Emergency contraception prevents pregnancy after unprotected sexual intercourse. If widely used, it could substantially reduce the number of unintended pregnancies that occur each year in the United States.1 Emergency contraceptives available in the United States include regular oral contraceptive pills containing the hormones estrogen and progestin, less common birth control pills called minipills that contain progestin only, and the copper-T IUD.

The purpose of this article is to examine rigorously the efficacy of one method of emergency contraception: the Yuzpe method. The Yuzpe regimen has replaced the older postcoital therapy of high doses of the estrogen diethylstilbestrol, primarily because the Yuzpe regimen has fewer side effects. It involves taking two doses of pills containing a combination of estrogen and progestin, with each dose containing 100 mcg of ethinyl estradiol and 1.0 mg of norgestrel.2 One dose is taken within 72 hours after unprotected coitus, and the second is taken 12 hours later. The total dosage is therefore 200 mcg of ethinyl estradiol and 2.0 mg of norgestrel.3 In this article, we use the terms Yuzpe regimen, Yuzpe method and emergency contraceptive pills (ECPs) interchangeably.

Investigators have often evaluated the efficacy of the Yuzpe method by calculating its failure rate (pregnancies divided by the number of women treated), which we consider an inappropriate measure. In this article, we discuss why the failure rate is a poor measure of ECP efficacy and present estimates of a better measure, the effectiveness rate (the proportionate reduction, due to treatment, in the risk of pregnancy). We then discuss four methodological issues involved in measuring effectiveness—the appropriateness of pooling data across studies, the effects of loss to follow-up and protocol violations, and the accuracy of the expected number of pregnancies reported in the clinical trials we review.

**Data**

This research updates and extends an earlier analysis5 in which we challenged the conclusion of Silvestre, Bouali and Ulmann,3 based on five clinical trials, that the Yuzpe regimen may not be effective. Here, we review in depth all published results from clinical trials of the Yuzpe method in which the data needed to calculate effectiveness rates were collected and reported. Summary data on the 10 trials meeting this criterion5 are reported in Table 1.

The criteria by which subjects were selected differed somewhat across trials. All but one trial required women to have experienced an act of unprotected intercourse within the previous 72 hours; the sample in that trial was limited to women who had had unprotected intercourse within the previous 48 hours.6 Trials also differed in whether they included women with more than one unprotected act within the past 72 hours, more than one unprotected act since the last menses or more than one act (protected or unprotected) since the last menses. Some trials required that women abstain or use condoms until the results of the therapy could be verified. In all trials, most subjects were young and had never been pregnant (at least in the trials reporting this characteristic).

The fraction of treated women subsequently lost to follow-up varied from 0.0% to 22.7%. Several investigators reported that they were confident that women lost to follow-up were not pregnant because they would have heard of failures from colleagues7 or because they believed pregnant women would have returned to the same clinics for management of their pregnancy.8

**Failure Rates**

The fractions of women completing each clinical trial who became pregnant despite emergency postcoital therapy are reported in Table 2 (page 60). The failure rates for these trials, which range from 0.2% to 2.8%, and their associated confidence intervals were calculated under two assumptions. First, they assume that women lost to follow-up became pregnant at the same rate as those under observation. However, if women lost to follow-up were, as some investigators claimed, less likely to become pregnant than those followed, the rates will be too high. Second, they assume that all women treated in each trial had an equal probability of failure, although the risk of pregnancy depends on the timing of unprotected intercourse relative to the day of ovulation and thus varies widely.

We can test the assumption that all women treated had an equal probability of failure by pooling observations from all 10 studies. The resulting pregnancy rate is 1.5%, with a 95% exact confidence interval extending from 1.2% to 1.9%.9 That nine of the 10 studies fall outside the confidence interval suggests strongly that this assumption is not true. Indeed, based on Fisher’s exact test, the null hypothesis of equality of failure rates across studies would be decisively rejected ($\chi^2=25.8$ with $9\, df$, asymptotic $p$-value = .002 and Monte Carlo estimate of $p$-value = .001). These calculations suggest not only that pooling failure rates from different clinical trials is inappropriate, but that the failure rate is, in itself, a poor measure of ECP efficacy.

**Effectiveness Rates**

By basing our analysis of the Yuzpe regimen on the effectiveness rate—the proportionate reduction caused by use of ECPs in the probability of conception—

James Trussell is professor of economics and public affairs and associate dean, Woodrow Wilson School of Public and International Affairs, and director, Office of Population Research, Princeton University, Princeton, N. J. Charlotte Ellertson is program associate, The Population Council, New York. Felicia Stewart is deputy assistant secretary for population affairs, Department of Health and Human Services, Bethesda, Md. The authors are grateful to Anna Glasier, Ho Pak Chung, Robin Percival Smith, Beryl Tully, M. R. Van Santen, Anne Webb, Albert Yuzpe and Guglielmo Zuliani for providing information not included in their original publications and to Robin Foldesy, Paul Van Look and Anne Webb for constructive comments. The authors have no financial interest whatsoever in the commercial success or failure of the Yuzpe method.

