Pubertal changes in hormone levels and depression in girls

A. Angold, E. J. Costello, A. Erkanli and C. M. Worthman

From the Department of Psychiatry and Behavioral Sciences, Duke University Medical Center; and Department of Anthropology, Emory University, NC, USA

ABSTRACT

Background. Throughout their reproductive years, women suffer from a higher prevalence of depression than men. Before puberty, however, this is not the case. In an earlier study, we found that reaching Tanner Stage III of puberty was associated with increased levels of depression in girls. This paper examines whether the morphological changes associated with puberty (as measured by Tanner stage) or the hormonal changes underlying them are more strongly associated with increased rates of depression in adolescent girls.

Methods. Data from three annual waves of interviews with 9 to 12-year-olds from the Great Smoky Mountains study were analysed.

Results. Models including the effects of testosterone and oestradiol eliminated the apparent effect of Tanner stage. The effect of testosterone was non-linear. FSH and LH had no effects on the probability of being depressed.

Conclusions. These findings argue against theories that explain the emergence of the female excess of depression in adulthood in terms of changes in body morphology and their resultant psychosocial effects on social interactions and self-perception. They suggest that causal explanations of the increase in depression in females need to focus on factors associated with changes in androgen and oestrogen levels rather than the morphological changes of puberty.

INTRODUCTION

The excess (more than two to one) of females with unipolar depression diagnosis (Weissman et al., 1996; Bebbington, 1998; Bebbington et al., 1998) does not appear until adolescence (Angold et al., 1998). Competing explanations for this phenomenon variously emphasize negative effects of early puberty on girls, negative effects of changes in cognitive processing and body image, changes in rates of life events and altered sensitivity to life events or other stressors, and direct effects of endocrine changes on the brain (Nolen-Hoeksema & Girgus, 1994; Bebbington, 1996, 1998). In previous work (Angold et al., 1998), we found that the transition to Tanner stage III of puberty was associated with a sharp increase in rates of DSM-IV unipolar depressions in girls, but that neither the timing nor reactivity of this transition, nor the appearance of menarche or its timing, had any significant effect (see Angold & Worthman, 1993; Angold et al., 1998 for reviews of the literature on morphological pubertal status and timing and depression). Tanner stage had a much larger effect on rates of depression than did age, suggesting that changes in rates of depression in adolescence were specifically related to the physical changes of puberty, rather than simply indexing a 'time of life' when the prevalence of depression rates in girls.

Associations of depression with concentrations of gonadal hormones and gonado-trophins have been reported in both adults and adolescents, but it remains unclear whether such
Associations arise from central-organizational changes that initiate and drive puberty, from endocrinological changes in endocrine milieu, or from psychological and social effects of the visible pulsations of puberty (Young & Korsten, 1998). Abnormalities in gonadotrophin secretion and of gonadotrophin-releasing hormone have been found in depressed adult patients (see, Tolis & Stefani, 1983) while follicle stimulating hormone (FSH) levels have been found to correlate positively with negative affect in normal men (Houwer, 1979). Oestrone levels have been implicated in menstrual cycle-associated mood changes which have, in turn, been linked with increased susceptibility to major depression (reviewed in Warne & et al., 1991). Of course, the premenstrual period of the menstrual cycle is a time of relatively low oestrone levels (Brodzicki, 1976; Warne & et al., 1991). This may appear to contradict the notion that the increases seen in oestrone levels in adolescence might be associated with depression, but it should be remembered that prior to the premenstrual oestrone drop, there is a substantial elevation of oestrone that peaks with ovulation. It could be that perimenstrual dysphoria is actually a delayed result of previous high oestrone levels. High oestrone oral contraceptives have long been associated with depression as a side effect, but on the other hand increases in oestradiol have been found to improve depressive symptoms in hypo-oestrogenic women (Montgomery & et al., 1987). Overall, the findings from this literature are rather inconsistent and difficult to interpret (Bancroft, 1993). In later life (after age 55), the female rates of depression diminishes, mostly because of falling rates in women at a time when their oestrone levels are again low (Bebbington, 1996; Bebbington & et al., 1998).

Studies of adolescence lend some weight to the idea that hormonal changes in puberty are 'bad' for girls, but 'good' for boys as far as mood is concerned, but no clear picture has emerged so far. The NIMH study of puberty and psychopathology (Nourmanna et al., 1987a, b; Sussman et al., 1987a, b) found negative associations between testosterone: oestradiol ratios, sex hormone binding globulin, and androstenedione concentration and negative emotional tone in boys. These workers also reported an association of early maturation (measured by oestradiol and testosterone: oestradiol ratios) with reduced negative emotional tone in boys, but more negative emotional tone in girls. Adrenal androgen levels correlated with negative emotional tone in boys, but not girls, while the opposite was the case for FSH (Sussman et al., 1985; Nourmanna & et al., 1986; Brooks-Gunn & Warren, 1989) found that negative affect increased in 16- to 14-year-old girls during rapid oestrone rise. A 1-year follow-up of 72 girls (Puckoff & et al., 1993) found a significant linear relationship between oestradiol level at time 1 and depression 1 year later according to one depression scale, but no such effect in relation to two other depression scales. Administration of sex steroids to adolescents with hypogonadism had little effect on mood or behavioral problems in either sex (Sussman & et al., 1998).

