CALL FOR PAPERS | Sex and Gender Differences in Pain and Inflammation

D"es experimental pain response vary across the menstrual cycle?
A methodological review

Jeffrey J. Sherman and Linda LeResche
Department of Oral Medicine, University of Washington, Seattle, Washington

Sherman, Jeffrey J., and Linda LeResche. Does experimental pain response vary across the menstrual cycle? A methodological review. Am J Physiol Regul Integr Comp Physiol 291: R245–R256, 2006. First published February 16, 2006; doi:10.1152/ajpregu.00920.2005.—The findings on sex differences in human experimental pain research are inconsistent. One possible factor contributing to the inconsistent findings is the female hormonal cycle, as hormone levels may affect pain sensitivity. A number of studies suggest that women’s responses to experimentally evoked pain vary across the menstrual cycle. However, at least an equal number of studies suggest a lack of variability. The purpose of this article is to review the literature with emphasis on what we believe could be the reasons for the inconsistent findings, namely, differences in populations sampled, timing of experimental sessions across the menstrual cycle, and nomenclature used to identify the time (phases) in the cycle when measurements were done, nature of the pain stimuli chosen, and outcomes measured. These inconsistencies and other methodological problems associated with most experimental pain studies make it difficult to draw inferences from this literature. For the science to improve, replication of significant findings using standardized timing of sessions, pain stimulus procedures, outcomes, and hormonal assessment is necessary.

laboratory pain; acute pain; estrogen

IT IS WIDELY ACCEPTED THAT human males and females differ in their experience of and responses to pain. Epidemiological evidence on clinical pain suggests a female preponderance for most chronic pain conditions (38, 70). Women also appear more likely to experience multiple pain conditions (72). In the laboratory, there is evidence to suggest that females experience greater pain intensity, lower thresholds, and lower tolerance to experimentally induced pain (5, 16, 58). However, as Berkley (5) has noted, sex differences in human experimental studies are inconsistently observed and relatively minor. Fillingim and Maixner (16) reviewed 34 human experimental pain studies and found sex differences in only two-thirds of them.

A multitude of factors—biological, psychological, and social—likely contributes to gender differences in clinical and experimental pain response (39). As Derbyshire (14) and others (5) have pointed out, however, gender differences in the prevalence and severity of clinical pain appear more consistent and robust than the differences observed in the laboratory. One factor that could help explain the differential effects of sex on clinical and experimental pain, as well as the inconsistencies observed across human experimental studies, is the influence of female reproductive hormones on pain perception and response. Few studies on sex differences before 1995 recorded the time in the menstrual cycle when experimental manipulations took place or accounted for the variability associated with female reproductive hormones during the cycle.

Evidence from animal studies suggests that gonadal steroids may exert a substantial influence on female responsivity to painful stimuli. Specifically, animal models suggest a role for female reproductive hormones in both nociceptive and pain modulatory systems. Exogenous administration of luteinizing hormone increases sensitivity of female rats to nociceptive stimuli and diminishes the analgesic effects of morphine by desensitizing brain opiate receptors (3, 4, 55). On the other hand, estradiol has been found to play a critical role in the pain modulation system of female mice (50, 64), and the administration of estradiol and progesterone to mimic the hormonal milieu of pregnancy in rats produces analgesia that is modulated through the spinal cord κ-opiate receptor analgesic system (11, 12). Although a preponderance of evidence from the nonhuman animal experimental pain literature demonstrates hormonal influences on nociceptive responses, findings on the pattern and the direction of the relationship are still somewhat inconsistent (19). Methodological differences such as the type of nociceptive stimuli used (e.g., thermal, electrical, mechanical, chemical), the responses measured (e.g., flinch, tail flick, escape, grooming behaviors), and precise timing of the estrous cycle likely explain those inconsistencies (19).

It appears that some clinical pain conditions in human females vary with the menstrual cycle and during pregnancy. Specifically, temporomandibular pain (40, 41), migraine (30, 74), and possibly tension-type headache (30, 31, 74) appear to be more likely to occur or to be more intense at times of low
or rapidly fluctuating estrogen. However, the data on hormonal effects on human clinical pain are limited.

In contrast, numerous human experimental studies have examined changes in perceptual and behavioral responses to noxious stimuli across the menstrual cycle (for reviews, see Refs. 20 and 58). Presumably, these studies have the implicit or explicit goal of determining the relationship between hormonal levels, or changes in hormonal levels, and pain response in women. An early review (26) concluded that the research demonstrated a consistent pattern of highest pain intensity during days 15–22 of the prototypical 28-day cycle. More recently, Riley and colleagues (58) suggested that a pattern of menstrual cycle effects emerged for pressure stimulation, cold pressor pain, thermal heat stimulation, and ischemic muscle pain, with subjects in the follicular phase (days 6–11) demonstrating higher thresholds than at other phases. For electrical stimulation, pain thresholds appeared to be higher during the luteal phase (58). Despite these seemingly definitive findings, another review (20) concluded that even after numerous empirical studies and reviews on the topic, the strength and the nature of the relationship between female reproductive hormones and experimental pain intensity and response is far from clear. As with the nonhuman animal studies, methodological differences abound in the human experimental pain research and may account, at least in part, for inconsistent findings.

