Levonorgestrel emergency contraception: a joint analysis of effectiveness and mechanism of action

Rafael T. Mikolajczyk, M.D.,a and Joseph B. Stanford, M.D., M.S.P.H.b

a Department of Public Health Medicine, School of Public Health, University of Bielefeld, Bielefeld, Germany; and b Department of Family and Preventive Medicine, University of Utah, Salt Lake City, Utah

Objective: To model the effectiveness that can be obtained if levonorgestrel-only emergency contraception (EC) acts only through disrupting ovulation, in relation to other effects that may occur before or after fertilization and accounting for delays in administration.

Design: We modeled follicular growth as a function of follicular size, using known day-specific probabilities of conception and known disruption of ovulation by levonorgestrel-only EC, to estimate the expected effectiveness of EC.

Setting: Combined data from multiple clinical studies.

Patient(s): Simulation models.

Intervention(s): Disruption of ovulation.

Main Outcome Measure(s): Effectiveness in the form of proportion of pregnancies prevented.

Result(s): With disruption of ovulation alone, the potential effectiveness of levonorgestrel EC ranged from 49% (no delay) to 8% (72-hour delay). With complete inhibition of fertilization before the day of ovulation, the potential effectiveness of levonorgestrel EC ranged from 90% (no delay) to 16% (72-hour delay).

Conclusion(s): The gap between effectiveness of levonorgestrel EC estimated from clinical studies and what can be attributed to disruption of ovulation may be explained by overestimation of actual effectiveness and supplementary mechanisms of action, including postfertilization effects. Additional data with follicular ultrasound and precise measures of delay between intercourse and EC administration would yield greater insight into effectiveness and mechanisms of action. (Fertil Steril 2007;–:–. ©2007 by American Society for Reproductive Medicine.)

Key Words: Emergency contraception, ovulation, follicular growth, effectiveness, postfertilization effects, levonorgestrel

Despite a growing body of literature on levonorgestrel (LNG) emergency contraception (EC), uncertainty remains regarding its effectiveness to prevent pregnancy. Clinical effectiveness trials have found that LNG EC is more effective than the Yuzpe EC regimen (1, 2). The actual effectiveness of LNG EC has been estimated between 58% and 95% under different delays in administration (1–6). In some of the studies, a decrease of effectiveness with delay was noted, in others not. Additional studies of effectiveness were analyzed in a Cochrane review by Cheng et al. (7), but the review did not look into the impact of delay, and the original reports were published in Chinese language.

Statistical procedures that have been used generally overestimate effectiveness (8–10). In the absence of trials with a placebo control group, the effectiveness of LNG EC was estimated to be at least 50%, assuming zero effectiveness of the second treatment arm consisting of women receiving the Yuzpe regimen (9).

Similarly, uncertainty remains regarding the mechanisms of action of LNG EC. Animal studies have found effects on ovulatory function but no significant effects after ovulation (11, 12). In humans, physiologic studies have consistently found that LNG EC inhibits, delays, or modifies hormone profiles accompanying ovulation (13–17). In our study, inhibition of ovulation, delay of more than 5 days, and modification of hormone profiles will all be referred to collectively as “disruption of ovulation.” There is limited evidence that the subsequent luteal phase may be shorter or have lower progesterone levels (13, 15). Evidence is mixed with regard to effects on the endometrium (13, 15). Effects on human spermatozoa were not demonstrated in vitro for LNG concentrations as used in EC (18–21).

These two issues of effectiveness and mechanism of action are directly related, but the relationship has not been assessed systematically (22, 23). We propose a model that estimates the effectiveness related to disruption of ovulation by LNG-EG, shows what level of clinical effectiveness is possible to achieve based on the prevention of fertilization, and illustrates the effectiveness with or without postfertilization mechanisms of action. We also propose additional research that may yield more insights into the relationship between effectiveness and mechanisms of action.
MATERIALS AND METHODS

Modeling the Potential Contribution of Ovulation Disturbance

An approach to estimating the contribution of ovulatory-function disruption to the effectiveness of EC was originally proposed by Trussell and Raymond (24). They estimated the effectiveness of Yuzpe method of EC under the assumption that its sole mechanism of action is disruption of ovulation and compared it with the effectiveness of EC observed in clinical trials. The analysis was conducted before clinical data on ovulatory-function disruption were available and was therefore based on crude assumptions, assigning preovulatory and postovulatory status based on the calendar day of the menstrual cycle.