None of these hormonal studies had sufficient power to tease apart the possible contributions of age itself, the indirect psychosocial impacts of morphological pubertal status, and the direct impact of the different groups of hormones that change at puberty. In addition, all of them used only depression scale scores, rather than interview-based diagnoses. Yet, it is a notable feature of the adolescent depression literature that, while diagnostic studies have consistently shown that rates of depression are higher in adolescents girls than adolescent boys, studies based on depression scale scores have produced quite inconsistent results (Kazdin et al., 1988; Kovacs, 1983, 1982; Heise & Matson, 1984; Fanti et al., 1985; Fitch et al., 1985; Haley et al., 1985; Bartell & Reynolds, 1986; Steiner et al., 1986; Huntley et al., 1987; Doerfler et al., 1988; Gates et al., 1988; Reinherz et al., 1990).

Our aim was to overcome some of these problems by examining the relationships of testosterone (which, in adolescent girls, is predominantly produced by the adrenal cortex), FSH and LH (from the pituitary) and oestradiol (predominantly from the ovary) with rates of depression, over multiple observations, in a relatively large representative general population sample of girls assessed with a psychiatric diagnostic interview. Having already shown that Tanner stage was more strongly associated with depression than was age, our strategy here involves comparisons of the sizes of effect of the four hormones and Tanner stage on depression rates. The analyses presented test the hypothesis
that the inclusion of hormonal measures in models with Tanner staging will reduce or eliminate the apparent effect of Tanner stage on depression, indicating that we are dealing with a pubertal effect not psychosocially mediated by morphological development. If this proves to be the case then we can go on to ask whether it appears to be an effect depending on the peripheral hormonal products of puberty (indexed by testosterone and oestradiol), or whether it is more directly related to central nervous systems changes (as indexed by FSH or LH).

**METHOD**

The data came from the Great Smoky Mountains Study (GSMS) of children and adolescents. A detailed account of the study design and instrumentation used can be found in an earlier paper (Costello et al. 1996). We present a summary here.

**Sampling frame**

A representative sample of 450, 9-, 11- and 13-year-olds, recruited (through the Student Information Management System of the public school systems of 11 counties in western North Carolina) was selected using a household equal probability design. At close as possible to the child’s birthday, a screening questionnaire was administered to a parent (usually the mother), by telephone or in person. This consisted of 55 questions from the Child Behavior Checklist about the child’s behaviour (externalizing) problems, together with some basic demographic and service use questions. All children scoring above a predetermined cut-off score of 50 (designed to include about 25% of the population) on the behavioural questions, plus a 1-in-10 random sample of those scoring below the cut-off, were recruited for the longitudinal study. Eighty percent of eligible families agreed to participate in the interview at least once (1073 of 1346).

Shortly after being screened, eligible children and one of their parents were interviewed. They were re-interviewed, using very similar assessment protocols, 1, 2 and 3 years later. Interviews were conducted between 1992 and 1996. Between 80% and 94% of the sample participated in each wave. In this paper we present data from the first three waves of data collection: because only for these waves did we have funding to complete hormone assays. The sample considered here, therefore, consists of 465 girls on whom we had a total of 1783 interview observations from the first three waves of data collection.

Because sexual development is a sensitive topic, we showed the Tanner stage assessment to parents before giving it to the children, and specifically asked permission to use it. At each wave, between 2.6% and 7.5% of parents refused to have the scale administered to their female children. However, refusal to compute the Tanner stage assessment was not significantly related to depression scores or diagnosis, so it seems unlikely that this additional source of missing data was a source of bias in the results.

At each wave, between 22.8% and 24.9% of female participants refused to give blood for hormone measurements (separate consent for the finger prick procedure used here was sought). Again, there was no significant relationship between depression status and missing hormone data, so again it seems unlikely that the results will be substantially biased by missing data in this area. Table 1 shows the numbers of girls interviewed at each wave, and the numbers for whom Tanner stage and hormonal data were available.

**Measures**

**Psychiatric symptoms and disorders.**

Children and parents were interviewed using the Child and Adolescent Psychiatric Assessment (CAPA) (Angold et al. 1995), which generates Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) (APA. 1994) diagnoses and a range of symptom scale scores, measures of functional impairment, and family burden (for details, see Costello et al. 1996). Diagnoses were generated from symptom codings by
computer algorithms. If either parent or child reported a symptom as present in the past 3 months, it was counted towards the relevant CAPA/DSM-IV scale score or diagnosis. Thus 3 months "primary period" was selected rather than, say, a 1-year lifetime period, because shorter recall periods are associated with more accurate recall (see Angold et al. 1996). We considered three depression diagnoses: DSM-IV major depressive episode, dysthymia, and depression not otherwise specified (NOS). The last of these diagnostic categories comprised individuals who met the DSM-IV experimental criteria for minor depressive disorder (APA, 1994, p. 719).