The purpose of this paper is threefold. First, we will review the literature relating to menstrual cycle effects on experimental pain in women with emphasis on what we believe could be the reasons for the inconsistent findings, namely, differences in populations sampled, timing of experimental sessions across the menstrual cycle, and nomenclature used to identify the time (phases) in the cycle when measurements were done, nature of the pain stimuli chosen, and outcomes measured. We will also review methodological problems associated with much of the human literature in the field, specifically, studies with between-subject designs, small sample sizes, sampling over only a single menstrual cycle, and lack of collection of biomarkers to confirm ovulation and track hormonal changes across the cycle. We will conclude with recommendations for future research on experimental pain response across the menstrual cycle and the clinical implications of such research.

**BACKGROUND**

**Hormonal Changes Across the Menstrual Cycle**

A menstrual cycle is conventionally defined as the time from the beginning of one menstrual flow (day 1) to the beginning of the next. Although there is a great deal of variability in menstrual cycle length between women and within individual women over time (69), the prototypical menstrual cycle is usually described for heuristic purposes as 28 days in length. Gynecologists divide the menstrual cycle into phases based on physiological events. Although the terminology for these phases varies somewhat, one common convention divides the cycle into three phases: follicular, ovulatory, and luteal. During the follicular phase, a sequence of actions of hormones and autocrine/paracrine peptides typically results in the maturation of a single mature follicle ready for ovulation. The follicular phase begins on the first day of menses and generally lasts 10–14 days. Ovulation is defined as the rupture of the follicle from the ovary. Approximately 24–36 h before ovulation a peak in estradiol (the predominant estrogen) occurs, followed by a peak in luteinizing hormone (LH) 10–12 h before ovulation. The LH surge is thought to be the most reliable indicator of impending ovulation (62). The luteal phase constitutes the remainder of the cycle, which lasts about 14 days. (Variability in cycle length is thought to be primarily due to variability in the length of the follicular, rather than the luteal phase.) The luteal phase is characterized by the differentiation of the corpus luteum, which then secretes progesterone and estradiol. Progesterone levels normally rise sharply after ovulation, peaking about 8 days after the LH surge; a secondary estradiol peak occurs at about the same time. During this phase, progesterone acts both centrally and locally to prevent further follicular development. The corpus luteum rapidly declines beginning 9–11 days after ovulation, resulting in rapid drops in estrogen and progesterone, which reach their lowest levels at the start of the menstrual period. These rapid declines in estrogen and progesterone have been associated with a constellation of symptoms, including pain, bloating, and mood changes colloquially referred to as premenstrual syndrome. The lower panels in Figs. 2–5 depict the patterns of change in estradiol and progesterone over an idealized 28-day cycle.

**Identification and Review of Studies**

Computer-based searches in PubMed, PsycINFO, and PsycARTICLES were conducted using combinations of the following keywords: pain, experimental, laboratory, acute, hormone, menstrual cycle, menstrual phase, estrogen, and progesterone. Reference sections from published articles in the field were also used as sources. Only articles published in peer-reviewed journals, using a within-subject design, that provided results for pain intensity, threshold, or tolerance outcomes are included in this review. Table 1 summarizes the methodology and results of these studies. In the next section, we begin by describing differences in experimental methodology that in and of themselves are not problematic. However, such inconsistencies make efforts to summarize and draw conclusions from the literature quite difficult. We then describe some methodological issues that are more problematic because they affect the basic quality of the research data, thus making it difficult to draw valid inferences regarding relationships between hormonal levels and pain. We conclude with recommendations for methods and procedures for future research that will mitigate these inconsistencies and improve our understanding of the relationship between experimental pain and reproductive hormones.

**DIFFERENCES IN EXPERIMENTAL METHODOLOGY**

**Variability in Populations Sampled**

**Chronic pain vs. pain free.** Numerous studies have found consistent differences in experimental pain intensity between chronic pain patients and normal controls (7, 44). Those differences have been attributed to alterations in the central and peripheral nervous systems and pain regulatory systems among persons with chronic pain (46, 57). Because chronic pain may alter responses to experimental pain, thus adding another possible source of variation among studies and because of the relatively small number of studies of menstrual cycle...
Effects on experimental pain in women with specific chronic pain conditions, this review will focus only on literature assessing experimental pain response in women without chronic pain.

Normally cycling women vs. oral contraceptive users. A few studies (15, 29, 60, 67, 71) have compared pain response across the cycle in normally cycling women vs. women who use oral contraceptives (OCs). The purpose of these comparisons is to assess pain response under the relatively stable hormone levels generated by OC use compared with the higher hormonal variability among normally cycling women. Indeed, some studies have found that OC use obviates menstrual cycle influences on experimental pain perception (29, 67). However, several of the studies (29, 67, 71) do not report the class and dosage of the specific formulations of OC. Such dosages and formulations alter the hormonal milieu and add a significant source of unreported variability. A separate review of findings of these studies would be useful but because of the small number of studies that have included OC groups and the even smaller number that report on type of OC, we have chosen to include in our summary only findings from normally cycling women.

