



Published in final edited form as:

Psychol Med. 2008 December ; 38(12): 1749–1757. doi:10.1017/S0033291708002997.

Ovarian hormones and binge eating: exploring associations in community samples

K. L. Klump^{1,*}, P. K. Keel², K. M. Culbert¹, and C. Edler²

¹ Department of Psychology, Michigan State University, East Lansing, MI, USA

² Department of Psychology, University of Iowa, Iowa City, IA, USA

Abstract

Background—Significant associations between changes in ovarian hormones and binge eating are present across the menstrual cycle in women with bulimia nervosa. However, no study has examined these relationships in a non-clinical sample, despite the need for these data for designing risk-factor studies.

Method—In study 1, we modified several continuous measures of binge eating and identified those that were most sensitive to menstrual-cycle fluctuations in a non-clinical sample of 10 women who completed measures for 35 days. In study 2, we explored associations between ovarian hormones and binge-eating scores in nine women who completed these same measures for 65 days and provided daily saliva samples for assays of estradiol and progesterone concentrations.

Results—In study 1, the Emotional Eating subscale of the Dutch Eating Behavior Questionnaire exhibited superior reliability and was most sensitive to predicted menstrual-cycle changes in binge eating (i.e. increased scores in the mid-luteal/premenstrual compared with follicular/ovulatory phases). In study 2, this scale showed predicted inverse associations with estradiol and positive associations with progesterone across the menstrual cycle that could not be accounted for by changes in negative affect.

Conclusion—Associations between ovarian hormones and binge eating are robust and present in clinical and non-clinical samples. Findings support the ability to examine the role of ovarian hormones as risk factors for binge eating in large-scale prospective studies and twin studies.

Keywords

Binge eating; estrogen; menstrual cycle; ovarian hormones; progesterone

Introduction

Emerging data show significant associations between changes in ovarian hormone levels and binge eating across the menstrual cycle in women with bulimia nervosa (BN) (Edler *et al.* 2007). Increased frequency of binge episodes has been found during the mid-luteal and premenstrual phases compared with the follicular and ovulatory phases in women with BN (Gladis & Walsh, 1987; Price *et al.* 1987; Lester *et al.* 2003; Edler *et al.* 2007), and a recent longitudinal study found that increases in estradiol, and decreases in progesterone, prospectively predict increases in binge frequency (Edler *et al.* 2007). These associations

* Address for correspondence: K. L. Klump, Ph.D., Department of Psychology, Michigan State University, 107B Psychology Building, East Lansing, MI 48824-1116, USA. (klump@msu.edu).

Declaration of Interest: None.

accounted for changes in binge frequency over the course of the menstrual cycle and were independent of baseline body weight and daily changes in negative affect (Edler *et al.* 2007). Moreover, changes in ovarian hormones accounted for 24% of the variance in binge frequency (Edler *et al.* 2007).

The strength of these within-subject designs comes from the fact that changes in estrogen and progesterone across the menstrual cycle occur in all normally menstruating women in preparation for conception. Thus, changes in ovarian hormones that precede changes in binge eating likely reflect the impact of hormones on binge eating rather than the reverse. This hypothesis is consistent with a vast animal literature demonstrating that experimental manipulations of ovarian hormones cause changes in food intake (Asarian & Geary, 2006). For example, ovariectomy of adult female rats removes the primary source of estrogen and causes immediate and sustained increases in food intake and weight (Wade, 1975). Exogenous estradiol treatment reverses these effects (Varma *et al.* 1999). Progesterone causes increased food intake in several species (e.g. rodents; Roberts *et al.* 1972; Varma *et al.* 1999), in part, by antagonizing the effects of estrogen. These patterns of food intake are similar to the binge-eating patterns observed in women with BN, where increased binge eating is observed during menstrual-cycle phases when estradiol levels are low and progesterone levels are high (e.g. the mid-luteal phase).

Despite support for a role of ovarian hormones, no study has examined associations between ovarian hormone levels and binge eating across the menstrual cycle in non-clinical samples of women. There is a strong need for such research to clarify whether hormone-behavior associations in women with BN can be examined as etiologic rather than maintenance factors. Twin research provides preliminary support for an etiologic role, as estrogen activation at puberty appears to increase genetic influences on disordered eating (Klump *et al.* 2003, 2007b). However, measurement issues complicate our ability to examine genetic effects across the menstrual cycle. Specifically, menstrual-cycle studies of women with BN (Edler *et al.* 2007) have used categorical measures of binge frequency that lack sensitivity in non-clinical samples where binge frequency would be expected to be lower. Continuous measures provide greater power to detect binge behaviors in community-based samples, but none of these scales assesses binge behaviors on a daily basis, limiting their usefulness for studying associations between daily fluctuations in hormones and binge eating across the menstrual cycle.

We conducted a two-part study to address these issues and lay the groundwork for twin studies of hormone-binge-eating associations across the menstrual cycle. In study 1, we modified existing binge-eating measures to assess daily fluctuations in binge episodes and compared their reliability and sensitivity for detecting ovarian hormone and binge-eating associations across menstrual-cycle phase in a non-clinical sample of women. In study 2, we used the most sensitive binge-eating measure identified in study 1 to confirm prospective associations between ovarian hormone concentrations and binge eating in a non-clinical sample.