A more precise adaptation of this approach is to incorporate data from ultrasound on follicular diameter. These data can be used to calculate the reduction in fecundity due to the disruption of ovulatory function and to compare it with fecundity in the absence of EC. The estimate of effectiveness thus obtained can be compared with estimates from clinical trials. Within this framework, the effect of an increasing delay between intercourse and use of EC also can be studied.

Follicular Growth

The analysis of the impact of the disruption of ovulatory function (as assessed by ultrasound) requires data on the normal development of ovarian follicles. A number of studies have reported mean diameters of follicles for days preceding ovulation, but provide insufficient data to estimate the distribution of follicular size for each day (25–30). We therefore use a detailed database of follicular measurements. The database was described in detail elsewhere (31, 32). In brief, there were 533 ultrasound measurements of follicular size from 327 cycles in 107 women who were not taking hormonal medication and had no history that would suggest subfertility. The participants were recruited from eight European centers in Aix-en-Provence, Dijon, Lyon, Milan, Verona, Duesseldorf, Liège, and Madrid.

The inclusion criteria were healthy menstruating women, aged 18 to 45 years inclusive, with previous menstrual cycle lengths of 24 to 34 days inclusive, and experience in natural family planning methods (basal body temperature and signs of cervical mucus). Women with frequent anovulatory cycles, on any hormonal treatment, with disturbances of follicular development or history of infertility were excluded. A total of 107 women were recruited whose mean age (± SD) was 33.1 ± 5.9 years, and 50% of the sample was aged between 29 and 39 years. More than 60% had proven fertility with a median parity of 2 (range: 0 to 6) in the whole sample. They were observed for up to four cycles (most women for three cycles). The study protocol was approved the institutional review board (the Comité Consultatif de Protection des Personnes dans Recherche Biomedicale de Lyon, France).

From these data, we derived an equation that gives a distribution of expected follicular size for a given day in relation to ovulation. We used random effects models to account for the differences among different women and cycles of the same woman (33, 34). For direct comparability with the published data from a study of follicular development and LNG EC by Croxatto et al. (15), we used the same measure of follicular size: the mean diameter of two planes. We applied the convention that the first day when follicle disappeared is the day of ovulation (0) and the day of largest follicular diameter is day −1 in relation to ovulation.

Daily Fecundity

Data on daily fecundity (the probability of clinical pregnancy for coitus on a given day relative to ovulation) have been extensively discussed in the context of the estimation of the overall effectiveness of EC (8, 10, 35–40); however, such data are available for few groups of women only. One of these data sets is from a cohort of 221 women planning pregnancy, the North Carolina Early Pregnancy Study (41, 42). Another widely used data set comes from 241 British users of basal body temperature for family planning (43, 44).

Using a Bayesian approach that permitted control of uncertainty in the biomarkers, Dunson et al. (45) showed that the underlying fecundity window observed in the North Carolina study was similar to that of the British users of temperature for natural family planning. Estimation of the effectiveness of EC in clinical trials of LNG EC has been based on the North Carolina data alone, or on both of these data sets combined (“pooled recognized conceptions”), according to the procedures developed by Trussell et al. (39, 40). We use both sets of published estimates in our analysis.

Disruption of Ovulatory Function

For the simulation of the disruption of ovulatory function by LNG EC, we use the data reported by Croxatto et al. (15), as summarized in Table 1. It is not entirely clear by how much fecundity is reduced in cycles in which ovulation is present, but the hormonal profiles are disturbed. We assumed that fertilization is not possible in such cycles, a conservative assumption that maximizes the potential effects of ovulation disruption.