---

*Ovral (two Ovral pills constitute one dose) is the only oral contraceptive marketed in the United States that contains the exact amounts of ethinyl estradiol and levonorgestrel in the Yuzpe regimen. Two Ovral pills contain 100 mcg of ethinyl estradiol and 0.5 mg of levonorgestrel. Other oral contraceptives that contain the same estrogen and progestin are often prescribed for emergency contraception. Four Nordette, Levlen or Lo/Ovral pills contain 120 mcg of ethinyl estradiol and 0.6 mg of levonorgestrel, and four yellow Triphasil or Triplevlen pills contain 120 mcg of ethinyl estradiol and 0.5 mg of levonorgestrel.*

†All statistical tests reported in this paper were performed with StatXact-Turbo software.

---

**Table 1:**

<table>
<thead>
<tr>
<th>Trial</th>
<th>Subjects</th>
<th>Successes</th>
<th>Failures</th>
<th>Effectiveness Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial 1</td>
<td>100</td>
<td>95</td>
<td>5</td>
<td>95.0</td>
</tr>
<tr>
<td>Trial 2</td>
<td>200</td>
<td>190</td>
<td>10</td>
<td>95.0</td>
</tr>
<tr>
<td>Trial 3</td>
<td>300</td>
<td>285</td>
<td>15</td>
<td>95.0</td>
</tr>
<tr>
<td>Trial 4</td>
<td>400</td>
<td>380</td>
<td>20</td>
<td>95.0</td>
</tr>
<tr>
<td>Trial 5</td>
<td>500</td>
<td>485</td>
<td>15</td>
<td>95.0</td>
</tr>
</tbody>
</table>

---

[1] Emergency contraception is a method of preventing pregnancy after unprotected sexual intercourse. Its effectiveness is assessed by calculating the proportionate reduction in the rate of pregnancy, which is a measure of the effectiveness of the contraceptive method.

[2] The Yuzpe regimen includes two doses of pills containing a combination of estrogen and progestin. Each dose contains 100 mcg of ethinyl estradiol and 1.0 mg of norgestrel.

[3] The Yuzpe regimen was first described by Albert Yuzpe and colleagues in 1975.

[4] The criteria by which subjects were selected differed somewhat across trials. All but one trial required women to have experienced an act of unprotected intercourse within the previous 72 hours; the sample in that trial was limited to women who had had unprotected intercourse within the previous 48 hours.

[5] Trials also differed in whether they included women with more than one unprotected act within the past 72 hours, more than one unprotected act since the last menses or more than one act (protected or unprotected) since the last menses.

[6] Some trials required that women abstain or use condoms until the results of the therapy could be verified.

[7] In all trials, most subjects were young and had never been pregnant.

[8] The failure rate is the proportionate reduction in the risk of pregnancy due to treatment. It is calculated as 1 - (success rate / failure rate).

[9] The 95% exact confidence interval is calculated using the Clopper-Pearson method.

---

Family Planning Perspectives
rather than the failure rate, we can avoid the problem of differences among women in the risk of pregnancy due to variations in the timing of unprotected intercourse. However, to compute the effectiveness rate, we need information not only on the observed number of pregnancies but also on the expected number of pregnancies. For this reason, we use data only from trials for which such information was provided in or could be derived from the published reports.

It is not possible to assess the effectiveness of the Yuzpe method from the data shown in Table 2 because they do not tell us how many women would have become pregnant in the absence of therapy. The expected number of pregnancies can be computed from external estimates of the risk of conception for each day of the cycle if information on the cycle day of intercourse relative to the cycle day of ovulation is available for subjects in the trial. The cycle day on which intercourse occurred relative to the expected cycle day of ovulation can be ascertained only among women with regular cycles.

As the data in Table 3 (page 61) show, estimates of the proportionate reduction in the risk of pregnancy resulting from the Yuzpe therapy range from 55.3% to 94.2%. If results from all 10 trials are pooled, the effectiveness of ECPs is 74.0%. If we treat the expected number of pregnancies as fixed, we can construct a 95% exact confidence interval of 68.2%–79.3% for the true effectiveness rate. By constructing a one-sided confidence interval, we can state with 95% confidence that the effectiveness rate is no lower than 69.1%.

These results change very little if we exclude the studies that might seem to differ from the others. For example, if we discard the two studies with the highest and lowest effectiveness rates—Yuzpe (1977) and Tully—the pooled effectiveness rate in the other eight studies is 74.0% (with a two-sided 95% confidence interval of 67.8%–79.7%, or a one-sided confidence interval with a lower bound of 68.6% if the expected number of pregnancies is regarded as fixed). If we discard the three studies—Yuzpe (1977), Bagshaw and Tully—in which only the number of women with unprotected midcycle intercourse (days –3 to +3) was reported, the effectiveness rate declines slightly to 72.9%
Effectiveness of the Yuzpe Regimen

Table 2. Calculation of failure rates (with 95% confidence intervals) in trials of the Yuzpe regimen of emergency contraception