Table 1. Summary of methods for studies assessing experimental pain across the menstrual cycle

<table>
<thead>
<tr>
<th>Ref. No.</th>
<th>Authors</th>
<th>Number of Subjects</th>
<th>Method of Determining Cycle</th>
<th>Timing of Sessions</th>
<th>Hormones Assayed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Amodei and Nelson-Gray, 1989</td>
<td>12</td>
<td>Self-report of menses onset</td>
<td>Premenstrual: 1–4 days before menses Menses: 1–4 days after start of cycle Menstrual: 1st day of menses Follicular: 7 ± 2 days after menses Periovulatory: 14 ± 2 days after menses Luteal: 21 ± 2 days after menses Perimenstrual: 28 days after menses to 3 days after next menses</td>
<td>No</td>
</tr>
<tr>
<td>9</td>
<td>Cimino et al., 2000</td>
<td>24</td>
<td>Self-report of menses onset</td>
<td>Premenstrual: 1–4 days before menses Menses: 1–4 days after start of cycle Menstrual: 1st day of menses Follicular: 5–12 days after menses Luteal: 16–27 days after menses Perimenstrual: 28 days after menses to 3 days after next menses</td>
<td>No</td>
</tr>
<tr>
<td>15</td>
<td>Drobek et al., 2002</td>
<td>10</td>
<td>Self-report of menses onset</td>
<td>Premenstrual: 1–9 days before menses Menstrual: Days 1–3 Luteal: 6–8 from LH Surge</td>
<td>No</td>
</tr>
<tr>
<td>18</td>
<td>Fillingim et al., 1997</td>
<td>11</td>
<td>PRISM Calendar LH surge by urine test to confirm ovulation</td>
<td>Midfollicular: Days 5–8 Ovulatory: Within 24 h of LH surge, mean day 14-7 Premenstrual: 1–9 days before menses Menstrual: Days 1–3 Luteal: 6–8 from LH Surge</td>
<td>Plasma estrogen, progesterone, LH, and beta-endorphin</td>
</tr>
<tr>
<td>29</td>
<td>Hapidou and Rollman, 1998</td>
<td>36</td>
<td>Self-report of menses onset</td>
<td>Menses: Days 1–3 Ovulatory: LH surge + 2 days Midluteal - LH + 7–8 days Late luteal - LH + 12–14 days Follicular: 4–9 days after menses Luteal: 8–12 days after LH surge</td>
<td>No</td>
</tr>
</tbody>
</table>

LH, luteinizing hormone.
Timing of Experimental Sessions Across the Menstrual Cycle and Cycle Phase Nomenclature

For most studies, the method used for identifying cycle phase was self-report based on the starting day of menses. Phases were then quantified by counting forward the days from the start of menses. Ovulation was confirmed in three studies using changes in basal body temperature, in four studies using urine tests to confirm LH surge, and in four studies using serum or saliva estrogen and progesterone assays. The relative merits of these approaches are discussed below. In addition, although nearly all studies standardized their presentation of results to a 28-day cycle, few reported the statistical methods used to standardize the results. Because of the great variability in timing of experimental manipulations, the figures in this paper use the convention of describing results by the days of a standardized 28-day menstrual cycle rather than by the exact timing and terminology used by the original authors. On the basis of this standardization, ovulation is presumed to occur on day 14 of the cycle.

Figure 1 summarizes the timing and nomenclature for the 14 studies on pain response and menstrual cycle relationships in pain-free controls meeting our inclusion criteria. This figure illustrates the lack of standardized operational definitions for identifying menstrual cycle phases. Researchers have used at least nine terms to define various cycle phases, and in nearly every case, each study defines the phase using a different range of days. For example, of the 14 studies reviewed, 10 measured pain responses during the “menstrual” phase. There is some consistency across these studies in that the timing of the experimental sessions, for this phase occurred between days 1 and 7 of the cycle. However, the exact timing of these sessions differs across nearly all of the studies. A review of the bottom panels of Figs. 2–5, which show fluctuations in estradiol and progesterone, however, suggests that this variability in the menses phase nomenclature may not be very important in assessing for associations between estradiol and progesterone levels and experimental pain because hormone levels are low with very little variability during this time. In contrast, of the 14 studies reviewed, 9 measured pain responses during what was termed the “follicular” phase. The exact timing for this phase differs across all of these studies with assessments occurring over a range of 11 days, between days 4 and 14 of the cycle. During this period, estradiol concentration may vary eightfold, from its lowest concentration in the cycle to its highest. Similarly, there is substantial variability across studies in the timing of experimental sessions and the terminology used to describe this timing during the second half of the cycle. Again, this results in a lack of hormonal comparability across studies (and possibly even across subjects within a study) for “luteal phase” assessments.

Variability in Outcome Measures

Three types of pain measurement outcomes are reported: pain intensity, threshold, and tolerance. Pain intensity is usually operationalized by a visual analog scale (VAS) or numeric rating system (NRS) asking the subject to rate the pain using anchors, such as “no pain” and “worst possible pain.” In some cases (29, 60), pain intensity is operationalized as the number of sites on the body reported as painful to palpation at a standardized pressure. Threshold measures are typically operationalized as the amount of time elapsed from the beginning of administration of a painful stimulus to the time at which the subject reports the stimulus turns to pain. Tolerance is typically the amount of time, increased pressure, electrical pulse, or heat that a subject can withstand. Pain intensity and threshold outcomes have historically been considered to be more
strongly associated with the sensory-discriminative aspect of pain, whereas tolerance may reflect the affective-motivational dimension of pain (54). Although these generalizations have been questioned, it is still plausible that hormonal changes across the menstrual cycle might affect intensity, threshold, and tolerance measures differently.