Method

Study 1: examining modified binge-eating measures

Participants—Participants were recruited through public advertisements at the University of Iowa and surrounding community, including posters placed on campus, community buses and grocery store bulletin boards. Inclusion criteria were the following: (1) age between 18 and 45 years; (2) menstruation every 22 to 35 days; (3) body mass index (BMI) between 19 kg/m² and 29 kg/m²; (4) no hormonal contraceptive, psychotropic medication (i.e. medications that alter brain functioning, e.g. antidepressant medication), or steroid medication use within the past 8 weeks; (5) no pregnancy or lactation within the past 6 months; (6) no history of genetic (e.g. Turner's syndrome) or medical conditions (e.g. diabetes, congenital adrenal

hyperphasia) known to influence hypothalamic–pituitary–gonadal or hypothalamic–pituitary–adrenal axis functioning, weight, or appetite. All inclusion criteria were assessed via self-report during telephone screens and were confirmed during intake and follow-up assessments.

A total of 13 women were recruited for the study. One of these women (8%) did not menstruate during the 35-day collection period and was thus ineligible for participation. One woman (8%) did not return for study visit 2 and was considered to have dropped out of the study. One additional woman (8%) completed only 17 of the 35 days of data collection and was excluded from analyses due to non-adherence with data collection procedures. Thus, 77% of women [$n=10$, mean age=28 ($s.d.=8.1$) years, mean BMI=22.9 ($s.d.=4.2$) kg/m²] enrolled in the study, completed data collection and remained eligible for inclusion in analyses. Among participants, 80% ($n=8$) were Caucasian and 20% ($n=2$) were Asian. Most participants ($n=9$, 70%) were in college or had completed a college degree ($n=2$, 20%).

Procedures—Participants came to the laboratory for two study visits separated by 35 days. Prior to participation, participants completed an informed consent document. Study visit 1 was used to confirm study eligibility and describe in-home data collection procedures. Participants completed daily assessments of binge eating and mood each evening for 35 consecutive days to capture one full menstrual cycle. Evening assessments were chosen based on previous research (Lester *et al.* 2003; Edler *et al.* 2007) where they were shown to maximize convenience for participants and reduce potential for recall biases that may be introduced with morning recall of the previous day's behaviors. Participants were given email addresses and phone numbers to contact project staff to discuss problems or difficulties that arose between study visits. During study visit 2, participants returned their assessments and staff reviewed these with participants.

Measures

Binge eating: We examined binge-eating measures that have been commonly used in the literature, exhibit good psychometric properties in their unmodified form and are relatively brief. The measures included: the Three Factor Eating Questionnaire (TFEQ) Disinhibition and Hunger subscales (Stunkard & Messick, 1985), the Dutch Eating Behavior Questionnaire (DEBQ) Emotional Eating and External Eating subscales (Van Strien *et al.* 1986), the Eating Disorders Examination Questionnaire (EDEQ) Eating Concern subscale (Fairburn & Beglin, 1994) and the Eating Disorder Inventory (EDI) Bulimia subscale (Garner *et al.* 1983).

Binge-eating measures were modified with permission so that instructions and questions, as appropriate, referred to experiences of that day. For example, the instructions of the original DEBQ asked participants to decide on a scale from 1 (=never) to 5 (=very often) whether items are 'true in relation to you' and thus were originally designed to tap stable individual differences on items such as, 'Do you have the desire to eat when you are irritated?'. These instructions were modified so that participants were to decide whether items were 'true in relation to you TODAY'.

Mood: The Positive and Negative Affect Schedule (PANAS; Watson *et al.* 1988) was used to assess fluctuations in mood. This measure was specifically designed to assess mood states (e.g. feelings of depression and anxiety) over different time periods, including a single day. Thus, it is ideal for examining daily fluctuations in mood and required no modifications. The PANAS demonstrates good convergent and discriminant validity (Watson *et al.* 1988), and internal consistencies for the Negative Affect scale on a single day were >0.85 in study 1 as well as in other published work (Edler *et al.* 2007).

Height and weight: BMI (weight in kg/height in m²) was calculated from height and weight measurements obtained at the intake assessment with a wall-mounted ruler and digital scale.

Menstrual bleeding: In order to determine menstrual-cycle phase, participants indicated dates of menstrual bleeding using a daily log book (see Lester *et al.* 2003).

Data analyses—Menstrual-cycle phase was determined using previously described methods (Lester *et al.* 2003). Briefly, the first day of menstrual bleeding was labeled ‘day + 1’, while the previous day was labeled ‘day -1’. The ovulatory phase then included days -15 to -12 (i.e. 12–15 days before the first day of menstrual bleeding), the mid-luteal phase included days -9 to -5, the premenstrual phase included days -3 to +1, and the follicular phase included days +5 to +10. Consistent with methods reported previously (Lester *et al.* 2003; Edler *et al.* 2007), scale scores were averaged across the days of the menstrual-cycle phase and converted into Z scores based on each participant’s mean and standard deviation score for that measure over the course of data collection.