<table>
<thead>
<tr>
<th>Lead author and year</th>
<th>No. followed</th>
<th>No. of pregnancies</th>
<th>Failure rate (%)</th>
<th>Incidence of nausea and vomiting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yuzpe, 1977</td>
<td>608</td>
<td>1</td>
<td>0.2 (0.0–0.9)</td>
<td>Two-thirds with nausea and 19% with vomiting (data from only one of four clinics, N unknown)</td>
</tr>
<tr>
<td>Yuzpe, 1982</td>
<td>634*</td>
<td>7</td>
<td>1.1 (0.4–2.3)</td>
<td>22% with nausea only, 4% with vomiting only, 25% with nausea and vomiting (N unknown)</td>
</tr>
<tr>
<td>Glasier, 1992</td>
<td>384</td>
<td>4</td>
<td>1.0 (0.3–2.6)</td>
<td>60% with nausea and 17% with vomiting on treatment day (N=346)</td>
</tr>
<tr>
<td>Bagshaw, 1988†</td>
<td>1,015</td>
<td>13</td>
<td>1.3 (0.7–2.2)</td>
<td>Antiemetics prescribed: 28% with nausea, 10% with vomiting (N=1,015)</td>
</tr>
<tr>
<td>Van Santen, 1985</td>
<td>571 ‡</td>
<td>5</td>
<td>0.9 (0.3–2.0)</td>
<td>Antiemetics offered for use prophylactically or only when symptoms developed: 38% with nausea, 18% with vomiting (N=434)</td>
</tr>
<tr>
<td>Percival-Smith, 1987</td>
<td>774</td>
<td>18</td>
<td>2.3 (1.4–3.7)</td>
<td>Antiemetics offered for use with second dose if nausea experienced with first dose; 30% with nausea only, 20% with nausea and vomiting (N=593)</td>
</tr>
<tr>
<td>Zuliani, 1990</td>
<td>407</td>
<td>9</td>
<td>2.2 (1.0–4.2)</td>
<td>55% with nausea, 17% with vomiting (N=407)</td>
</tr>
<tr>
<td>Ho, 1993</td>
<td>347 §</td>
<td>9</td>
<td>2.6 (1.2–4.9)</td>
<td>47% with nausea, 22% with vomiting (N=424)</td>
</tr>
<tr>
<td>Webb, 1992</td>
<td>191**</td>
<td>5</td>
<td>2.6 (0.9–6.0)</td>
<td>70% with nausea (26% mild, 25% moderate, 19% severe), 22% with vomiting (N=191)</td>
</tr>
<tr>
<td>Tully, 1983</td>
<td>395</td>
<td>11</td>
<td>2.8 (1.4–4.9)</td>
<td>30% with nausea, 12% with vomiting (N=296)</td>
</tr>
<tr>
<td>Total</td>
<td>5,326</td>
<td>82</td>
<td>1.5 (1.2–1.9)</td>
<td>Not applicable</td>
</tr>
</tbody>
</table>

*Excludes 13 women who had other acts of unprotected intercourse in that cycle before 72 hours prior to treatment, four of whom became pregnant. †Trial included both standard Yuzpe treatment (two doses 12 hours apart) and modified treatment (three doses 12 hours apart); results pertain to all women in trial because failure rates were not significantly different. (Failure rate for higher dosage in the study by Bagshaw to 29% in Yuzpe’s 1982 study; however, rates in seven of the 10 studies fell in the narrow range of 17–22% (see the last column of Table 2). In the studies in which women were routinely offered antiemetics (Van Santen and Percival-Smith), rates of vomiting were no lower than average (18% and 20%, respectively). However, if antiemetic medication was taken only after symptoms developed (an option in Van Santen’s study) or only with the second dose when nausea was experienced with the first dose (the instructions in Percival-Smith’s study), little reduction in the overall incidence of nausea or vomiting would be expected. In contrast, the one study (Bagshaw) in which all women were instructed to take antiemetics before each dose of ECPs had the lowest rates of vomiting (10%) and nausea (28%).

Methodological Issues

Although the use of effectiveness rates rather than failure rates to measure the efficacy of ECPs avoids the problem of differences among women in the risk of pregnancy created by variations in the timing of unprotected intercourse, other methodological questions remain. In this section, we discuss four methodological issues in the measurement of effectiveness of ECPs: the appropriateness of pooling observations, and the impact on our analysis of loss to follow-up, protocol violations and the use of external estimates of conception rates by cycle day to compute the expected number of pregnancies.

Pooling Observations

Is it legitimate to pool all observations from all 10 studies? The major reason for treating the samples as relatively homogeneous is that they share one important characteristic: All women had unprotected intercourse during only one menstrual cycle. Moreover, the timing of that act in relation to the estimated day of ovulation is in effect held constant by weighting observations by external estimates of the risk of conception specific to the day of the cycle on which unprotected intercourse occurred. The definition of a regular cycle and the algorithm for determining the expected day of ovulation were similar across studies. Finally, most women in the trials shared two key demographic characteristics: They were young and had never been pregnant.

Another reason for treating the samples as homogeneous is the similarity in rates of vomiting; vomiting that occurs before the steroids in ECPs are absorbed (within approximately three hours after treatment) can lower the efficacy of the treatment. None of the reports on the 10 studies included rates of vomiting within three hours following treatment. Overall rates of vomiting ranged from 10% in the study by Bagshaw to 29% in Yuzpe’s 1982 study; however, rates in seven of the 10 studies fell in the narrow range of 17–22% (see the last column of Table 2).
ies. Furthermore, information collected from women at the time of treatment about the cycle day of unprotected intercourse and about the length of prior cycles is unlikely to be reported without error.

