Br. J. clin. Pharmac. (1992), **33**, 293–298

A placebo-controlled study of effects of oral progesterone on performance and mood

ELLEN W. FREEMAN[1,2], LISA WEINSTOCK[1], KARL RICKELS[1,2], STEVEN J. SONDHEIMER[1] & CHRISTOS COUTIFARIS[1]
Departments of [1]Obstetrics/Gynecology and [2] Psychiatry, University of Pennsylvania, Philadelphia, PA 19104, USA

1 Oral micronized progesterone (P) is proposed for the treatment of certain endocrine gynaecological disorders. To examine the effects of P on performance and mood, a randomized, placebo-controlled study of 24 healthy females ages 18–24 years on low-dose oral contraceptives was conducted.

2 Subjects were admitted to the Clinical Research Center on four occasions and received single doses of oral P (300, 600, 1200 mg) or placebo. Blood sampling, psychometric tests and mood scales were administered at baseline and at hourly intervals for 6 h.

3 P doses produced significant dose-related but highly variable increases in plasma P concentrations. Fatigue increased with P doses, although few subjects were objectively drowsy. Very high peak plasma P concentrations, achieved by some subjects at the 1200 mg dose, were associated with decreased information processing and verbal memory function as well as fatigue.

4 We conclude that oral P can safely be prescribed at higher than previously-reported doses, based on evidence of transient behavioural effects only at the highest doses in some subjects who achieved high plasma P concentrations.

Keywords oral progesterone dose-response plasma concentrations behavioural effects mood effects

Introduction

Oral micronized progesterone (P) is suggested for treatment of a number of gynaecological disorders, including luteal phase defect, support of the luteal phase in treatment of infertility and postmenopausal hormonal replacement therapy. Progesterone has been widely used in treatment of the premenstrual syndrome (PMS) although recent controlled studies of luteal phase suppository administration found no effectiveness compared with placebo (Freeman *et al.*, 1990) and no physiological rationale for progesterone therapy for PMS in the late luteal phase (Schmidt *et al.*, 1991).

Although P is commonly prescribed and studies of oral administration of 200 mg P are reported (Arafat *et al.*, 1988; Chakmakjian *et al.*, 1987; Hargrove *et al.*, 1989; Maxson & Hargrove, 1986), there is widespread concern that oral P may be unacceptably sedating, based on clinical reports of this response (Arafat *et al.*, 1988; Tapanacinen *et al.*, 1989) and the reported sedative-hypnotic actions of one of its metabolites, 3α-hydroxy-5α-dihydroprogesterone (3α-OH-DHP) (Majewska *et al.*, 1986). Although animal data show that progesterone metabolites are barbiturate-like modulators of the GABA receptor, and that these characteristics may provide a mechanism for the anaesthetic and hypnotic action of the natural and synthetic steroids (Majewska *et al.*, 1986; Mendelson *et al.*, 1987), these processes have not been demonstrated in humans, and the dose-response relationship of oral P and behavioural effects is not well-understood.

We conducted a placebo-controlled, double-blind study of oral micronized P in a sample of healthy young women to determine whether acute doses, ranging from 300 to 1200 mg, produced performance deficits. Results reported here are the plasma progesterone concentrations achieved at each dose and their associations with the psychological measures of performance and mood.

Methods

Subjects

Twenty-four healthy females were randomized to receive three single dosages of oral micronized progesterone (300, 600, 1200 mg) and placebo, each dose at a minimum interval of 1 week. Prior to treatment, the subjects

Correspondence: Dr Ellen W. Freeman, Department of Obstetrics/Gynecology, Hospital of the University of Pennsylvania, 3400 Spruce Street, Philadelphia, PA 19104, USA (215) 622-3329

signed consent for study participation approved by the University Review Board and were placed on low-dose oral contraceptives (1 mg norethindrone, 35 mg ethinyl-oestradiol) to create a standardized hormonal milieu approximating the luteal phase of the menstrual cycle.