Simulations of EC Effectiveness

We simulated random samples of 10,000 women presenting for EC for a single cycle each, and we calculated the number of “expected” pregnancies for each simulated cohort of women using both sets of the daily fecundity data. We assumed that women “presented” for EC treatment with equal probability on days −10 to +5 around the day of ovulation (day 0).

For each of the women within the fecundity window, we used the follicular growth equation to estimate a follicular diameter, which in turn was used to estimate the disruption of ovulation by LNG EC based on the data from the Croxatto study (Table 1). We assumed that effects observed for
12–14 mm, 15–17 mm, and ≥18 mm groups reported by Croxatto et al. (15) apply to follicles of size up to 11.51–14.5 mm, 14.51–17.5 mm, and ≥17.51 mm, respectively. When LNG EC was administered on a day when follicular size was below 11.5 mm, we assumed that there was zero probability of pregnancy. These conservative assumptions maximized the possible effects of LNG EC to disrupt ovulation and prevent fertilization. With this information, we estimated the “observed” pregnancies within the simulated cohorts.

According to accepted procedure, we calculated effectiveness of EC as 1 – (observed/expected) pregnancies. This yielded the effectiveness that would be expected according to the effect of LNG EC on ovulation for each simulation. We then compared this effectiveness with what might be obtained if LNG EC had additional effects that also prevent clinically identified pregnancy.

Complete Prevention of Fertilization

Levonorgestrel EC might have additional effects to prevent fertilization besides disrupting ovulation. For example, it might delay ovulation less than 5 days, inhibit sperm migration, or reduce sperm capability for fertilization (although this potential effect has not been supported in physiologic studies). Therefore, we ran simulations in which the probability of observed pregnancy was zero whenever LNG EC was administered at any time before the day of ovulation.

Delay between Index Intercourse and the Administration of EC

In practice, there is often a delay between the index intercourse and the administration of EC. We simulated the effects of delay in simulations with the daily fecundity selected from the day of index intercourse and the reduction in fecundity selected according to the day of EC administration. Separate simulations were performed for delays of 0, 24, 72, 96, and 120 hours. In these simulations, we postulated that LNG EC cannot work by disrupting ovulation on the day of ovulation itself, which is the day of follicular rupture.

RESULTS

We found that the rate of follicular growth is reasonably constant across cycles (1.5 mm per day) but that follicles in different cycles grow for different durations and rupture at different final sizes (with a mean [± SD] follicular size before rupture of 21.7 ± 2.8 mm). The main source of variation was between cycles, and accounting for the clustering of cycles within women did not improve the fit of the model. The resulting equation for the growth trajectory in an individual cycle is:

\[
\text{Follicular diameter (mm)} = 21.6747 - 1.5032 \\
\times (\text{Days to ovulation}) + e
\]

where \(e\) is normally distributed random error with mean = 0 and SD = 2.755.

Accordingly, in our simulations, we assigned follicular sizes based on this equation to each day relative to ovulation. Because the index intercourse for each woman was equally distributed among the days of the −10 to + 5 window, there were about 625 women on each day of the fecund window, resulting in 669 “expected” pregnancies for the pooled daily fecundities based on the British and North Carolina samples (39, 40).

We calculated a reduction in fecundity due to disruption of ovulation by LNG EC by dividing the proportion of cycles with ovulation after EC by the proportion of placebo cycles with ovulation, reported in the study by Croxatto et al. (15), for each stratum of follicular size. We assumed that