Internal consistency for the daily binge-eating measures was calculated across subjects for each day and then averaged across days. Test–retest reliability analyses using the Spearman–Brown coefficient and intra-class correlations were conducted for 2 days within a cycle phase to estimate the degree of variance over time that could be attributed to sources unrelated to changes in ovarian hormones, including random sources of error. Repeated-measures analysis of variance (ANOVA) was then used to examine mean differences in binge scores between cycle phases. This final set of analyses was used to identify scales that demonstrated variability across menstrual-cycle phase that has been observed for changes in binge frequency in patients with BN (Lester *et al.* 2003; Edler *et al.* 2007).

Study 2: prospective associations between ovarian hormones and binge eating

Participants—Participants were a convenience sample of women drawn from the Michigan State University Twin Registry (MSUTR; Klump & Burt, 2006). The MSUTR sample is well characterized in terms of age, BMI, medical conditions and medication use (including oral contraceptive use). Thus, we capitalized on the availability of MSUTR participants and assessed 15 female adult twins from the MSUTR who met inclusion criteria for study 1. Although 50% (152/304) of the MSUTR female twins met these inclusion criteria, we limited our assessments to a subgroup (9%, 15/152) due to the high cost of daily hormone assays and the limited resources of this internally funded study. Importantly, the 15 assessed twins were not significantly different from MSUTR female twins in terms of ethnicity, socio-economic status, age, BMI, baseline levels of binge eating, compensatory behaviors, weight preoccupation or negative affect (data not shown). Thus, our subgroup appears to be representative of the larger MSUTR sample which is broadly representative of young women in Michigan (see Culbert *et al.* in press).

The assessed sample included 14 twins from seven complete twin pairs and one twin from an incomplete pair whose co-twin did not meet BMI criteria and thus was not eligible for the study. Although several complete twin pairs were assessed, small samples prohibited us from conducting twin analyses. Twins were instead treated as individuals in analyses, with appropriate modifications for correlated data (see Data analyses).

Of the original 15 twins enrolled in the study, two twins subsequently dropped out within the first week due to difficulties generating the required volume of saliva. Four additional twins did not follow study procedures (e.g. collected saliva on 1 day).

Complete data from a total of nine twins (nine out of 15, 60%) were included in analyses. Included twins were all full-time college students [mean age=23 (*s.d.*=23) years] of average

BMI [mean=22.9 ($s.d.=4.2$) kg/m²] and relatively diverse ethnicity [$n=6$ (67%), $n=2$ (22%) and $n=1$ (11%) Caucasian, African-American and Hispanic respectively]. Using previously collected data from the MSUTR, we compared twins who completed the assessments ($n=9$) to those who dropped out or were ineligible ($n=6$). Although participating twins exhibited higher compensatory behavior, weight preoccupation, and depression scores (d values=0.43–0.80), there were no significant differences in binge eating ($d=0.29$), BMI ($d=0.36$), age ($d=0.22$), ethnicity ($w=0.29$) or work status (i.e. all were college students). Thus, in aggregate, findings suggest that our participating twins were similar on key characteristics (e.g. binge eating) to those who did not complete the study, and that attrition was associated with less (rather than more) pathology.

Procedures—All participants completed an informed consent document prior to beginning the study. Study visit 1 was used to confirm study eligibility, obtain height and weight measurements, and describe data collection procedures. Participants were instructed to collect daily morning salivary hormone and evening behavioral data for 65 consecutive days (i.e. two menstrual cycles). Salivary samples were collected every morning within 30 min of waking before teeth brushing, eating, drinking or chewing. Participants passively drooled through a straw into a tube until at least 4 ml was produced. Immediately after collecting their saliva, participants recorded the date and time of collection on their tube. The timing of saliva collection and behavioral data recording ensured that a given day's ovarian hormone collection preceded events of that day (including levels of binge eating) which were recorded that evening.

Study visit 2 occurred approximately 30 days into data collection. Project staff met with participants to collect saliva specimens, behavioral questionnaires, and provide new data collection materials. Adherence to the study protocol and continued eligibility for the study also were assessed. Project staff called and emailed participants approximately every 2 weeks throughout data collection to answer questions. Participants were given email addresses and phone numbers of project staff to discuss problems or difficulties that arose between study contacts. Study visit 3 occurred after the 65-day data collection period. Participants returned their saliva and behavioral data which were reviewed by project staff for completeness.

Measures

Binge eating: Based on results from study 1, the modified version of the Emotional Eating subscale of the DEBQ was used to assess binge eating in study 2. The Emotional Eating subscale assesses the tendency to eat excessive amounts of food for reasons that are typically endorsed by individuals who experience recurrent binge-eating episodes. This subscale has demonstrated validity in distinguishing among normal controls, overweight women and bulimic patients, indicating that scores on this scale not only discriminate between normal eating and overeating but also overeating and binge eating (Wardle, 1987). Internal consistency of the modified version of the Emotional Eating subscale in study 2 was excellent ($\alpha=0.98$).