**Loss to Follow-Up**

In three of the published reports—Yuzpe (1982), Percival-Smith and Tully—women lost to follow-up were included in computations of the expected number of pregnancies, although they could not by definition contribute to the observed number of pregnancies. If any women lost to follow-up became pregnant, then the observed number of pregnancies in these trials is too low and the effectiveness of ECPs is overestimated. Even in these studies, not all women lost to follow-up were included in computations of the expected number of pregnancies because the calculation is based only on women with regular cycles and known cycle day of unprotected intercourse.

If we assume that women lost to follow-up and women followed in these three studies were equally likely to have regular cycles and known cycle day of unprotected intercourse and had identical distributions of cycle days of unprotected intercourse, we can statistically remove women lost to follow-up simply by multiplying the expected number of pregnancies by the proportion of the treated sample that was followed. Then the expected number of pregnancies for all 10 studies would be reduced from 254.2 to 243.7, and the effectiveness rate would be decreased from 74.0% to 72.9%.

Only Yuzpe’s 1977 study reported no women lost to follow-up. The other six studies excluded women lost to follow-up when the expected number of pregnancies was computed. If women lost to follow-up were less likely to become pregnant than those followed, then the effectiveness of ECPs in those six studies would be underestimated. The overall magnitude and even the direction of any bias caused by the analytical treatment of women lost to follow-up is unknown.

**Protocol Violations**

In computations of effectiveness rates, the following women who violate protocol should be removed from the analysis, whether or not they become pregnant: those who have unprotected intercourse not only within the 72 hours before treatment but earlier in the same cycle, those who have unprotected intercourse after treatment but before the result of treatment is known, and those who fail to take both doses of ECPs. If only such women who become pregnant are removed, the effectiveness rate will be too high. However, because the observed number of pregnancies will be the overwhelming determinant of the effectiveness rate, failure to omit women who violate protocol but do not become pregnant will bias the result very little.

On balance, the effectiveness rates we reviewed are likely to be biased downward because of the failure to remove certain pregnancies from the analysis. In particular, observed pregnancies probably include some contributed by women who were already pregnant from an act of intercourse that occurred more than 72 hours before they began treatment and some contributed by those who became pregnant from another act of intercourse following therapy. In the remainder of this section, we discuss the effects on effec-

---

**Table 3. Calculation of effectiveness rates among women with regular cycles and known cycle day of unprotected intercourse in trials of the Yuzpe regimen of emergency contraception**

<table>
<thead>
<tr>
<th>Lead author and year</th>
<th>No. treated</th>
<th>Observed no. of pregnancies</th>
<th>Cycle days* of unprotected intercourse</th>
<th>Expected no. of pregnancies</th>
<th>Effectiveness rate (%)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yuzpe, 1977</td>
<td>152</td>
<td>1</td>
<td>–3 to +3</td>
<td>17.1 †</td>
<td>94.2</td>
<td>None</td>
</tr>
<tr>
<td>Yuzpe, 1982</td>
<td>451</td>
<td>5†</td>
<td>≤–8 to ≥+5</td>
<td>31.8 ‡</td>
<td>84.3</td>
<td>Calculation of expected pregnancies includes women lost to follow-up</td>
</tr>
<tr>
<td>Glasier, 1992</td>
<td>384</td>
<td>4</td>
<td>≤–8 to ≥+5</td>
<td>23 †</td>
<td>82.6</td>
<td>Random assignment to Yuzpe or mifepristone</td>
</tr>
<tr>
<td>Bagshaw, 1988**</td>
<td>345</td>
<td>8</td>
<td>–3 to +3</td>
<td>38.8†</td>
<td>79.4</td>
<td>None</td>
</tr>
<tr>
<td>Van Santen, 1985</td>
<td>461</td>
<td>5††</td>
<td>≤–8 to ≥+5</td>
<td>23.4§</td>
<td>78.6</td>
<td>Combined data from open study and comparative study (Yuzpe vs. high-dose estrogen); for approximately half those receiving the Yuzpe regimen, the progesterin was norgestrel; for the other half, levonorgestrel</td>
</tr>
<tr>
<td>Percival-Smith, 1987</td>
<td>622</td>
<td>12</td>
<td>≤–8 to ≥+5</td>
<td>40.2§</td>
<td>70.1</td>
<td>Calculation of expected pregnancies includes women lost to follow-up</td>
</tr>
<tr>
<td>Zuliani, 1990</td>
<td>407</td>
<td>9</td>
<td>≤–8 to ≥+5</td>
<td>28.7§</td>
<td>68.6</td>
<td>Random assignment to Yuzpe or danazol</td>
</tr>
<tr>
<td>Ho, 1993</td>
<td>341 §§</td>
<td>9</td>
<td>≤–8 to ≥+5</td>
<td>22.0§</td>
<td>59.1</td>
<td>Random assignment to Yuzpe or levonorgestrel-only regimen; levonorgestrel was Yuzpe progesterin</td>
</tr>
<tr>
<td>Webb, 1992</td>
<td>191 †</td>
<td>5</td>
<td>≤–8 to ≥+5</td>
<td>11.3 §</td>
<td>55.8</td>
<td>Random assignment to Yuzpe, mifepristone or danazol; levonorgestrel was Yuzpe progesterin</td>
</tr>
<tr>
<td>Tully, 1983</td>
<td>159</td>
<td>8</td>
<td>–3 to +3</td>
<td>17.9 †</td>
<td>55.3</td>
<td>Calculation of expected pregnancies includes women lost to follow-up</td>
</tr>
<tr>
<td>Total</td>
<td>3,534</td>
<td>66</td>
<td>na</td>
<td>254.2</td>
<td>74.0</td>
<td>None</td>
</tr>
</tbody>
</table>