<table>
<thead>
<tr>
<th>Classification of follicular diameters</th>
<th>12–14 mm</th>
<th>15–17 mm</th>
<th>≥18 mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean diameter (± SD)</td>
<td>12.6 (0.6)</td>
<td>15.7 (0.5)</td>
<td>18.5 (0.7)</td>
</tr>
<tr>
<td>LNG group</td>
<td>N = 18</td>
<td>N = 22</td>
<td>N = 17</td>
</tr>
<tr>
<td>No rupture within 5 days</td>
<td>15</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Ovulatory dysfunction</td>
<td>2</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>Placebo</td>
<td>N = 18</td>
<td>N = 22</td>
<td>N = 16</td>
</tr>
<tr>
<td>No rupture within 5 days</td>
<td>10</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Ovulatory dysfunction</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Reduction factor for fecundity(^a)</td>
<td>0.143</td>
<td>0.167</td>
<td>0.605</td>
</tr>
</tbody>
</table>

\(^a\) Proportion of cycles with ovulation after administration of levonorgestrel EC, divided by proportion of placebo cycles with ovulation. This is based on the assumptions that follicles that do not rupture within 5 days result in no fertilization, and that cycles with ovulatory dysfunction result in complete inhibition of fertilization. This factor is used for the simulations in Figure 1.
follicles that do not rupture within 5 days cannot result in fertilization from the index intercourse because of the length of the fertile window (41), and that cycles with changed hormonal profiles accompanying ovulation likewise cannot result in fertilization. The resulting estimates of reduction in fecundity are shown in the bottom row of Table 1.

We then simulated the reductions in pregnancy (the “observed” pregnancies of the simulated EC trials) based on the disruption of ovulation observed with LNG EC (see Table 1) for delays of 0, 24, 48, 72, and 120 hours between index coitus and administration of EC. From this, we calculated effectiveness due to disruption of ovulation; the results are shown in Table 2. In Figure 1, selected results are displayed in relation to hypothetical additional effects (other than on ovulation) that would be necessary to reach a particular level of clinical effectiveness. If EC were taken immediately after intercourse, disruption of ovulation could result in almost 49% effectiveness; any higher observed effectiveness would require other mechanisms of action to operate in addition. With a delay of 24 hours or more, other mechanisms of action are required to obtain an effectiveness of 50% or more.

Similarly, we simulated effectiveness based on the premise that LNG EC always prevents fertilization when administered anytime before ovulation for delays of 0, 24, 72, 48, 96, and 120 hours between index coitus and administration of EC. We calculated the maximum theoretically possible effectiveness due to prevention of fertilization by any means; the results are shown in Table 2. In Figure 2, selected results are displayed in relation to hypothetical additional postfertilization effects that would be necessary to reach a particular level of clinical effectiveness. If EC were taken immediately after intercourse, prefertilization effects could result in up to 90% effectiveness. With a delay of 24 hours or more, the maximum observed effectiveness from all prefertilization mechanisms of action would be 60%, and a delay of 72 hours or more would require postfertilization mechanisms to be operative to obtain an effectiveness of 50% or more.

### Table 2

<table>
<thead>
<tr>
<th>Delay between coitus and EC (hours)</th>
<th>Effectiveness due to simulated effect (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Disruption of ovulation</td>
</tr>
<tr>
<td>0</td>
<td>49</td>
</tr>
<tr>
<td>24</td>
<td>31</td>
</tr>
<tr>
<td>48</td>
<td>16</td>
</tr>
<tr>
<td>72</td>
<td>8</td>
</tr>
<tr>
<td>96</td>
<td>2</td>
</tr>
<tr>
<td>120</td>
<td>0</td>
</tr>
</tbody>
</table>

Under either set of assumptions, if LNG EC is taken 96 hours or more after intercourse, the maximum possible ability of LNG EC to intervene before fertilization is essentially zero. This is because, with the increasing delay between the index intercourse and EC the use, the proportion of women who have already ovulated (and therefore are likely to have already undergone fertilization) increases.