Mood symptoms: The PANAS was used to measure negative affect and exhibited high internal consistency ($\alpha=0.91$) in the study 2 sample.

Height and weight: Height and weight were measured with a wall-mounted ruler and digital scale.

Menstrual bleeding: As in study 1, participants reported dates of menstrual bleeding in a daily log book.

Ovarian hormones: Similar to previous research (Edler *et al.* 2007), saliva samples were used to examine estradiol and progesterone concentrations. Monitoring biomarkers in saliva has

distinct advantages over doing so in other biological fluids (e.g. serum) (Shirtcliff *et al.* 2000). Sampling saliva represents a less-invasive method for repeated sampling schedules, and salivary levels reflect unbound hormones that provide more accurate indicators of active forms of estradiol and progesterone. Finally, saliva sampling is associated with higher compliance and more robust hormone– behavior associations than bloodspot samples (Edler *et al.* 2007).

To minimize variance in assay quality due to differences in assay reagents, participant samples were stored until all of the participant's samples were received. Each assay included data from one participant. Estradiol assays were conducted by Salimetrics, LLC (State College, PA, USA) using radioimmunoassay (Diagnostic Systems Laboratory, Webster, TX, USA), while progesterone assays were conducted by the Michigan State University Genetics Service Laboratory (East Lansing, MI, USA) also using radioimmunoassay (Diagnostic Systems Laboratory). All assay procedures followed previously described protocols (Jasienska *et al.* 2004). The average intra- and inter-assay coefficients of variation (CV) for estradiol were <6% and <9%, respectively, while assay sensitivity was 0.25 pg/ml. The CV values for progesterone were 12% and 14%, while the lower limit of sensitivity was 10 pg/ml.

Data analyses—Menstrual-cycle days were determined following procedures outlined in Edler *et al.* (2007). As before, the first day of self-reported menstrual bleeding was designated as 'day +1' and the previous day as 'day -1'. Inspection of hormone profiles confirmed the presence of ovulation in each participant, as evidenced by a peak in estradiol 12–18 days before the onset of menses, and a peak in progesterone 5–9 days prior to menses.

Statistical analyses of associations between ovarian hormones and binge eating were modeled after those used by Edler *et al.* (2007). Five-day rolling averages were calculated and converted to Z scores based on each participant's overall mean and standard deviation across data collection. Rolling averages are commonly used in menstrual-cycle studies of hormones (Gladis & Walsh, 1987) because they minimize random variation due to environmental circumstances, reduce the influence of hormone pulsatility and smooth the pattern of hormone variability (Kassam *et al.* 1996; Waller *et al.* 1998).

Correlations between daily Z scores for hormones and binge eating were then calculated separately for each participant. We used partial correlations for these analyses to control for levels of negative affect and the effects of the other ovarian hormone (i.e. progesterone in correlations between estradiol and binge eating; estrogen in correlations of progesterone and binge eating). Within-subject, partial correlations were converted to Fisher's Z scores and combined across participants. The combined Fisher Z score was then converted back to a Pearson's *r* value to yield an overall effect size (Rosenthal & Rosnow, 1991). One-tailed significance levels were converted to Z scores and then assigned negative values if the direction of the effect was opposite to the predicted direction. Adjusted Z scores were summed and divided by the square root of the total number of observations. The resulting Z score was converted to a *p* value to yield an overall significance level. Thus, the cycle-day results represent combined correlation effect sizes and significance levels from within-subject analyses. This approach has the advantage of providing readily interpretable effect sizes for associations among variables of interest.

One complicating factor was that co-twin observations are not independent of each other. Thus, although twins were treated as individuals rather than pairs in analyses, the inclusion of co-twins results in a correlated data structure. Non-independence could be addressed using multi-level models (MLM) or similar approaches; however, the very small number of twin pairs makes implementation of such models difficult (Kenny *et al.* 2006). Thus, we instead calculated correlations using all twins (i.e. those from complete and incomplete pairs), and then recalculated them after randomly selecting one co-twin from each complete twin pair. The

second set of analyses ensures that results obtained in the full sample are not unduly influenced by the inclusion of twin pairs.

Results

Study 1

Reliability—All measures demonstrated adequate internal consistency across days, with mean α values greater than 0.70 (Table 1). Repeated-measures ANOVA using pairwise comparisons of measures' reliabilities over days revealed that the DEBQ Emotional Eating subscale was superior to four of the five other measures, including the TFEQ Disinhibition [mean difference 0.16, 95% confidence interval (CI) 0.13–0.18], TFEQ Hunger (mean difference 0.22, 95% CI 0.18–0.26), DEBQ External Eating (mean difference 0.04, 95% CI 0.001–0.08) and EDEQ Eating Concerns (mean difference 0.18, 95% CI 0.11–0.26) subscales. No significant difference emerged in comparisons between the DEBQ Emotional Eating and EDI Bulimia subscales with regard to mean internal consistency across days (mean difference 0.02, 95% CI –0.02 to 0.06); however, the EDI Bulimia subscale also did not differ significantly from the DEBQ External Eating subscale (mean difference 0.02, 95% CI –0.01 to 0.05).