The female subjects, ages 18–25 years, were recruited from the University community. They reported regular menstrual cycles within normal range of 23–35 days. None had evidence of pill-induced amenorrhea. General good health was confirmed by physical examination, complete blood count and blood chemistries. Subjects had no complaints of premenstrual symptoms or mood disorders. They took no psychotropic medications for 6 months and no other medications other than oral contraceptives for 2 weeks prior to starting the study. They had no caffeine or alcohol for 36 h prior to each study visit. Subjects were paid for their participation.

Study procedure

At the initial screening visit, subjects had a physical examination, blood drawn for laboratory tests and were acquainted with the psychometric tests. The testing sessions were scheduled without regard to phase in the oral contraceptive-induced standardized cycle at minimum intervals of 1 week. Subjects were randomly assigned to the double-blind treatment sequences, with the order of the four doses determined by a Latin-square design. At each 6 h testing session, subjects arrived at the clinical research centre at 08.00 h after overnight fast. An i.v. catheter was placed in the non-dominant arm, and blood samples were drawn to measure baseline progesterone concentrations. The psychometric test battery and mood scales were administered to assess baseline status on the psychometric tests and mood measures. Subjects then ingested four gelatin capsules, each containing either placebo or 300 mg micronized P. Blood sampling was repeated at 15 min intervals for the first hour, and then hourly at 2, 3 and 4 h. Psychological testing was administered at each hour for the first 4 h of the study. Administration time for each test battery was approximately 20 min. Subjects were served lunch at hour 4, a final psychological test battery was administered at hour 6 for all subjects. Blood sampling was repeated at hour 6 for the last 12 subjects only. Subjects were not permitted to sleep during the 6 h session.

Blood was allowed to clot and serum was separated and stored at −70° C. Serum P concentrations were measured in duplicate by standard radioimmunoassay kits (Diagnostic Products, Los Angeles, CA). The inter- and intra-assay variabilities were 6.8% and 3.05% respectively.

Capsules containing 300 mg P were prepared and individually weighed by the pharmacist. The contents from two capsules were dissolved in ethanol and P was measured in quadruplicate. These eight assays indicated that the mean (±s.d.) was 306.28 (±39.76) mg per capsule.

Assessments

The behavioural tests used to assess performance changes included the digit symbol substitution test (DSST), symbol copying (SCT), immediate word recall, delayed word recall and critical flicker fusion frequency (CFF). These tests are reported in the literature, have been used in previous studies to evaluate benzodiazepine treatment (Lucki *et al.*, 1986, 1987), and were previously reported to be sensitive to the effects of benzodiazepines in normal volunteers (Kleinknecht & Donaldson, 1975; Peturrson *et al.*, 1983; Wittenborn, 1979).

The DSST evaluates information processing, visual scanning and motor speed. The test form consists of a page with a grid containing the numbers 0 to 9. Under each row of numbers is attached a row of empty boxes. Subjects insert a symbol in each empty box to correspond with the number above the box, as indicated in the reference grid at the top of the page. Subjects were given 90 s to fill in as many symbols as possible. Different forms and codes were used each time the test was administered to control for learning effects. The score is the number of correct symbols completed.

The symbol copying test requires subjects to draw symbols immediately below a sample. The same nine symbols that were used in the DSST were presented in random order on the test form. Different forms were used each time the test was given to control for learning effects. The score is the number of symbols replicated in 90 s.

The tests of immediate and delayed recall evaluate verbal memory function. Different versions of the test were used for the repeated measurements, with word lists matched for their frequency of use. Subjects were read a list of 16 common English nouns at an approximate rate of 1 word/2 s. They then recorded as many words as they could recall immediately after the list was read for the test of immediate recall. Twenty minutes later, subjects again recorded as many words as they could recall for the delayed recall test. Scores are the number of words recalled correctly at each timepoint.

The CFF is a relatively uncontaminated behavioural measure of central nervous system functioning that has been reported as sensitive to a variety of psychoactive medications (Smith & Misiak, 1976). The test uses the Leeds Tester, a small screen that displays four flashing red diodes. The flashing frequency of the diodes changes continuously in increasing and decreasing directions. The screen is placed approximately 1 m from the subject's eye level and trials of both ascending and descending frequency are administered. Subjects press a stop button when the flashing diode light appears to be continuous, thereby determining the critical fusion frequency measured in Hz. The score in the present report is the average Hz of two increasing and two decreasing flicker tests.