*The range of days of intercourse recorded on either side of the estimated day of ovulation; estimated date of ovulation is usual cycle length minus 14 days in studies by Yuzpe (1977 and 1982), Glasier, Zuliani, Webb and Tully; average cycle length minus 14 days in study by Bagshaw; modal cycle length minus 14 days in study by Van Santen; mean cycle length minus 14 days in study by Ho; and actual cycle length minus 14 days for women who did not become pregnant and usual cycle length minus 14 days for women who did in study by Percival-Smith. †Not reported by the original investigators; estimated under the assumption that women are distributed uniformly across the reported range of days of intercourse using data from Dixon (see reference 12). ‡Observed pregnancies exclude four women who had other acts of unprotected intercourse in that cycle more than 72 hours prior to treatment. §Reported by the original investigators based on Dixon’s method. ††Trial contained both standard Yuzpe treatment (two doses 12 hours apart) and modified treatment (three doses 12 hours apart); results pertain to all women in trial since failure rates were not significantly different. (Failure rate for higher dosage was actually higher.) †‡Excludes one observed pregnancy because subject took only one of the two doses. §§Excludes seven women who had further acts of intercourse after treatment; six became pregnant. *Excludes five women who used emergency contraception twice in one cycle.
Effectiveness of the Yuzpe Regimen

tiveness rates of differential treatment of pregnancies and exposure to the risk of pregnancy among women who violate protocol, inclusion of women who were already pregnant when treated and inclusion of women who became pregnant after treatment. We close with a discussion of the importance to efficacy of the failure to take both doses of the treatment.

• Differential treatment of pregnancies and exposure. We discovered two instances in which women who violated protocol and became pregnant might have been treated differently in the analysis from those who did not become pregnant. In Van Santen’s study, a woman who took only one dose of ECP’s and who became pregnant was excluded from all analyses. If there were other women who did not complete the treatment but did not become pregnant, the study’s reported effectiveness rate would be biased upward.

In Yuzpe’s 1982 study, four women who became pregnant despite receiving ECP’s had not only had unprotected intercourse within 72 hours before treatment but had also had unprotected intercourse earlier in the treatment cycle. These women were excluded from the analysis of effectiveness rates; if other women who violated protocol in the same way but did not become pregnant remained in the analysis, the effectiveness rate would be biased upward. However, the impact of removing all women who violated protocol from calculations of the expected number of pregnancies would lower the effectiveness rate in that study very little—from 84.3% to no less than 83.1%. *

• Women pregnant at treatment. Webb’s study excluded one of six observed pregnancies because it did not occur during the act of unprotected intercourse for which treatment was given. The woman was pregnant despite having had a period at the expected time within the three weeks prior to treatment; pregnancy at the time of treatment was confirmed by a blood test. All of the other studies would have included this pregnancy even though it was not attributable to ECP failure.

If, as in this study, only 83% of observed pregnancies in all studies should be attributed to failure of emergency contraception, the number of observed pregnancies attributable to ECP failure in the other studies would be reduced from 61.0 to 50.8 and the total number of pregnancies attributable to ECP failure would fall from 66.0 to 55.8. The pooled effectiveness rate for the 10 studies would rise from 74.0% to 78.0%.

Percival-Smith abandoned an exclusion criterion used in an earlier study (the 1982 Yuzpe study) because of concerns about study bias. In the earlier study, subjects were excluded if they had had unprotected intercourse more than 72 hours before treatment. This exclusion criterion was abandoned because subjects subsequently admitted that they had lied to obtain treatment. Although it seems appropriate to treat such women but exclude them from the analysis, none of the investigators in the other studies reported any distinction between clinical treatment protocol and protocol for an efficacy study. Therefore, some subjects in the other studies may also have concealed earlier acts of unprotected intercourse to obtain treatment. Such behavior would bias effectiveness rates downward.

• Women becoming pregnant after treatment. In the study by Ho, 77 women were excluded from the analysis because they violated the protocol for the trial by having intercourse after treatment but before the results were known; six of them became pregnant. The failure rates of these women (78%) and the remaining 347 women (2.6%) are significantly different ($\chi^2=4.55$ with 1 df, p-value = .037 and mid-p-value = .023 based on Fisher’s exact test). Altogether, six of the 15 observed pregnancies were excluded.

If, as in this study, only 60% of observed pregnancies should be attributed to failure of emergency contraception, then the number of observed pregnancies attributable to ECP failure in the other studies would be reduced from 570 to 342 and the total number of pregnancies attributable to ECP failure would fall from 66.0 to 43.2. The pooled effectiveness rate for the 10 studies would rise from 74.0% to 81.5%.

If we adjust both for those pregnant at treatment (based on information from Webb’s study) and for those who had further acts of intercourse following treatment (using information from Ho’s study), the pooled effectiveness rate for the 10 studies would rise from 74.0% to 84.5%.