We examined test–retest reliability by calculating the correlation between scores on the same measure for adjacent days within the same menstrual-cycle phase. We chose the day before and the first day of menstruation because both days are assigned to the premenstrual phase, are characterized by similar hormone profiles, and could be reliably identified for each participant. All correlations were statistically significant and were associated with a large effect size. As with internal consistency, the DEBQ Emotional Eating subscale demonstrated the highest test–retest reliability (Table 1).

Menstrual-cycle phase comparisons—Of the measures employed, the DEBQ Emotional Eating subscale demonstrated the most robust pattern of menstrual-phase fluctuation (see Table 2). Contrast analyses indicated that scores increased in the mid-luteal/premenstrual phases compared with the follicular/ovulatory phases. These findings are consistent with mid-luteal/premenstrual exacerbation of binge eating found previously for women with BN (Gladis & Walsh, 1987; Price *et al.* 1987; Lester *et al.* 2003; Edler *et al.* 2007). The remaining scales all demonstrated greater scores in the premenstrual compared with the follicular phase, as would be expected; however, no other scales conformed to *a priori* hypotheses regarding levels across all four cycle phases.

Although the name of the DEBQ Emotional Eating scale implies that findings might be attributable to premenstrual changes in mood, negative affect did not demonstrate mid-luteal/premenstrual exacerbation [$t(36)=0.87, p=0.39$]. Instead, increases in DEBQ Emotional Eating scores occurred before premenstrual increases in negative affect were observed (see Fig. 1).

Study 1 discussion—Findings supported our ability to modify existing measures of disordered eating to examine mid-luteal/premenstrual exacerbation of binge eating in a community-based sample. Across various measures examined, the DEBQ Emotional Eating scale demonstrated superior psychometric properties with regard to internal consistency and test–retest reliability. In addition, the DEBQ Emotional Eating scale demonstrated predicted associations between menstrual-cycle phase and binge level. Nonetheless, we did not directly assess ovarian hormone levels. Thus, we cannot confirm that all participants ovulated. Anovulatory cycles are not associated with mid-luteal increases in food intake (Barr *et al.* 1995), and this may have constrained our ability to find predicted associations with this small sample. A need for direct assessment of ovarian hormones was addressed in study 2.

Study 2

As shown in Fig. 2, increases in binge eating were associated with decreases in estradiol and increases in progesterone. These impressions were confirmed by the partial correlations (see Table 3) which indicated a significant inverse association between estradiol and binge eating, and a significant positive association between progesterone and binge eating. Correlations were remarkably similar to a previous study of women with BN by Edler *et al.* (2007) (i.e. estradiol–binge eating $r=-0.25$, progesterone–binge eating $r=0.37$ in the previous study). Notably, our correlations remained essentially unchanged when recalculated using one randomly selected twin from each twin pair (estradiol–binge eating $r=-0.27$, $p<0.001$; progesterone–binge eating $r=0.22$, $p<0.01$).

Study 2 discussion—Findings from study 2 extend those of study 1 by directly assaying ovarian hormone concentrations and examining their longitudinal associations with binge eating. Results confirm that changes in ovarian hormones predicted changes in DEBQ Emotional Eating scores across the menstrual cycle in a non-clinical sample of women.

Discussion

Our findings replicate and extend previous work examining associations between binge eating and ovarian hormones in women with BN (Lester *et al.* 2003; Edler *et al.* 2007). Using two independent samples of women, our data show that: (1) the DEBQ Emotional Eating subscale is a reliable and sensitive measure of menstrual-cycle changes in binge eating appropriate for use in non-patient samples; (2) binge-eating scores show significant longitudinal associations with estrogen and progesterone across the menstrual cycle in a non-clinical sample.

Our studies are the first to examine associations between changes in binge eating and ovarian hormones in a non-clinical sample. Findings provide strong support for the feasibility of examining these associations in community-based samples using the modified DEBQ Emotional Eating scale as a daily binge-eating measure. Indeed, the magnitude of associations between binge eating, menstrual-cycle phase and ovarian hormones was remarkably similar to those found in previous studies of women with BN (Lester *et al.* 2003; Edler *et al.* 2007). In addition, all participants in study 2 demonstrated ovulatory cycles. Thus, unlike studies of women with BN (Edler *et al.* 2007), no participants were lost to analyses due to menstrual irregularities.

Several limitations of our studies should be noted. First, sample sizes were small. However, sample sizes were very similar to those in previous studies of women with BN ($n=8-11$; Lester *et al.* 2003; Edler *et al.* 2007) utilizing similar designs in which daily saliva samples were collected. Small samples would be expected to increase error and decrease the probability of replication and detection of significant effects. That was clearly not the case for our studies. Thus, despite small sample sizes, four independent studies [two clinical (Lester *et al.* 2003; Edler *et al.* 2007), two nonclinical] obtained the same pattern of results. Future studies should replicate our findings with larger samples.