Mood changes were assessed with the Profile of Mood State (POMS) (McNair & Lorr, 1964), a self-report questionnaire assessing 6 factor-analytically derived mood dimensions: tension-anxiety, depression, anger-hostility, vigour, fatigue and confusion. This measure is intended to assess transient mood states and has been shown to be a sensitive measure of the effects of experimental manipulations on normal subjects. Subjects rated 65 mood items for their severity at the time of rating with values ranging from 0 (not at all) to 4 (extremely). Scores are the sums of item scores for each mood dimension.

Data analysis

The data were analyzed using standard Statistical Analysis System (SAS) programs (SAS Institute, Inc.). There were no differences in the baseline scores compared between the four visits. Change scores (the same-day baseline score subtracted from the test score at each timepoint) were used to control for baseline variability in the peformance and mood measures. The nonparametric Friedman analysis of variance of ranked data with repeated measures (Siegel, 1956) was used to compare each outcome measure between the three P doses and placebo. Comparisons were made at each hour (with no statistically significant results) and at the time of peak P at each dose. Multivariate tests and contrast transformations for significant MANOVA results are reported. To examine the effects of P in subjects who achieved very high plasma concentrations, the outcome measures were compared between P and placebo in high and low plasma P groups using Wilcoxon signed rank tests. In all analyses, results of $P < 0.05$ with conservative two-tailed interpretation were considered statistically significant.

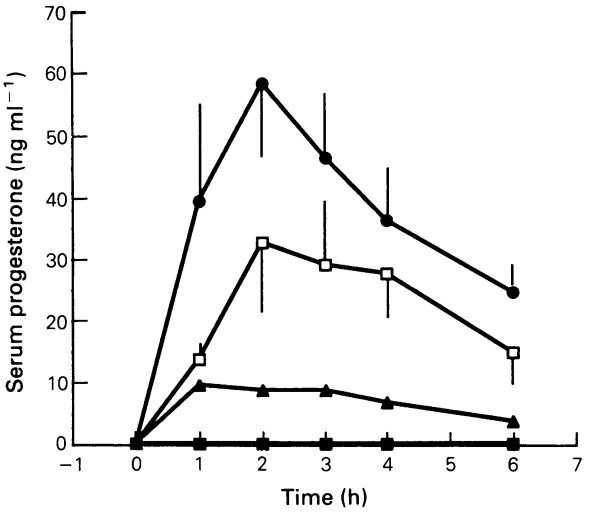

Figure 1 Mean (±s.e. mean) of serum P concentrations by dose and times, $n = 24$ (two outliers omitted at 1200 mg). P value for repeated measures ANOVA on ranked data is $P = 0.0001$ at each hour after baseline. ■ 0 mg, ▲ 300 mg, □ 600 mg, ● 1200 mg progesterone.

Results

Progesterone concentrations

All subjects absorbed the oral progesterone and evidenced large increases in plasma P concentrations. The means (s.e. mean) for each dose are depicted in Figure 1. There was large variability between subjects with the peak plasma concentrations ranging from 0.15 to 0.79 ng ml^{-1} on the placebo dose; 1.86 to 46.54 ng ml^{-1} at dose 300 mg; 5.20 to 340.00 ng ml^{-1} at dose 600 mg; and 10.00 to 2360.00 ng ml^{-1} at dose 1200 mg. Two subjects were considered outliers at the 1200 mg dose, with peak P concentrations of 2360.00 ng ml^{-1} and 960.00 ng ml^{-1}. The mean peak of plasma P was at 1 h with a dose of 300 mg and at 2 h with doses of 600 and 1200 mg, again with large variability. Subjects reached peak plasma P levels at each of the 6 h studied, and the same subjects achieved peak plasma P concentrations at different times across the four treatments. Analysis of plasma P concentrations (repeated measures ANOVA of ranked data, Table 1) showed highly significant increases in plasma P at each hour measured after baseline ($P < 0.0001$). All *post hoc* pairwise comparisons between doses and timepoints were statistically significant.