• Failure to take both doses. In his original pilot study of postcoital contraception, Yuzpe used a single dose that was equivalent to one of the two doses in the current therapy, and women were treated up to five days following unprotected intercourse. 13 There were three pregnancies in 148 treatment cycles (a failure rate of 2.0%). He later conducted another trial to evaluate the regimen currently in use, in which women take a second dose 12 hours after the first dose and treatment can be initiated only within 72 hours following unprotected intercourse. 14 In that trial, only one pregnancy occurred in 608 treatment cycles (see Table 2), resulting in a failure rate (0.2%) one-tenth that in the pilot study.

Although the failure rate in the second trial is significantly lower than that in the first study ($\chi^2=6.1$ with 1 df, p = .025 and mid-p = .013, based on Fisher’s exact test), the failure rate in the pilot study is lower than the rates shown in Table 2 for five of the 10 studies. Given that the effectiveness rate of the regimen in the pilot study is not available and therefore cannot be compared with the effectiveness rates of the current regimen in subsequent trials, we cannot evaluate the importance of the second dose to the efficacy of the current regimen.

Expected Number of Pregnancies

Accurate estimates of the number of pregnancies that would be expected without emergency contraceptive treatment are needed to calculate the treatment’s effectiveness. If estimated conception probabilities by cycle day are too low, then the expected number of pregnancies will also be too low, and the apparent effectiveness of ECP’s will be underestimated.

The expected numbers of pregnancies shown in Table 3 are based on conception probabilities by cycle day calculated by Dixon and colleagues12 by averaging the results from studies by Schwartz and colleagues,13 Barrett and Marshall,14 and Vollman.15 To determine if the expected numbers of pregnancies shown in the table are accurate, we examine in this section each of the three original studies as well as the methodology Dixon used in his calculations.

Schwartz. This study examined conception rate by cycle day in a series of 821 cycles (contributed by 529 women), each...
with a single donor insemination. The women were presumed to be fertile; all had a male partner with azoospermia or oligospermia. The overall success rate was 12%, where success was defined as clinical confirmation of pregnancy after at least 21 days of hyperthermia (elevated temperature caused by pregnancy). Among the 631 cycles in which ovulation could be detected from temperature charts, the pregnancy rate was 13%.

Because the women were trying to conceive, insemination days were clustered at midcycle. No pregnancies were observed from 71 inseminations before cycle day –4 or after cycle day +2 (with cycle day 0 the last day of hypothermia). Estimated conception rates were 8%, 20%, 13%, 21%, 15%, 11% and 9% on cycle days –4, –3, –2, –1, 0, +1 and +2, respectively. Thus, the average seven-day midcycle conception rate would be 13.9%. Estimated conception probabilities from this study are unbiased downward, because artificial insemination with frozen semen is not as likely as coitus to result in pregnancy.16

• Barrett and Marshall. These investigators analyzed data contributed by 241 British married couples with demonstrated fertility who were regulating conception by the basal body temperature (BBT) method. The couples recorded each morning BBT and all acts of intercourse. The analysis is based on all available data reported by those seeking to avoid pregnancy as well as by those seeking to become pregnant.

Because couples generally reported more than one act of intercourse per cycle, a more sophisticated statistical procedure than simple division (appropriately employed by Schwartz) was needed to estimate conception probabilities by cycle day. For each cycle, the probability of no conception is the product of the probabilities of not conceiving on each day of the cycle when intercourse occurred. Barrett and Marshall therefore used logistic regression to compute maximum likelihood estimates of the probability of conception specific to each cycle day. They expressed the probability of conception \( P \) in a given cycle as: \( P = 1 - \prod (1 - p_i)^{x_i} \), where \( p_i \) is the probability of conception on day \( i \) and \( x_i \) equals 1 if coitus occurred on day \( i \) and 0 otherwise (Model 1). The resulting estimated conception rates were 13%, 20%, 17%, 30%, 14% and 7% on cycle days –4, –3, –2, –1, 0 and +1, respectively,* and the average six-day midcycle conception rate was 16.8%.

This model assumes that sperm entering the female reproductive tract on different days have statistically independent probabilities of fertilizing the ovum and producing a pregnancy. However, in Barrett and Marshall’s study, conceptions were defined as pregnancies lasting at least two months from the last menstrual period (or six weeks from ovulation); hence, fetal losses before the second month of pregnancy were not included in the calculation of conception rates. Estimated conception rates in this study are therefore biased downward.

The model can easily be generalized to allow for nonindependence of risks of conception, where lack of independence arises from fertilized ova that do not implant or fetal losses before pregnancy is observed and from nonfertilizable ova.17 In the revisied model (Model 2), the probability of conception is expressed as: \( P = P_o P_f P_v \), where \( P_o \), \( P_f \) and \( P_v \) are, respectively, probabilities that a fertilizable ovum is produced, that it is fertilized and that the conceptus survives to the point of observation. Because ovulation is verified by BBT in the data from Barrett and Marshall’s study, \( P_o \) can be taken to be 1.0,† although strictly speaking, \( P_f \) and \( P_v \) are not separately identifiable and can be expressed jointly as a factor \( k = P_f P_v \). Thus, a final model (Model 3) expresses the probability of conception per cycle as \( P = k[1 - \prod (1 - p_i)^{x_i}] \).