Variability in Pain Stimuli

The choice of experimental pain stimulus is complex for a variety of reasons. In general, the assessment of pain requires a reproducible pain stimulus that is strong enough to produce a measurable response, moderate enough to reveal individual differences, meaningful enough to bear some resemblance to pain in everyday life, or theoretically precise enough to elucidate a basic mechanism of pain responsiveness. Specific to gender and hormone research, certain pain stimuli may result in different physiological stress responses that are influenced by hormones of interest. Some pain stimuli may activate peripheral mechanisms (e.g., electrical pain stimulation), others may activate endogenous pain regulatory mechanisms (e.g., ischemic pain), others may be characterized by intense autonomic or hypothalmo-pituitary axis (HPA) axis arousal (e.g., ischemic pain, cold pressor), and of course, pain stimuli may operate along several of these pathways. Importantly though, each of these mechanisms may be differentially affected by gonadal hormones. For example, some studies have shown enhanced HPA response to stress during the luteal phase (34, 68) compared with the follicular phase.

SUMMARY OF EXPERIMENTAL FINDINGS BY PAIN STIMULUS

Because of the probability that choice of experimental stimulus affects results, we have chosen to summarize, in a series of figures, the results of studies that used similar stimulus modalities and similar outcome measures. In addition, for the reasons stated above, although many studies investigated more than one group of subjects, results reviewed here are limited to women who were free of chronic pain and had normal (unmedicated) menstrual cycles (i.e., healthy normal subjects). Most experimental studies have used a within-subject design. As discussed in the section METHODOLOGICAL PROBLEMS, between-subject designs add considerable error variance, so we also limited the review to studies using within-subject designs. Fourteen studies met these criteria and are reviewed in the sections that follow.

Ischemic Pain

Of the 14 studies reviewed, five used an ischemic pain stimulus (1, 18, 53, 60, 65). This experimental pain stimulus produces a deep, aching pain similar to many clinical pain syndromes (17). It is thought to activate endogenous opioid pain regulatory systems (45, 52), so menstrual cycle effects might be expected for this stimulus. All five studies measured ischemic pain intensity using numeric or verbal descriptors, or VAS ratings, but none found significant cyclic differences. Figure 2 illustrates the findings from the studies using ischemic pain thresholds and tolerance time outcomes. In this and all figures that follow, only differences at a trend level of significance or better (i.e., \( P < 0.10 \)) are displayed. Further, in this and all figures that follow, red lines indicate highest sensitivity to pain. In the case of threshold and tolerance outcomes, these indicate lowest threshold and tolerance measures. Blue lines indicate lowest sensitivity or highest threshold and tolerance outcomes. Interestingly, within each study on ischemic pain thresholds and tolerance, the pattern of results was identical for each outcome. Thus both are combined within a single figure. Two studies found highest threshold and tolerance times (indicating lowest sensitivity) at the follicular phase (days 5–8 or 4–9) when compared with ovulatory (within 24 h of LH surge) or luteal phases (1–9 days before menses and 5–10 days after confirmed LH surge). Three studies found no significant differences in ischemic pain threshold and tolerance times across the cycle.

In their review of the literature available at the time (1, 18, 53), Riley and colleagues (58) concluded that the largest cyclic effect on ischemic pain threshold and tolerance was that higher threshold and tolerance (i.e., lower pain sensitivity) occurred in the follicular compared with the luteal phase. Since that review, two additional studies using a similar ischemic pain stimulus found no significant cyclic effects. Further, of the two studies that found cyclic differences in threshold and tolerance times, one (18) had a 9-day window corresponding to the luteal phase. During this 9-day window, progesterone levels typically fluctuate from their highest levels (~8 ng/ml) to their lowest

![Fig. 2. Pain threshold and tolerance to ischemic stimulation. Solid lines indicate highest sensitivity to pain (lowest threshold and tolerance outcomes). Dashed lines indicate lowest sensitivity to pain (highest threshold and tolerance outcomes). Solid vs. dashed line differences are significant at \( P < 0.10 \). Hollow lines (i.e., ---) indicate that no differences significant at \( P < 0.10 \) were found. Length and placement of the lines indicate timing of the experimental sessions.](image-url)
levels (<1 ng/ml). Estradiol fluctuates from a secondary peak of ~400 pg/ml to its lowest level in the cycle. Because of the addition of two large sample, nonsignificant studies reviewed here, we believe that it is difficult to conclude that there are cyclic differences in ischemic pain threshold or tolerance times.

**Thermal Stimuli**

Four studies used a heat stimulus, and one study used a cold pressor task. Heat stimulation produces a brief, sharp, cutaneous pain that is attributable to both A-delta fiber activity and C-afferent activity (32). Heat-stimulated pain may not be influenced to the same degree as ischemic pain by endogenous opioid pain regulatory systems (18). In support of this hypothesis, naloxone, an opioid receptor antagonist has been shown to increase ischemic but not thermal pain (37, 63). Cold pressor pain is produced by the immersion of a limb in cold water. Like the ischemic pain task, the cold pressor is thought of as a natural pain that increases in intensity throughout exposure (32). There is considerable debate as to whether noxious cold is processed similarly or differently than noxious heat by the nervous system. There is some evidence that hot and cold nociception are genetically correlated in mice (49) and that the two stimuli are processed similarly at the level of the peripheral nervous system (35, 61). However, these modalities may be separable at higher levels of the nervous system (8, 10). We group the two stimuli together here only for illustrative purposes and not because we fall on the convergent side of the debate.