Performance and mood measures at peak P

Results for the performance and mood measures showed a dose-related response for the fatigue ($P < 0.007$) and vigour ($P < 0.04$) factors, with greater fatigue at each P dose and less vigour at the 1200 mg P dose compared with placebo (Table 2). The verbal memory test of immediate recall showed deficits with P doses ($P < 0.003$), but only the 600 mg differed significantly from placebo.

Effects at high plasma P concentrations

Since the MANOVA results for the test scores showed that apparent effects of P were at the highest doses only, and since there was large variability in the outcome measures as indicated by the standard errors (Table 2), the 1200 mg sample was split at the median of peak plasma P, with results shown in Table 3. Term pregnancy

Table 1 Plasma concentrations (ng ml^{-1}) by dose and time, $n = 24$

	Dose (mg)				
Time (h)	*Placebo*	*300*	*600*	*1200*[2]	P *value*[1]
0	0.35 (0.03)	0.36 (0.03)	0.38 (0.03)	0.40 (0.05)	NS
1	0.28 (0.03)	9.84 (2.80)	13.83 (3.40)	39.69 (16.80)	0.0001
2	0.24 (0.02)	9.01 (2.20)	32.81 (13.90)	58.54 (13.90)	0.0001
3	0.22 (0.03)	8.98 (1.70)	29.16 (10.10)	46.69 (10.40)	0.0001
4	0.21 (0.02)	6.97 (1.10)	27.80 (8.90)	36.45 (8.20)	0.0001
6	0.25 (0.04)	3.95 (0.52)	15.07 (5.40)	24.78 (4.60)	0.0001

Values are means (± s.e. mean). Two outliers are omitted at dose 1200 mg. At 6 h, $n = 12$.

[1]Friedman repeated measures MANOVA based on ranked data. All pairwise comparisons at each hour are significant.

[2]Including two outliers at dose 1200, the means (s.e. mean) from 2–6 h are 106.28 (36.4), 153.72 (100.9), 133.83 (69.3), and 92.65 (46.5) respectively.

Table 2 Psychological measures at peak progesterone concentrations, $n = 24$, mean (s.e. mean) of change scores[1] at peak P for each dose

Measure	*Baseline (Placebo)*	*Placebo*	*300 mg*	*600 mg*	*1200 mg*	P *value*[2]
Performance tests						
Immediate recall	10.5(0.61)	−0.67(0.48)	−1.39(0.58)	−2.71(0.31)*	−1.71(0.63)	0.003
Delayed recall	9.0(0.73)	−2.95(0.56)	−2.13(0.63)	−3.79(0.43)	−3.88(0.72)	0.17
Digit symbol substitution	78.8(2.10)	−1.50(1.25)	−0.17(1.17)	−1.38(1.43)	−1.42(2.24)	0.77
Symbol copying	155.0(4.80)	2.08(4.60)	−1.09(3.02)	0.38(2.54)	−4.00(3.59)	0.16
Critical flicker frequency	27.9(0.60)	2.64(3.39)	2.78(3.57)	2.04(4.94)	5.25(6.76)	0.13
Mood scores (POMS)						
Tension-Anxiety	8.2(0.67)	−1.75(0.50)	−0.82(0.39)	−0.92(0.53)	−1.65(0.76)	0.20
Depression	1.1(0.43)	−0.38(0.28)	−0.14(0.10)	−0.67(0.33)	−1.13(0.43)	0.60
Anger-Hostility	0.9(0.35)	−0.71(0.25)	0.59(0.40)	−0.58(0.47)	−0.48(0.27)	0.17
Vigour	13.3(1.50)	0.92(1.15)	−1.73(0.89)	−0.54(0.84)	−2.22(0.83)*	0.04
Fatigue	5.8(1.10)	−2.17(0.99)	2.05(0.79)*	1.00(0.72)*	3.78(1.38)*	0.007
Confusion	6.4(0.35)	−0.54(0.20)	−0.41(0.33)	0.08(0.28)	0.78(0.46)	0.21

[1]Change = Peak P time minus baseline. Negative scores, except for fatigue and confusion, indicate direction toward sedative effect.
[2]Friedman analysis of variance ranked data with repeated measures.
*Indicates significant difference from placebo in *post hoc* pairwise tests.

