$. Combined, the two subgroups have a contraceptive failure rate of

$$\frac{1200(1-m)P(t)}{mt + (1-m)E(t)}.$$