Of the five studies that used thermal stimuli, one measured pain intensity (28) and found no cyclic differences. Three studies (18, 28, 71) measured heat pain threshold and two (18, 71) measured tolerance but no significant cyclic differences were found for either outcome.

Two additional studies measured heat pain discrimination using signal detection methodology (24, 25). Although these studies did not include measures of pain intensity, threshold, or tolerance, they are mentioned here because they provide some data regarding the ability to discriminate sensations. Three stimulus intensities were presented randomly and subjects were assigned each stimulus to one of the following response categories: 1, nothing; 2, warm; 3, hot; 4, faintly painful; 5, moderately painful; and 6, strongly painful. These studies found greatest pain discriminatory ability at ovulation (ovulation pinpointed using basal body temperature) compared with menses (days 1–7), postmenses (days 8–14), and premenses (days 22–28).

While Riley and colleagues (59) concluded that highest heat thresholds and tolerance were observed during the follicular phase, they base that conclusion on one study (18) that did not demonstrate significant effects. In reviewing two additional studies, we find it difficult to concur. None of the studies using thermal heat or cold stimuli found significant cyclic differences on pain outcomes. Two studies (24, 25), however, suggest that heat pain discrimination may be greater during the ovulatory phase than at other times in the cycle.

**Electrical Pain**

Figure 3 illustrates the results of the three studies that used an electrical pain stimulus. Electrical pain has been criticized as an experimental pain induction procedure because it activates all classes of primary afferent nerve fibers producing both painful and nonpainful sensations (27). Further, some have described the stimulation as less natural than other pain stimuli (20, 27). In their review of pain perception across the menstrual cycle, Riley and colleagues (59) found that electrical pain was different from other pain modalities, with electrical pain showing highest thresholds during the luteal phase.

Electrical stimulation has been applied in menstrual cycle research cutaneously (23, 67, 71), subcutaneously, and intramuscularly (23). Veith and colleagues (71) found no cycle effects to cutaneous shock. Tedford and colleagues (67) found highest threshold values to cutaneous shock during days 15–21 compared with days 1–14. Giamberardino and colleagues (23) applied electrical stimulation at all three tissue depths (skin, subcutis, muscle) at three sites (abdomen, arm, leg). In their study, the highest pain thresholds occurred at the luteal phase regardless of tissue depth or site but statistical significance was achieved for only the abdominal subcutis and muscle sites with luteal phase threshold (days 17–22) greater than menstrual (days 2–6) and premenstrual (days 25–28) phases. It appears that there is some consistency using electrical stimulation with highest thresholds occurring after ovulation.

**Subcutaneous Injection of Capsaicin**

One study (22) used a novel pain stimulus involving an intradermal injection of capsaicin into the forehead to create a
trigeminal sensitization phenomenon simulating migraine. Capsaicin elicits burning pain and cutaneous neurogenic vaso-dilation by causing release of substance P and CGRP from sensory C-fibers (43, 66). The pain-related-dependent measures in this study included a pain intensity rating for the site around the injection and pressure pain threshold (PPT) using an electronic device (Algometer, Horby, Sweden AB) at bilateral frontalis and left deltoid sites before and after the capsaicin injection. We present the data on PPT after the injection in this section because thresholds were likely uniquely influenced by the capsaicin-induced sensitization phenomenon. Data on PPT before the injection are presented in the next section. These authors found cyclic differences in PPT around the injection site with thresholds during menses (days 1–3) significantly lower than during the luteal phase (6–8 days after confirmed ovulation). Consistent with the PPT findings, intensity ratings after the injection were higher during the menstrual phase at both frontalis sites.

Mechanical Pressure

There are numerous forms of mechanical pressure pain stimuli. One of the earliest was the Forgione Barber finger pressure device used in two studies (1, 36). This device uses a weight balanced on a fulcrum usually placed on a middle phalanx of a finger. Other forms of pressure reviewed include digital palpation at standardized body sites and digital pressure used to evaluate fibromyalgia tender points (29), as well as use of a pressure-transducing measurement device (algometer) (9, 15, 22, 60). Although pressure-transducing devices result in more reliably reproducible pain stimulation, digital palpation has more clinical relevance, being more frequently used in clinical examinations.

All studies using mechanical pressure measured pain thresholds (Fig. 4). Two studies (9, 15) measured pain pressure threshold at masseter and temporalis sites. Cimino and colleagues (9) found cyclic differences at both masseter and temporalis sites. Highest masseter thresholds were found at menses (first day of menstrual bleed), follicular (days 5–9), and luteal (days 19–23) compared with lowest masseter threshold during the periovulatory phase (days 12–16). For the temporalis site, the highest threshold was at the luteal (days 19–23) compared with the lowest thresholds at periovulatory (days 12–16) phase. Drobek and colleagues (15) found highest thresholds at the masseter site during the perimenstrual phase (28 days after menses to 3 days after next menses) compared with the follicular phase (days 5–12). There were no cyclic differences at the temporalis sites. Gazerani and colleagues (22) measured PPT at bilateral frontalis and left deltoid sites and found cyclic differences only in the left frontalis muscle with highest PPT during the luteal (6–8 days after confirmed ovulation) compared with the menstrual phase (days 1–3). No other studies found cyclic differences in pressure pain thresholds. Two studies using mechanical pressure measured pressure pain tolerance (36, 60) but neither found significant cyclic effects.