Table 3 Psychological measures in the high peak P group at dose 1200 mg compared between peak P and placebo (mean (s.e. mean) change[1])

	High P group, n = *12*			*Low P group,* n = *12*		
Measure	*Placebo*[2]	*Peak plasma P*	P *value*[3]	*Placebo*[2]	*Peak plasma P*	P *value*
Performance tests[5]						
Delayed recall	−1.83(0.98)	−5.33(1.06)	0.009	−2.83(0.71)	−2.42 (0.81)	0.60
Digit symbol substitution	−0.83(1.90)	−7.25(2.63)	0.03	−0.17(1.63)	4.42 (2.81)	0.22
Symbol copying	1.75(5.08)	−12.50(4.57)	0.10	4.42(2.70)	4.50 (4.46)	0.99
Mood scores (POMS)[5]						
Vigour	−0.08(1.50)	−3.58(1.19)	0.09	−3.33(1.56)	−0.73 (1.02)	0.09
Fatigue	0.83(1.91)	5.58(1.73)	0.003	−0.33(0.90)	1.82 (2.12)	0.47
Peak P concentrations[5]	−0.07(0.02)	387.50(193.76)[4]	0.0001	−0.15(0.04)	24.70 (3.35)	0.0005

[1]Change = Peak P time minus baseline.
[2]The change for placebo is calculated at the same hour as the peak change for 1200 mg P.
[3]Wilcoxon signed rank *t*-test. Results shown to $P < 0.10$ in high P group.
[4]Omitting two outliers, the mean (s.e. mean) is 133.04 (31.71).
[5]Comparisons between the high and low group (Wilcoxon Rank Sum) are: Delayed recall ($P = 0.058$), Digit symbol substitution ($P = 0.008$), Symbol copy ($P = 0.03$), Vigour ($P = 0.07$), Fatigue ($P = 0.06$) and Peak P ($P = 0.0001$).

levels of P (> 100 ng ml^{-1}) were achieved by 8 of the 24 subjects at the 1200 mg dose. Fatigue at peak plasma P following the 1200 mg dose was clearly greater compared with placebo at the same hour ($P = 0.003$). Delayed recall ($P = 0.009$) and digit symbol substitution ($P = 0.03$) were decreased following the P dose compared with placebo, indicating diminished information processing and verbal memory function in this group with very high plasma P concentrations. There were trend relationships indicating slower symbol copying ($P = 0.10$) and less vigour ($P = 0.09$). While the number of comparisons increased the likelihood of significant tests, the results were consistent with the MANOVA results (Table 2) and indicate that the most clear-cut effects of P were present in the small group with the highest plasma P concentrations. These results were not found in the other half of the sample with lower plasma P concentrations at the 1200 mg dose.

Discussion and conclusions

The results show that oral P, administered at higher doses than previously reported in the literature, was associated with significant dose-related increases in plasma P concentrations, although there was large variability as observed in previous studies. The plasma P concentrations at the 300 mg dose were consistent with results reported for 200 mg doses (Arafat *et al.*, 1988; Chakmakjian & Zachariah, 1987; Hargrove *et al.*, 1989; Maxson & Hargrove, 1986; Sitruk-Ware *et al.*, 1987), while the P concentrations at the 1200 mg dose reached levels of term pregnancy in one-third of the subjects. The variability in plasma P concentrations was also consistent with previous reports of peak P values ranging from 4.6 to 44.9 ng ml^{-1} following a 200 mg P dose (Sitruk-Ware *et al.*, 1987). The reasons for the large range in plasma P concentrations might be due to differ-

ences in the rate of absorption or the extent of absorption of P into fatty tissues (Sitruk-Ware *et al.*, 1987). To our knowledge this is the only explanation for the variability of P that has been corroborated by other reports.