Second, although we controlled for baseline BMI and daily changes in negative affect, we could not rule out other third variables that may account for ovarian hormone–binge eating associations. For example, it is possible that dieting could cause reduced estrogen levels and increased binge eating. Notably, acute (i.e. 5 days) dieting would not affect these relationships, as acute dieting and weight loss have no effect on estradiol levels (Loucks & Heath, 1994). By contrast, sustained (e.g. ≥ 6 weeks) dieting has been shown to reduce estradiol levels (Pirke *et al.* 1985). However, this would result in blunted fluctuations in estradiol levels over the menstrual cycle and would have reduced our ability to detect significant within-person associations between changes in ovarian hormones and changes in binge levels.

Third, a subgroup of participants in both studies dropped out or provided data that could not be used in analyses. Our studies were designed to establish the presence of associations between binge eating and ovarian hormones in non-clinical samples. Consequently, we used small, convenience samples of women. Our studies therefore do not address the feasibility of recruitment of representative samples. Although data from study 2 suggest that women who completed the study were not significantly different on key study variables (e.g. binge eating) from twins who did not, completers exhibited more eating and mood symptoms. Thus, additional research is needed to confirm the representativeness of our samples and generalizability of our findings. Future studies may benefit from shorter data collections (e.g. 45 days), as studies using a shorter period of data collection have demonstrated higher completion rates (range=77–89%) (Lester *et al.* 2003; Edler *et al.* 2007).

In addition to the need for large-scale studies, there is a particular need for prospective twin studies to replicate and extend our findings by examining etiologic associations between ovarian hormones and binge eating. Indirect evidence suggests that ovarian hormones may contribute to the genetic diathesis of binge eating, as genetic risk for disordered eating becomes prominent during estrogen activation at puberty (Klump *et al.* 2003, 2007a, b). Ovarian hormones are potent regulators of gene transcription (Ostlund *et al.* 2003) in neurotransmitter systems that are implicated in the etiology of disordered eating (e.g. serotonin) (Klump & Culbert, 2007). Associations between ovarian hormones and binge eating may therefore reflect genomic effects on the production of neurotransmitters, their receptors, or their signal transduction mechanisms. Additional twin and molecular genetic research is needed to explore these possibilities.

References

- Asarian L, Geary N. Modulation of appetite by gonadal steroid hormones. *Philosophical Transactions of the Royal Society – Biological Science* 2006;361:1251–1263.
- Barr SI, Janelle KC, Prior JC. Energy intakes are higher during the luteal phase of ovulatory menstrual cycles. *American Journal of Clinical Nutrition* 1995;61:39–43. [PubMed: 7825535]
- Culbert KM, Breedlove SM, Burt SA, Klump KL. Prenatal hormone exposure and risk for eating disorders: a comparison of opposite- and same-sex twins. *Archives of General Psychiatry*. in press.
- Edler C, Lipson SF, Keel PK. Ovarian hormones and binge eating in bulimia nervosa. *Psychological Medicine* 2007;37:131–141. [PubMed: 17038206]
- Fairburn CG, Beglin SJ. Assessment of eating disorders: interview or self-report questionnaire? *International Journal of Eating Disorders* 1994;16:363–370. [PubMed: 7866415]
- Garner DM, Olmsted MP, Polivy J. Development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. *International Journal of Eating Disorders* 1983;2:15–34.
- Gladis MM, Walsh BT. Premenstrual exacerbation of binge eating in bulimia. *American Journal of Psychiatry* 1987;144:1592–1595. [PubMed: 3688285]
- Jasienska G, Ziolkiewicz A, Ellison PT, Lipson SF, Thune I. Large breasts and narrow waist indicate high reproductive potential in women. *Proceedings of the Royal Society of London Series B* 2004;271:1213–1217. [PubMed: 15306344]
- Kassam A, Overstreet JW, Snow-Harter C, De Souza MJ, Gold EB, Lasley BL. Identification of anovulation and transient luteal function using a urinary pregnanediol-3-glucuronide ration algorithm. *Environmental Health Perspectives* 1996;104:408–413. [PubMed: 8732951]
- Kenny, DA.; Kashy, DA.; Cook, WL. *Dyadic Data Analysis*. The Guilford Press; New York: 2006.
- Klump KL, Burt SA. The Michigan State University Twin Registry (MSUTR): genetic, environmental and neurobiological influences on behavior across development. *Twin Research and Human Genetics* 2006;9:971–977. [PubMed: 17254439]
- Klump KL, Burt SA, McGue M, Iacono WG. Changes in genetic and environmental influences on disordered eating across adolescence: a longitudinal twin study. *Archives of General Psychiatry* 2007a;64:1409–1415. [PubMed: 18056549]