### Figure 1

Relationship between effectiveness of levonorgestrel emergency contraception (EC) that may be estimated from a clinical trial and the contribution of mechanisms other than ovulation disturbance (based on the levels of ovulation disturbance described in Table 1), by different delays between index intercourse and EC administration. The intercept with the x axis represents the effectiveness of levonorgestrel EC if the only action is ovulation disturbance for a given delay between index intercourse and EC administration.
After intercourse, it was 58% effective (1). Similarly, in it was 85% effective; and if it was administered 49 to 72 hours after intercourse, it was 95% effective within 24 hours after intercourse. If EC was administered within 24 hours after intercourse, it was 95% effective (1). Recently, the effectiveness of LNG EC in 976 women was estimated to be 79% effective when administered up to 72 hours after intercourse, whereas our simulations suggest that the effectiveness with a full 72-hour delay for all women would be 64%, and 60% effective when administered between 72 and 120 hours after intercourse (46).

However, in another study of 1021 women receiving LNG EC, the overall effectiveness was estimated to be 64%, and there was no evidence for a trend of reduced effectiveness with increasing delay between intercourse and EC administration (an unexplainable result if disruption of ovulation plays a major role) (4). In a Chinese study of 2018 women receiving two different formulas of LNG EC, the proportion of observed pregnancies was higher when EC was administered after a 72-hour delay, but the difference was not significant, and it was not presented in the form of effectiveness (6). These contradictory findings could be explained by different composition in the subsamples of women who were receiving LNG with delay, but the necessary information to evaluate this is not available in the published data.

There are two possible explanations for these discrepancies. The first is that the effectiveness rates estimated in effectiveness trials of LNG EC could be inflated; the actual effectiveness of LNG EC may be substantially lower than what has been estimated from clinical trials. Recent analyses of the procedures used to estimate effectiveness of LNG EC have suggested that the overall effectiveness has indeed been overestimated by approximately 10% (absolute difference) (9, 10, 38).

However, this relatively modest amount of overestimation could not account for the discrepancies noted. For example, even if estimates of effectiveness from clinical trials are adjusted downward by 25% (absolute), they would still suggest that LNG EC would be 33% to 45% effective up to 72 hours after intercourse, whereas our simulations suggest that the effectiveness with a full 72-hour delay for all women would be 8% without any additional effects beyond ovulation disturbance, or 16% with complete inhibition of fertilization before ovulation but no postfertilization effects. Furthermore, although there has been overestimation of the effectiveness of LNG EC, we have recently confirmed in a simulation study that the methods commonly used to estimate the effectiveness of EC provide reasonably consistent results across a reasonable range of underlying assumptions (8). Hence, it seems improbable that overestimation from clinical trials alone can account for the discrepancies noted.

The other possibility is that mechanisms other than disturbance of ovulation contribute to the reduction of clinical pregnancy, including mechanisms acting after fertilization. Figures 1 and 2 quantitate the absolute contribution needed from these other mechanisms to reach a defined level of clinical effectiveness.

Recent studies in Cebus monkeys and in rats found that LNG EC strongly inhibited ovulation and fertilization in these species (11, 12). However, results from nonhuman studies do not necessarily extrapolate to humans. For example, the mechanisms of action in animals for the intrauterine device have been shown to differ markedly from mechanisms of action in humans (47, 48).

**DISCUSSION**

Our analysis shows a direct relationship between potential mechanisms of action of LNG EC and the observed clinical effectiveness. If ovulation disruption is taken to be the only significant mechanism of action, the total effectiveness could be not much higher than 50% if EC is administered immediately after intercourse; with delays, it would be substantially less than 50%. Similarly, if LNG EC given before the day of ovulation were able to completely prevent fertilization by any mechanism, if there were no postfertilization effects, and if there were a 24-hour delay in administration, the highest possible effectiveness would be about 60%. The sharp reduction in effectiveness of LNG EC with increasing delay between the time of intercourse and of LNG EC administration is attributable to the increasing proportion of women who are within the fecundity window at the time of intercourse and subsequently have already ovulated by the time of EC use.