The mood measures showed that the POMS 'fatigue' factor increased significantly with P doses. However, it is important to emphasize that clinically, only 4 of the 24 subjects appeared objectively drowsy at 1 or 2 h after medication intake. Two of these women proved to be the two outliers with extremely high P concentrations at dose 1200 mg and showed consistent changes in the direction of sedation on their psychological test scores at the time of drowsiness. Of the two remaining subjects who appeared objectively drowsy, one proved to be on placebo and showed no remarkable changes in the psychological test scores. The other had received the 600 mg dose, reached a peak plasma P concentration of 94 ng ml^{-1} at the time of the observed drowsiness and had small changes in the psychological test scores that showed less vigour, more confusion and poorer memory compared with her scores on the other doses. It is also important to emphasize that the study subjects were young healthy women who had little or no mood distress and therefore could not be expected to show decreases in negative affect such as anxiety, since such affect did not exist at baseline.

In the performance measures, some deficit in verbal memory function was observed, but was not consistently related to increasing P doses. However, those subjects who reached very high peak plasma P concentrations at the 1200 mg P dose showed statistically significant decreases in several performance tests including delayed recall and digit symbol substitution tests, indicating an association of high plasma P concentrations with some deficit in these measures.

In spite of a previous report of one subject who experienced hypnotic effects at an oral P dose of 400 mg (Arafat *et al.*, 1988), we did not find profound effects of oral P associated with performance or mood at the administered doses for reasons that are not obvious. Possibly acute doses were insufficient to produce sedative effects, particularly since effects were observed at the highest dose in the subgroup with the highest plasma P concentrations. Possibly these healthy young volunteers were not very sensitive to the effects of P. Familiarity with the psychometric tests could possibly have mitigated the effects of P, but the Latin Square design was implemented to control for this problem, and tests of each of the 12 measures by visit order showed no statistically significant learning effects. Perhaps differences in circulating levels of gonadal steroids between these young women and the post-menopausal women who were the subjects in previous reports of hypnotic (Arafat *et al.*, 1988) and drowsy states (Tapanacinen *et al.*, 1989) following oral P doses are associated with responses to P. Possibly the oral contraceptives, given to standardize the hormonal milieu, altered the metabolism of P or the metabolites of the oral contraceptives had some unknown competitive effect with the metabolites of the oral P doses. Another study, where test occasions are standardized in the menstrual cycle, could support or refute these results based on a standardized hormonal milieu.

We administered P doses of 1200 mg and placebo to four males who produced concentrations of P and psychological test results consistent with those of the women (unpublished data). Only one male had any indication of fatigue and poorer delayed recall, an incidence similar to the observed effects of P administration in the study sample.

Other important questions remain for further study. For example, the bioavailability of the P metabolite 3α-OH-DHP having sedative-hypnotic actions is not known, although metabolite studies are now in progress. Since the earlier report of sedation with P occurred in a post-menopausal female (Arafat *et al.*, 1988), the question remains as to whether the response to oral P differs in cycling women compared with post-menopausal women with lower levels of circulating oestrogen and progesterone. Another important question is whether these results from single dose administration would be altered with longer durations of P administration, particularly if high P concentrations were sustained. We previously found a threshold effect with a 400 mg P dose in the luteal phase (Myers *et al.*, 1987), i.e. that treatment duration and P increment were inversely related, but additional data are needed.

The present study indicates that orally administered micronized progesterone can be prescribed at higher than previously reported doses to young cycling women without debilitating effects on performance or mood. Evidence of transient fatigue with these P doses of 300 to 1200 mg indicates that patients should be observed closely at the onset of treatment, particularly at P doses of 600 mg or higher.

The authors acknowledge with appreciation the contributions of James Goin, Ph.D. as statistical consultant, Irwin Lucki, Ph.D., as consultant for the psychometric testing, and the technical assistance of Barbara Altimari, Peggy Bulley, Beatriz Garcia-España and Richard Hall.

This study was supported by NIH grants HD-18633 and 5M01 RR000040 (Clinical Research Center), a fellowship from NIMH (#5-T32-MH-18902) and the Benjamin and Mary Siddons Measey Foundation to Lisa Weinstock.