- Klump, KL.; Burt, SA.; Sisk, C.; Keel, PK. Estrogen moderates genetic effects on disordered eating during puberty. Paper presented at the Eating Disorder Research Society Meeting; Pittsburgh, PA. 2007b.
- Klump KL, Culbert KM. Molecular genetic studies of eating disorders: current status and future directions. *Current Directions in Psychological Science* 2007;16:37–41. [PubMed: 20539827]
- Klump KL, McGue M, Iacono WG. Differential heritability of eating attitudes and behaviors in prepubertal versus pubertal twins. *International Journal of Eating Disorders* 2003;33:287–292. [PubMed: 12655625]
- Klump KL, Perkins P, Burt SA, McGue M, Iacono WG. Puberty moderates genetic influences on disordered eating. *Psychological Medicine* 2007c;37:627–634. [PubMed: 17335640]
- Lester NA, Keel PK, Lipson SF. Symptom fluctuation in bulimia nervosa: relation to menstrual-cycle phase and Cortisol levels. *Psychological Medicine* 2003;33:51–60. [PubMed: 12537036]
- Loucks AB, Heath EM. Dietary restriction reduces luteinizing hormone (LH) pulse frequency during waking hours and increases LH pulse amplitude during sleep in young menstruating women. *Journal of Clinical Endocrinology and Metabolism* 1994;78:910–915. [PubMed: 8157720]
- Ostlund H, Keller E, Hurd YL. Estrogen receptor gene expression in relation to neuropsychiatric disorders. *Annals of the New York Academy of Sciences* 2003;1007:54–63. [PubMed: 14993040]
- Pirke KM, Schweiger U, Lemmel W, Krieg JC, Berger M. The influence of dieting on the menstrual cycle of healthy young women. *Journal of Clinical Endocrinology and Metabolism* 1985;60:1174–1179. [PubMed: 3923022]
- Price WA, Torem MS, DiMarzio LR. Premenstrual exacerbation of bulimia. *Psychosomatics* 1987;28:378–379. [PubMed: 3438453]
- Roberts S, Kenney NJ, Mook DG. Overeating induced by progesterone in the ovariectomized, adrenalectomized rat. *Hormones and Behavior* 1972;3:267–276. [PubMed: 4680513]
- Rosenthal, R.; Rosnow, RL. *Essentials of Behavioral Research: Methods and Data Analysis*. 2nd. McGraw-Hill; New York: 1991.
- Shirtcliff EA, Granger DA, Schwartz EB, Curran MJ, Booth A, Overman WH. Assessing estradiol in biobehavioral studies using saliva and blood spots: simple radioimmunoassay protocols, reliability, and comparative validity. *Hormones and Behavior* 2000;38:137–147. [PubMed: 10964528]
- Stunkard AJ, Messick S. The three-factor eating questionnaire to measure dietary restraint, disinhibition, and hunger. *Journal of Psychosomatic Research* 1985;29:671–680.
- Van Strien T, Frijters JE, Bergers GP, Defares PB. The Dutch Eating Behavior Questionnaire (DEBQ) for assessment of restrained, emotional, and external eating behavior. *International Journal of Eating Disorders* 1986;5:295–315.
- Varma M, Chai JK, Meguid MM, Laviano A, Gleason JR, Yang ZJ, Blaha V. Effect of estradiol and progesterone on daily rhythm in food intake and feeding patterns in Fischer rats. *Physiology and Behavior* 1999;68:99–107. [PubMed: 10627068]
- Wade GN. Some effects of ovarian hormones on food intake and body weight in female rats. *Journal of Comparative and Physiological Psychology* 1975;88:183–193. [PubMed: 1120795]
- Waller K, Swan SH, Windham GC, Fenster L, Elkin EP, Lasley BL. Use of urine biomarkers to evaluate menstrual function in healthy premenopausal women. *American Journal of Epidemiology* 1998;147:1071–1080. [PubMed: 9620051]
- Wardle J. Eating style: a validation study of the Dutch Eating Behaviour Questionnaire in normal subjects and women with eating disorders. *Journal of Psychosomatic Research* 1987;31:161–169. [PubMed: 3473234]
- Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology* 1988;54:1063–1070. [PubMed: 3397865]

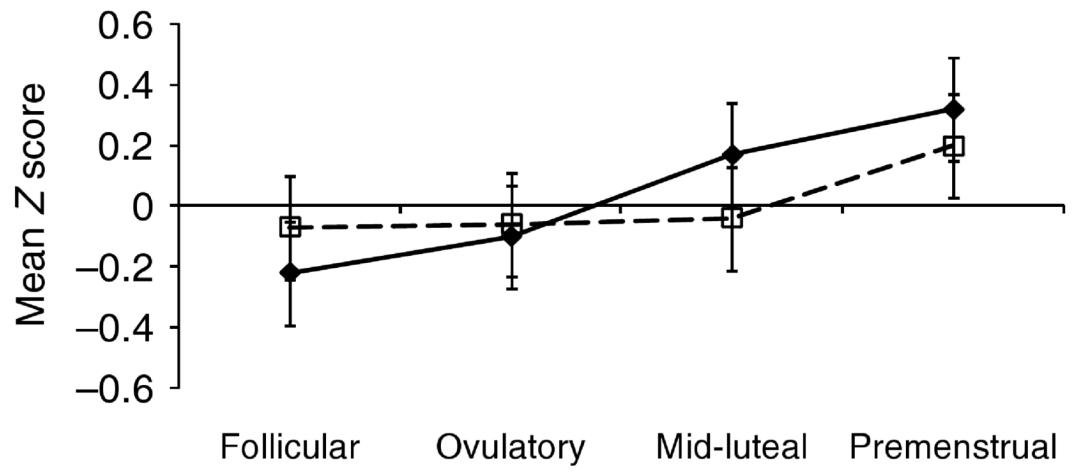


Fig. 1. Levels of Emotional Eating (—◆—) and Negative Affect (- -□- -) across menstrual-cycle phases in study 1 ($n=10$ women).