Using data‡ from Barrett and Marshall’s study, Schwartz, MacDonald and Heuchel estimated \( k \) to be 0.52,18 their estimates of \( k \) were similar in pattern to—but on average 12% higher than—Barrett and Marshall’s earlier estimates of \( k \). Thus, Models 1 and 3 yield similar estimates of conception when intercourse occurs only once during the fecund period; estimates from the two models differ greatly when multiple acts of intercourse occur.§ When Royston19 used the same data and the third model but did not constrain conception probabilities to be zero outside the window from cycle day –4 to day +1, the estimated conception probabilities were, on average, 26% higher than those estimated by Barrett and Marshall.

• Vollman. The description of the procedure for collecting information for Vollman’s study is not very clear. The data were from 74 couples who “for some time” had been successfully using periodic abstinence to space pregnancies; they then “agreed to have intercourse only once in the cycle for the next planned pregnancy.” Vollman not only reports the cycle days on which the 43 who had thus far conceived had done so, but also reports the cycle days during which 1,395 acts of intercourse had occurred in 561 cycles during which no pregnancy was observed. This information clearly shows that many of these 1,395 acts occurred in cycles with more than one act, apparently because experience contributed while the couples were still avoiding pregnancy was also included.

One problem arises because the study group had already successfully been using periodic abstinence for some time. Thus, these couples may have been selected for low fecundity. The more serious problem is that the retrospective experience of these couples (prior to their agreement to have intercourse only once per cycle) is included, but the experience of other couples who had already had an unintentional pregnancy is excluded, thereby compounding the selection bias.

Even if there were no selection bias, the estimates of conception probabilities would be biased downward because Vollman simply divided the number of conceptions resulting from intercourse on day i by the total number of acts of intercourse on day i to obtain the conception rate on day i. But because there was more than one act of intercourse per cycle, maximum likelihood techniques should have been used to estimate the parameters \( k, P_o \) and \( P_v \) in Model 3.**

The estimates of conception probabilities specific to each cycle day can be summarized as:

\[ P_i = P_o P_f P_v \]

*Cycle days were identified differently by Barrett and Marshall; the first day of raised temperature was designated –1 and the last day of hypothermia was designated +1, with no day designated 0. We redefined the days to correspond to the convention used by Schwartz, in which the last day of hypothermia (cycle day –1 in Barrett and Marshall) is designated as day 0.

†The analysis is limited to cycles in which ovulation occurred, or at least appeared to occur; otherwise, cycle days cannot be identified because the reference day is day 0 (the last day of hypothermia).

‡The data used both by Schwartz and by Royston contain 2,192 ovulatory cycles and 103 conceptions, 294 more cycles and four more conceptions than in the data available to Barrett and Marshall at the time of their analysis.

§Let \( P_i \) denote the probability of conception when intercourse occurs on both day \( i \) and day \( j \) and \( P_j \) denote the probabilities when intercourse occurs only on day \( i \) or day \( j \). In Model 1, \( P = P_o P_f P_v \) and \( P_i = P_o P_f P_v + P_o P_j + P_j P_v - P_o P_f P_v \). In Model 3, \( P = k P_o P_f P_v \) and \( P_i = k P_o P_f P_v + k P_o P_j + k P_j P_v - k P_o P_f P_v \). Correcting for the factor \( k \) ensures that nonindependence does not produce absurdly high predicted probabilities of conception when intercourse is very frequent.

**If all conception cycles contained only one act of intercourse, whereas sterile cycles contained more than one act, then simple division is indeed appropriate in Model 1. For example, suppose that a conception resulted from intercourse on day \( i \) in cycle 1, a conception resulted from intercourse on day \( j \) in cycle 2, and no conception resulted from intercourse on day \( i \) and day \( j \) in cycle 3. The Vollman estimates of \( P_i \) and \( P_j \) are 0.5. The likelihood for Model 1 would be \( P_i P_j \), the maximum likelihood estimates of \( P_i \) and \( P_j \) are also 0.5. However, the likelihood in Model 3 would be \( k P_i k P_j \). Correcting for the factor \( k \) gives \( k P_i k P_j \). The maximum likelihood estimates of \( k, P_i \) and \( P_j \) are 0.667.

---

Volume 28, Number 2, March / April 1996

63
marized as follows. Schwartz’s estimates are definitely too low because they are based on women artificially inseminated with frozen sperm. Vollman’s estimates are unambiguously biased downward both because of selection bias and because of use of an inappropriate statistical model. Barrett and Marshall’s estimates are clearly biased downward because they are based on a statistical model that incorrectly assumes independence.

Even the revised estimates by Royston might be too low because couples who use periodic abstinence over a long period of time are selected for low fecundity: More fecund couples are more likely to be unsuccessful users and consequently switch to other contraceptive methods, leaving only the less fecund couples as continuing users of periodic abstinence. On the other hand, the estimates might be too high because all the women in the study had borne at least one child. These two biases would tend to offset one another.