References

Arafat, E. S., Hargrove, J. T., Maxson, W. S., Desiderio, D. M., Wentz, A. C. & Anderson, R. N. (1988). Sedative and hypnotic effects of oral administration of micronized progesterone may be mediated through its metabolites. *Am. J. Obstet. Gynecol.*, **159**, 1203–1209.

Chakmakjian, Z. H. & Zachariah, N. Y. (1987). Bioavailability of progesterone with different modes of administration. *J. Reprod. Med.*, **32**, 443–448.

Freeman, E., Rickels, K., Sondheimer, S. J. & Polansky, M. (1990). Ineffectiveness of progesterone suppository treatment for premenstrual syndrome. *J. Am. med. Ass.* **264**, 349–353.

Hargrove, J. T., Maxson, W. S. & Wentz, A. C. (1989). Absorption of oral progesterone is influenced by vehicle and particle size. *Am. J. Obstet. Gynecol.*, **161**, 948–951.

Kleinknecht, R. A. & Donaldson, D. (1975). A review of the

effects of diazepam on cognitive and psychomotor performance. *J. Nerv. Ment. Dis.*, **161**, 399–411.

Lucki, I., Rickels, K. & Geller, A. M. (1986). Chronic use of benzodiazepines and psychomotor and cognitive test performance. *Psychopharmacology*, **88**, 426–433.

Lucki, I., Rickels, K., Giesecke, M. A. & Geller, A. (1987). Differential effects of the anxiolytic drugs, diazepam and buspirone, on memory function. *Br. J. clin. Pharmac.*, **23**, 207–211.

Majewska, M. D., Harrison, N. L., Schwartz, R. D., Barker, J. L. & Paul, S. M. (1986). Steroid hormone metabolites are barbiturate-like modulators of the GABA receptor. *Science*, **232**, 1004–1007.

Maxson, W. S. & Hargrove, J. T. (1986). Bioavailability of oral micronized progesterone. *Fertil. Steril.*, **44**, 622–626.

McNair, D. M. & Lorr, M. (1964). An analysis of mood in neurotics. *J. Abnorm. Soc. Psychol.*, **69**, 620–627.

Mendelson, W. B., Martin, J. V., Perlis, M., Wagner, R., Majewska, M. D. & Paul, J. M. (1987). Sleep induction by an adrenal steroid in the rat. *Psychopharmacology*, **93**, 226–229.

Myers, E. R., Sondheimer, S. J., Freeman, E.W., Strauss, J. F. & Rickels, K. (1987). Serum progesterone levels following vaginal administration of progesterone during the luteal phase. *Fertil. Steril.*, **47**, 71–75.

Peturrson, H., Gudjonsson, G. H. & Lader, M. H. (1983). Psychometric performance during withdrawal from long-term benzodiazepine treatment. *Psychopharmacology*, **81**, 345–349.

SAS Institute, Inc. (1989). SAS/STAT User's Guide, Version 6, 4th ed. Cary, NC: SAS Institute Inc.

Schmidt, P. J., Nieman, L. K., Grover, G. N., Muller, K. L., Merriam, G. R. & Rubinow, D. R. (1991). Lack of effect of induced menses on symptoms in women with premenstrual syndrome. *New Engl. J. Med.*, **324**, 1174–1179.

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.

Sitruk-Ware, R., Bricaire, C., Delignieres, B., Yaneva, H. & Mauvais-Jarvis, M. (1987). Oral micronized progesterone. *Contraception*, **36**, 373–402.

Smith, J. M. & Misiak, H. K. (1976). Critical flicker frequency (CFF) and psychotropic drugs in normal human subjects —a review. *Psychopharmacology*, **47**, 175–182.

Tapanacinen, J., Kauppila, A., Metsa-Ketela, T. & Vapaatalo, H. (1989). Prostanoids and catecholamines after oral administration of natural progesterone. *Gynecol. Endocrinol.*, **3**, 135–142.

Wittenborn, J. R. (1979). Effects of benzodiazepines on psychomotor performance. *Br. J. clin. Pharmac.*, **7**, 615–675.

(Received 1 July 1991, accepted 21 October 1991)