Of the studies using mechanical pressure stimuli, three (29, 36, 60) measured pain intensity (Fig. 5). Kuczmierczyk and colleagues (36) found highest pain intensity (numeric ratings to

![Fig. 4. Pain thresholds to pressure. Solid lines indicate highest sensitivity to pain (lowest threshold and tolerance outcomes). Dashed lines indicate lowest sensitivity to pain (highest threshold and tolerance outcomes). Solid vs. dashed differences are significant at P < 0.10. Hollow lines (i.e., ==) indicate that no differences significant at P < 0.10 were found. Length and placement of the lines indicate timing of the experimental sessions. aForgione Barber device-induced finger pressure. bAlgometer-induced pressure at right masseter belly. cAlgometer-induced pressure at right temporalis 2 cm behind anterior border of muscle. dAlgometer-induced pressure at 1.5 cm superior to first masseter site. eAlgometer-induced pressure at 2 cm superior to first temporalis site. fAverage of algometer-induced pressure at right and left temporals. gAverage of algometer-induced pressure at right and left frontalis. hAlgometer-induced pressure at left deltoid muscle. iAlgometer-induced pressure at left frontalis muscle. jAlgometer-induced pressure at right frontalis muscle. kDolorimeter-induced pressure at six unilateral active and three control sites from the list of diagnostic criteria for fibromyalgia syndrome. lForgione Barber device-induced finger pressure.

Additional annotations:
- Amoel-Nelson-Gray '89
- Cimino et al., '90
- Cimino et al., '90
- Cimino et al., '90
- Cimino et al., '90
- Drobek et al., '92
- Drobek et al., '92
- Gazereani et al., '95
- Gazereani et al., '95
- Gazereani et al., '95
- Hapiou et al., '99
- Kuczmerczyk et al., '86
- Sherman et al. '95
- Sherman et al. '95

![Graph](https://example.com/graph.png)

### Day of Menstrual Cycle

<table>
<thead>
<tr>
<th>Day</th>
<th>Estradiol</th>
<th>Progesterone</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>30</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>40</td>
<td>20</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>7</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>8</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>9</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>10</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>11</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>12</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>13</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>14</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>15</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>16</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>17</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>18</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>19</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>20</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>21</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>22</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>23</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>24</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>25</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>26</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>27</td>
<td>50</td>
<td>20</td>
</tr>
</tbody>
</table>

By 10.220.33.6 on October 14, 2017 http://ajpregu.physiology.org/ Downloaded from
finger pressure) in the intermenstrual (days 7–22) compared with the premenstrual (days 24–28) phase. Hapidou and Rollman (29) found the greatest number of tender points rated as painful in the follicular (days 8–14) compared with the early and late luteal phases (days 15–21 and 22–28, respectively). Sherman and colleagues (60) measured palpation pain intensity at fixed amounts and rates of pressure at sites used for clinical evaluation of temporomandibular disorders and fibromyalgia syndrome and found no significant differences in intensity across the cycle. It appears that the most consistent cyclic effect for mechanical pressure, albeit in only two studies, is in the intensity rating outcome and that pressure pain ratings are highest between days 7 and 21 of the cycle.

**SUMMARY OF FINDINGS BY OUTCOME**

All five (1, 18, 53, 60, 65) studies using an ischemic pain stimulus used nearly identical methodology and outcomes for delivering the stimulus. The delivery of the other stimuli, however, varied considerably by the body site stimulated, and the amount of pressure, technique for pressure delivery, amplitude of electricity, and intensity of heat administered. It is plausible that cyclic effects on experimental pain may be more apparent when illustrated by outcome measure rather than by pain stimulus used, so in an attempt to better understand the findings, we also graphed results for all studies by grouping them based on pain outcome (intensity, threshold, tolerance). There were no consistent cyclic patterns on any pain outcome.

**METHODOLOGICAL PROBLEMS**

High-quality studies of pain response across the menstrual cycle are very difficult to conduct, primarily because of the large inter- and intra-individual variability in the length of the menstrual cycle. The use of biological markers (e.g., urine testing for LH surge, blood sampling) to track the cycle increases certainty about the underlying hormonal state but also increases cost and subject burden. Furthermore, critical times of a subject’s cycle may not conveniently correspond with the technician’s normal 40-h work week. These are logistical problems that are often unavoidable. However, some methodological problems that are fairly easily avoided are listed below. Our primary purpose in enumerating these issues is not to critique earlier studies, many of which used the best approaches available at the time, but rather to improve the quality of future research.

**Between-Subject Designs**

Pain perception and tolerance are influenced by complex interactions between biological variables (e.g., body size, muscle mass, genetics, pain inhibitory pathways, and CNS variation) and psychosocial variables (e.g., depression, anxiety, culture, sex role expectancies, social learning factors, pain-related appraisal, and gender of experimenter). There are substantial differences in these variables within individuals and even greater differences between individuals. Statistically, large differences between individuals make it difficult to observe meaningful group differences in between-subject designs. For example, the variability in hormonal levels and lengths of menstrual cycles across women is profound. If one were to study one group of women in the luteal phase and another in the follicular phase, variability due to chance and other unaccounted for variables is likely to be so great as to obscure real cyclic effects. If the goal of the research is to detect meaningful phase differences in experimental pain response, it seems essential to use within-subject designs.