One final methodological issue merits attention. To reflect the uncertainty in the estimated day of ovulation in women receiving emergency contraception, Dixon assumed that if a woman’s estimated cycle day of unprotected intercourse was day i, the true cycle day could have been day i±1, day i±2, or day i±3. He estimated the probability P of conception on cycle day i as the weighted average \( P = 0.05p_{i-3} + 0.1p_{i-2} + 0.2p_{i-1} + 0.3p_i + 0.2p_{i+1} + 0.1p_{i+2} + 0.05p_{i+3} \), where the values \( p \) are the averages of the conception probabilities in the three original studies (Vollman; Schwartz; and Barrett and Marshall). The original studies show a much shorter period of nonzero conception probabilities (eight vs. 14 days) and higher average daily probabilities of conception during that period (12% vs. 7%).

Dixon provides no justification for the use of the weighted averages; it is by no means obvious that this procedure is an improvement over the unweighted values. Because the unprotected coital acts of the women in the trials of ECPs described above tend to be concentrated at midcycle,\(^{21}\) use of the Dixon table would result in underestimates of the expected number of pregnancies even if the estimates from the original studies were unbiased.

To gauge the impact of various biases identified thus far, let us consider the study by Tully, which had the lowest effectiveness rate (see Table 3). If the averages of the conception probabilities in the three original studies are used rather than Dixon’s estimates, the expected number of pregnancies rises from 17.9 to 19.0 and the effectiveness rate rises from 55.3% to 57.9%. If only the estimates from Schwartz and from Barrett and Marshall are used, then the expected number of pregnancies rises to 20.1 and the effectiveness rate to 60.2%. If Royston’s estimates based on the data from Barrett and Marshall are used, the expected number of pregnancies rises to 24.3 and the effectiveness rate to 67.1%.

One concern with these revised estimates is that estimated conception risks by cycle day are based only on the cycles in which ovulation was presumed to occur (based on BBT shift). Because the expected number of pregnancies is based on all cycles in the 10 trials contributed by women who had a history of regular cycles, and some of these cycles may not have been ovulatory, the effectiveness of ECP’s may be too high.

However, a study by Hilgers and Bailey found that all of the cycles judged by serum progesterone assay to be anovulatory nevertheless had classical biphasic BBT patterns characteristic of ovulatory cycles.\(^ {22}\) Hence, the data from Schwartz and from Barrett and Marshall probably also contain anovulatory cycles. Moreover, women in the ECP trials were overwhelmingly concentrated in the age group (younger than age 30) with the highest fecundity;\(^ {23}\) in contrast, more than half of the women in the Barrett and Marshall study were aged 30 or older.

In summary, we conclude that the conception probabilities proposed by Dixon are too low. Therefore, the estimated expected number of pregnancies based on these conception rates is too low and the effectiveness of ECP’s is understated.

**Conclusion**

Estimates of ECP efficacy can reasonably be compared only when they are based on expected numbers of pregnancies computed by matching the cycle day of unprotected intercourse relative to ovulation with conception rates specific to that cycle day. Results computed in this way are available for only 10 published studies. The effectiveness rates (the proportionate reductions in the expected number of pregnancies) range from a low of 55.3% to a high of 94.2%. The weighted average of the effectiveness rates in all 10 studies obtained by pooling all observations is 74.0%.

The true effectiveness rate of the Yuzpe regimen is probably higher—at least 75% and perhaps higher than 80%—because the observed number of pregnancies in these studies is likely to be too high and the expected number of pregnancies is probably too low. While not as effective as the ongoing use of any regular method of contraception, ECP use substantially reduces the risk of pregnancy after unprotected sexual intercourse. Even a 75% effectiveness rate would ensure that ECPs are very cost-effective.\(^ {24}\)

One finding of clinical relevance emerges from our review: The limited evidence currently available suggests that prophylactic use of antinausea medicine can reduce the incidence of nausea and vomiting. Reduced vomiting might increase the effectiveness of ECP treatment. Although antinausea medicine is not routinely offered with ECP treatment in the United States, using it might be preferable to taking ECPs with food, the advice currently offered by many providers of emergency contraception. This advice is based solely on anecdotal evidence; in addition, taking food with ECPs theoretically could lower the maximum concentration of hormones in the blood and thereby reduce the effectiveness of the therapy.\(^ {25}\)

Almost all women can safely use ECPs.\(^ {26}\) Treatment may not be appropriate for those who have an active migraine with marked neurological symptoms or crescendo migraine.\(^ {27}\) Combined oral contraceptives are usually not prescribed to women with a history of stroke or blood clots in the legs or lungs. Nevertheless, in the only study of clotting factors following treatment with ECPs, there were no discernable changes among women with normal results on thrombophilia screening tests.\(^ {28}\)

Lack of public awareness concerning emergency contraception is primarily due to inaction by the pharmaceutical companies that make birth control pills and IUDs. Manufacturers cannot market or advertise these products for postcoital use until they seek and gain formal approval from the U.S. Food and Drug Administration for this specific purpose; thus far, no company has applied. Without this commercial promotion, it is not surprising that physicians infrequently prescribe emergency contraception and fail to provide information about it to women during routine visits.\(^ {29}\) The end result is that very few women know that emergency contraception is effective, accessible and safe.\(^ {30}\)

**References**

1. J. Trussell et al., “Emergency Contraceptive Pills

corefined on page 87
Effectiveness of Yuzpe Regimen... (continued from page 64)


18. Ibid.