These results stand in sharp contrast to the higher effectiveness rates reported in clinical trials. In one large trial, the effectiveness of LNG EC in 976 women was estimated to be 85% within 72 hours after intercourse. If EC was administered within 24 hours after intercourse, it was 95% effective; if it was administered 25 to 48 hours after intercourse, it was 85% effective; and if it was administered 49 to 72 hours after intercourse, it was 58% effective. Similarly, in a study of 1356 women receiving LNG EC, it was estimated to be 79% effective when administered up to 72 hours after intercourse, and 58% effective when administered between 72 and 120 hours after intercourse (46).
In our simulations, we have conservatively overestimated the effect of ovulation disruption to an unknown degree in several ways. First, we have assumed that any disruption of ovulation results in complete prevention of fertilization, which may not be true. Studies that suggest reduced potential for fertilization in cycles with blunted LH peaks are based on observations from unmedicated cycles of subfertile women, which may display different behavior than cycles of women of normal fecundity receiving EC (49, 50). Second, when LNG EC was administered on a day below a follicular size of 11.5 mm we assumed that it resulted in zero probability of pregnancy, which is also probably not true in some fraction of cases.

Third, in simulating the effect of LNG EC to disrupt ovulation, we applied the estimates of effect from Croxatto et al. (15). Although they reported estimates for three ranges of follicular diameters, the distribution of diameters from women in each group of their study was clustered toward a lower diameter within the specified range (see Table 1). In our simulations, there was no such clustering of women within these ranges; we effectively applied results from a group of women with lower mean follicular diameters to simulated women with slightly higher follicular diameters. Because LNG EC has a stronger effect on ovulation at lower follicular diameters, this procedure systematically overestimated the effect of LNG EC on the disruption of ovulation to an unknown extent. Finally, the sample used for developing the follicular growth model included a sample of women among whom 60% had proven fecundity; this is likely to be higher than in a sample of women presenting for EC. It is also possible that follicular development could differ among the samples; however, in the model of follicular growth, the main source of variability was among cycles and not among women, suggesting that a strong systematic difference is not likely among groups of women with regard to follicular growth.

We were not able to estimate confidence limits for the estimates. A meta-analytic approach would be necessary to combine the estimates of effectiveness of LNG EC or estimates of daily fecundity. However, in the case of effectiveness studies, the published data are not sufficient for calculation of confidence limits; in the case of daily fecundity, no standard methods exist for meta-analysis. Furthermore, we acknowledge the limitation of the relatively small size of the study used for the evaluation of the effects of LNG EC on ovulation.

Two lines of clinical research would provide improved estimates of actual effectiveness of LNG EC and of the relative contribution of ovulation disruption and other mechanisms of action to its effectiveness. The first would involve designing studies to collect precise data on the delay between index coitus and EC administration and its effect on pregnancy rates and estimated effectiveness of EC. If the magnitude of the effect decreases over the time delay, this evidence would support preovulatory effects; to the extent that it does not, there is evidence of additional postovulatory effects.

The second and more valuable approach would be to focus on follicular ultrasound. This could include two components: [1] reference data on follicular diameter and probability of conception within the same group of women; [2] data from a clinical trial of LNG EC that would include a follicular ultrasound at the time of EC administration. Either or ideally both of these components would allow for more direct modeling of the relationship between follicular size, the time of ovulation, and the effects of EC. Additional studies specifically addressing the disruption of ovulation by LNG EC would also be helpful to confirm our model. Finally, further physiologic studies would be welcome to clarify what alternative mechanisms may be in place, including modulation of endometrial physiology and velocity of oviductal transport (51, 52).

Even with the limitations of available data, our model clearly shows that the questions of actual clinical effectiveness of LNG EC and its mechanisms of action are directly connected. Either the actual clinical effectiveness is far lower than has been estimated in the literature to date or mechanisms of action other than ovulation disruption must be contributing to the clinical effectiveness. In our opinion, both explanations are likely to be contributing to the observed discrepancy between the level of effectiveness that can be attributed to preovulatory effects and the effectiveness reported in clinical trials.

Acknowledgments: The authors thank Rene Ecochard, M.D., for supplying data on follicular size by ultrasound and time to ovulation.

REFERENCES