**Method for Tracking Cycle Phase**

Many of the studies (1, 9, 15, 23, 29, 36, 67) we reviewed did not determine when, or even whether, ovulation occurred, making it difficult to accurately assess phase of the cycle. The most commonly used method in the studies reviewed was to rely solely on self-report of the start of menses. Although report of menses onset is likely fairly reliable, this method is problematic for two reasons. First variability in cycle length makes precise prediction of hormonal changes and cycle phase problematic. Second, studies that intend to capture data at the late luteal phase will likely lose data if menses occurs early and collect data that are hormonally irrelevant if menses occurs late.

Additionally, few studies used LH testing or other biomarkers to confirm ovulation. Determination of ovulation is important, because ~1/4–1/3 of cycles may be anovulatory in normal women (47). Importantly, even higher rates of anovulatory cycles—up to 50%—have been found for young women living in dormitories and other group housing arrangements (47). This population may constitute a prime source of subjects for menstrual cycle studies. If ovulation does not occur, the hormonal milieu in the second half of the cycle differs markedly from that which occurs in a normal ovulatory cycle. In
those studies where ovulation was assessed, the method used most often was monitoring of basal body temperature. However, compliance with this method is problematic, and this method is less accurate than hormonal measurement (33, 42).

Lack of Hormone Status Measurement

If the underlying aim of menstrual cycle studies is to assess the relationship of pain to changes in hormone levels, it would seem useful to measure hormone levels directly. Although blood draws for assessment of hormone status provide the most accurate measurement of blood levels of hormones and their metabolites, the method is invasive, costly and unpleasant (59). Salivary assessment of estradiol and progesterone is a reliable and valid reflection of the unbound (bioactive) fraction of the hormone in blood (21, 56). Salivary concentrations of these hormones are independent of salivary flow rate and do not exhibit strong diurnal variability (6, 13, 73). The methodology has been used in observational and experimental studies (40, 60) and has been shown to be valid compared with serum levels. However, the expense of such assays may have limited their use.

Small Sample Size

As can be seen in Table 1, this literature is replete with small and likely underpowered studies. Thus the predominance of nonsignificant findings is not surprising and considering the other inconsistencies in study methodology, the conflicting results are not surprising. In the last several years, there have been at least four relatively large studies with 24 or more subjects (9, 60, 65), but only one of these found significant differences across the cycle. Further, we believe that an important methodological problem with most of the studies reviewed has been the limitation imposed on availability and reliability of data by resorting to collection during only one menstrual cycle. Sampling across multiple cycles increases power and allows for fewer subjects overall. Nevertheless, in our past research sampling pain outcomes at four distinct times corresponding to predicted peaks and nadirs of estradiol and progesterone in each of three consecutive cycles, our sample size estimates suggested that to detect a change in subject’s response due to cyclic variation on the order of 1/2 of a standard deviation, and considering the likelihood of about a 20% missing data, 33 subjects would be required to allow for power > 0.80. By this standard there has been only one (29) adequately powered study published to date.

DISCUSSION

It is understood that the body of scientific knowledge in a field grows through replication and elaboration of findings. The purpose of this report was to review the methodology and findings of the extant literature examining cyclic variation in experimental pain perception in the hopes of finding consistent effects, or lacking those effects, illustrating inconsistencies and methodological problems in the field that make replication of findings unlikely. In reviewing the literature and calculating effect sizes for 16 experimental pain stimuli studies, Riley and colleagues (59) concluded that the menstrual cycle affects responses to experimental pain stimuli, but that the impact of menstrual cycle phase on pain thresholds is different for different types of stimulation. In contrast to that conclusion, we see little evidence for a cycle effect on experimentally evoked response to specific pain stimuli, with the possible exception of electrical pain stimulation.

In her review of the experimental literature on sex differences, Berkley (5) noted that sex differences are “inconsistently observed, relatively minor, exist only for certain forms of stimulation, and can be affected by numerous . . . variables.” We believe that the same statement could describe the existing literature on menstrual cycle effects on experimental pain. If menstrual cycle-related differences in responses to experimental pain exist, these differences are likely not dramatic, for if dramatic effects occur, more consistencies would be apparent across the studies reviewed. However, if the size of differences is moderate or small, researchers must design tightly controlled studies to be able to detect cycle effects amidst the large behavioral variation within and between individuals. In addition, researchers need to pay better attention to the actual variations in hormone levels across the cycle. If menstrual cycle effects are moderate or small, inconsistencies and methodological problems may increase error variance to such an extent as to obscure experimental effects. We believe that the inconsistencies illustrated and methodological problems described make it difficult to draw conclusions about the relationship of female reproductive hormones and experimental pain at this time. For the state of the science to improve, more consistency in numerous methodological areas is needed. Only then can results from studies be quantitatively combined to determine whether the menstrual cycle exerts statistically significant, and more importantly, clinically meaningful effects on pain.

We have demonstrated the considerable variability across studies in the nomenclature of cycle phase and in the actual timing of experimental manipulation. In reviewing the literature, we observed that researchers have used at least nine terms to define the various cycle phases and that in nearly every case, there was almost no concordance across studies on exact timing of those phases. In conducting our own experimental pain research in this area, we noted the logistical difficulties associated with scheduling a subject into the laboratory with short notice (e.g., within 2 or 3 days of starting menses or a positive LH surge) or at times that can only be predicted based on the presumed time for beginning the next menstrual cycle (e.g., late luteal phase). Understandably, most studies use wide “windows of opportunity” for scheduling subjects into the laboratory, but the variability in the hormonal milieu within these windows likely obscures any clear experimental effect. Further obscuring true experimental effects, few studies confirm ovulation or assay hormones.