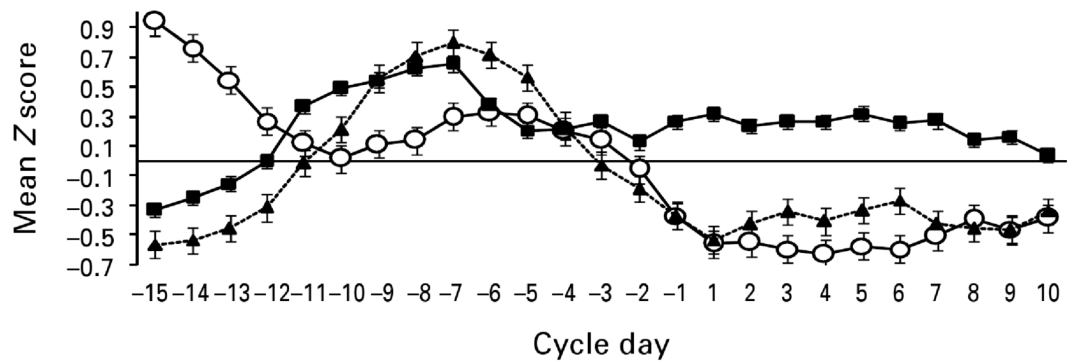


Fig. 2. Levels of binge eating (—■—), estradiol (—○—) and progesterone (—▲—) across the menstrual cycle in study 2. Mean Z scores represent 5-day rolling averages calculated within subjects, then averaged across participants ($n=9$ female twins).

Table 1
Internal consistency and test–retest reliability of binge-eating measures in study 1 (n=10 women)

Scale	Internal consistency: Cronbach's α , mean (s.d.)	Test–retest reliability	
		Spearman–Brown coefficient	ICC
DEBQ Emotional Eating	0.93 (0.12)	0.90	0.87**
DEBQ External Eating	0.90 (0.07)	0.83	0.83**
EDEQ Eating Concerns	0.76 (0.20)	0.90	0.84**
EDI Bulimia	0.92 (0.08)	0.90	0.84**
TFEQ Disinhibition	0.79 (0.10)	0.73	0.69*
TFEQ Hunger	0.72 (0.11)	0.80	0.74*

s.d., Standard deviation; ICC, Intra-class correlation; DEBQ, Dutch Eating Behavior Questionnaire; EDEQ, Eating Disorders Examination Questionnaire; EDI, Eating Disorders Inventory; TFEQ, Three Factor Eating Questionnaire.

*
 $p < 0.05$,

**
 $p < 0.01$.

Table 2
Mean Z score by cycle phase and contrast analyses of mid-luteal/premenstrual increase in scale scores in study 1 (n=10 women)^a

Scale	Follicular	Ovulatory	Mid-luteal	Premenstrual	t(36)	p
DEBQ Emotional Eating	-0.22	-0.10	0.17	0.32	2.30	0.03
DEBQ External Eating	-0.39	0.20	0.11	0.27	1.48	0.15
EDEQ Eating Concerns	-0.10	0.01	0.04	0.09	0.66	0.52
EDI Bulimia	-0.16	0.29	-0.04	0.03	-0.39	0.70
TFEQ Disinhibition	-0.23	0.16	0.05	0.14	-0.16	0.88
TFEQ Hunger	-0.07	-0.07	-0.04	0.20	0.94	0.35

DEBQ, Dutch Eating Behavior Questionnaire; EDEQ, Eating Disorders Examination Questionnaire; EDI, Eating Disorders Inventory; TFEQ, Three Factor Eating Questionnaire.

^aContrast analyses compare scores in the follicular/ovulatory phases with those in the mid-luteal/premenstrual phase.

Table 3
Correlations between ovarian hormones and binge eating in study 2(n=9)^a

	Twin	df	Correlation, <i>r</i>	
			Estradiol ^b	Progesterone ^c
All subjects	-	8	-0.23***	0.27***
Family 1	A		0.07	-0.21
	B		0.43	0.88
Family 4	A		-0.27	-0.33
	B		-0.37	0.13
Family 5	A		-0.24	0.31
	B		-0.43	-0.006
Family 7	A		-0.63	0.49
Family 8	A		-0.36	0.33
	B		-0.12	0.34

df, Degrees of freedom.

^a Although twins were treated as individuals in analyses, they are grouped by family here to allow examination of within-family similarities in correlations. Families 1, 5, 7 and 8 are monozygotic twin pairs, while family 4 is a dizygotic twin pair.

^b Values are partial correlations that control for negative affect [measured with the Positive and Negative Affect Schedule (PANAS)] and progesterone.

^c Values are partial correlations that control for negative affect (measured with the PANAS) and estradiol.

 $p < 0.001$.