Along with profound hormonal variation across the cycle, comes substantial variation in the emotional milieu. For example, along with dramatic declines in estradiol and progesterone during the days before menses, many women experience numerous physical and emotional changes such as bloating, cramping, pain, fatigue, irritability, and sadness. For many women, the anticipation, presence, severity, and combination of these symptoms is extremely stressful. Becker and colleagues (2) review the bidirectional relationship between gonadal steroids and stress and suggest that experiments examining sex differences consider the potential mitigating influences of stress and stress hormones. As stress and stress hormones can also impact experimental pain sensitivity, and

AJP-Regul Integr Comp Physiol • VOL 291 • AUGUST 2006 • www.ajpregu.org

Downloaded from http://ajpregu.physiology.org/ by 10.220.33.6 on October 14, 2017
many women experience additional stress during times of rapid hormone declines, the same suggestion can be made for menstrual cycle experiments. Measuring stress and emotional changes along with reproductive hormones and narrowing windows of data collection could further reduce the “noise” that obscures true experimental effects.

**Strategies for Future Research**

The Special Interest Group on Sex, Gender and Pain of the International Association for the Study of Pain has suggested that, to speed the progress of research on sex and gender issues related to pain, a task force of interested researchers should work together to develop standardized measures and conventions (29a). In the case of studies of experimental pain response across the menstrual cycle, we heartily concur with this recommendation. Knowing the biological mechanisms that underlie profound differences in health and reactivity between the sexes could be extremely helpful in elucidating the pathogenesis of various common pain problems. The animal literature suggests that hormonal differences are a likely mechanism for differences in pain perception and/or reactivity. However, to better understand the relationship between reproductive hormones and pain, experimenters must attend to several issues.

Recently, Becker and colleagues (2) published a methodological paper advocating strategies and methods for research on sex differences. Understandably, many of their recommendations apply to research on variations in experimental pain response across the menstrual cycle. For example, one basic question to be addressed is what relationships we are attempting to discover through menstrual cycle studies. We believe that the underlying goal of much previous menstrual cycle research has been to detect pain/hormone relationships in humans. In the absence of direct measurement of hormones, phases of the menstrual cycle have been used as surrogate measures of hormonal levels. Unless there are theoretical reasons for studying the menstrual cycle itself (e.g., mood changes linked to the cycle), this approach has several problems. First, the gynecological nomenclature for cycle phases is geared toward reproductive function. However, hormone levels within a gynecologically defined phase may vary greatly (2). If the true aim of menstrual cycle studies is to assess the relationship of pain to hormone levels, and menstrual cycle approaches are to be used, standardizing the timing of data collection to narrow, biologically relevant windows (e.g., times of highest and lowest hormone levels) is of utmost importance. We believe it would be productive to abandon gynecological terminology altogether or, in the case of theoretically appropriate reasons for studying the menstrual cycle (2), to adopt a standardized approach to defining and naming these critical times in the menstrual cycle.

Second, we suggest that prediction of ovulation using LH detection become a minimal standard in menstrual cycle studies. Although a positive LH test is not a perfect guarantee that ovulation will occur (and hence that the “expected” hormonal patterns will occur in the second half of the cycle), these tests are relatively easy for subjects to perform and have high sensitivity and specificity for detecting ovulation (48). Additional measurement of midluteal progesterone levels can provide actual confirmation of ovulation.

Third, as others have noted (2) if the true aim of menstrual cycle studies is to assess pain/hormone relationships, it seems logical and essential to measure hormones directly as a confirmation that expected changes in estradiol or progesterone have occurred. In the past, direct observation of hormone levels required drawing blood from subjects for serum assays. This is an invasive procedure, stressful to subjects and requiring an adaptation period so that the stress of the venipuncture does not influence outcome. Although saliva collection has its own technical issues explored in other papers (2) (e.g., need to time collections before, rather than after meals; use of cotton as a collection device may produce inaccurate readings), it is relatively noninvasive and several commercial laboratories in the United States and Europe are now equipped to provide assays for estradiol and progesterone in saliva. Thus hormone levels could be assayed daily or at least on the days of experimental sessions. Researchers should even consider the direct manipulation of hormone levels (2), e.g., using an estrogen patch, as has been done in a study of brain responses to experimental pain (75). Salivary assays could be used to confirm that the manipulations were successful.

Fourth, exogenous hormone use may attenuate the cycle effects on experimental pain and thus studies including an OC group should report on the type and dosage of OC administered.

Finally, we recognize that choice of stimulus and body site examined may be driven by specific theoretical or clinical considerations. However, to create a more coherent body of research, we suggest that researchers attempt to reach consensus on standardizing stimulus modalities, technique of delivery, and body site tested.

**GRANTS**

This study was supported by National Institute of Dental and Craniofacial Research Grant R-01DE-16212.

**REFERENCES**

PAIN RESPONSES ACROSS THE MENSTRUAL CYCLE


29. International Association for the Study of Pain Special Interest Group, Newsletter of the IASP. Special Interest Group on Sex, Gender, and Pain, November, 2004, p. 3–4.