Pubertal morphological status

Self-ratings of pubertal morphological status based on the standard Tanner staging system (Tanner, 1962) were performed with the aid of schematic drawings of secondary sexual characteristics (breasts and pubic hair). Such ratings correlate well with physical examination, based on Tanner stages (Duke et al. 1980; Morris & Ud key, 1980; Frankowski et al. 1987; Dorn et al. 1990; Schlossberger et al. 1992). Each child was provided with sex-appropriate schematic drawings and requested to rate herself on each dimension. Both self-ratings were averaged to yield a single individual score (ranging from 1—prepubertal, to 4—adult level of development).

Blood spot collection

Hormone samples were obtained at the beginning of the interview session, as follows. Two finger-prick samples were collected at 20-min intervals, applied to specifically prepared paper, immediately refrigerated prior to drying, and express shipped (without refrigeration) to the laboratory within 2 weeks of collection. Samples were then stored at −20°C until they were assayed.

Hormonal assays

Blood spot FSH and LH were measured using modifications of commercially available fluorimmunometric kits for assay of these hormones in serum or plasma (DELFIA; Wallac, Inc., Gaithersburg, MD). The blood spot testosterone (T) and estradiol (E2) assays are modifications of commercially available serum/plasma radioimmunoassay kits (Biax, South Portland, ME; Parsert, Santa Monica, CA; DSL; Webster, TN, respectively). Complete details of protocol, validation, performance, sample stability, and comparability to plasma or serum values for each blood spot assay are provided elsewhere (Worthman & Stallings, 1997a, b). To minimize effects of pulsatility, hormone values for each observation were taken as the average of the two blood spot samples. Formulas for conversion of blood spot assay values to serum/plasma equivalents are given in Worthman & Stallings (1997a). All values in the present report represent plasma/serum equivalents and hence are comparable to the extant literature based on that medium.

In brief, assay performance characteristics are as follows. Assay sensitivity, defined as the dose with 95% confidence for discrimination from zero dose, is: FSH 0.13 IU/L, LH 0.26 IU/L, T 63 ng/dl, E2 95 ng/ml. Intrassay coefficients of variation for low (FSH 2.7 IU/L, LH 1.5 IU/L, T 33 ng/dl, E2 246 pg/ml), medium (FSH 31.3 IU/L, LH 6.4 IU/L, T 299.2 ng/dl, E2 452 pg/ml), and high (FSH 38.0 IU/L, LH 19.5 IU/L, T 548.2 ng/dl, E2 124.5 pg/ml) controls are: low—FSH 7.8%, LH 10.9%, T 7.6%, E2 14.4%; medium—FSH 5.3%, LH 7.7%, T 9.3%, E2 74%; high—FSH 9.9%, LH 3.5%, T 7.0%, E2 40.4%. Interassay CVs for the same controls are: low—FSH 9.2%, LH 11.6%, T 13.9%, E2 150%; medium—FSH 8.6%, LH 7.2%, T 12.5%, E2 116%; high—FSH 5.9%, LH 7.8%; T 11.8%, E2 8.4%.

Timing of sample collection

Most of the hormones included in this study exhibit altered diurnal variation during puberty. Night-time pulsatile gonadotropin release escalates in pre- and very early pubertal and later spreads throughout the day. Adrenal androgens also show diurnal variation, with peak levels in the morning and an evening trough, leading to around 50% diurnal variability in testosterone (Mikulecky et al. 1995). Estradiol release is phase shifted in early puberty, showing a daytime acrophase and midnight nadir (Goij, 1993; Neurohr et al. 1996). Diurnal variation in E2 abates within one year after menarche.

Collection times for blood spots were constrained by the times that were convenient for families to be interviewed in their homes. Thus,
the range of times for sample collection was 8 a.m. to 11 p.m. However, 76.5% of samples were collected between 1 p.m. and 7 p.m.; that is, during periods when the diurnal variation curve is relatively flat. Nonetheless, it might seem, at first sight, as though it would be a good idea to apply some correction to "standardize" all the hormone values to a particular time of day. However, the changes described in the previous paragraph indicate that the pattern of diurnal variation is highly dependent on level of development, so no simple correction factor can be applied for any of the four hormones under consideration. Complex correction schemes would have to be dependent upon the level of each hormone, and would invariably result in some individuals being mis-correlated. We, therefore, decided not to attempt any corrections for the time of day of sample collection, recognizing that this decision would necessarily result in increased error variation in any regression analysis compared with a situation in which we had been able to collect all samples at a particular time of day. Such errors will tend to lead to underestimates of the effects of the hormones in comparison with Tanner stage.

**ANALYTICAL STRATEGY**

The presence of repeated measures and screen stratified sampling required the use of weighted analyses to generate unbiased population parameter estimates and of "sandwich" type variance corrections (Diggle et al. 1994; Pickles et al. 1995) to produce appropriate confidence intervals and $P$ values. These were obtained using mixed effects hierarchical linear logistic models (HLM; Diggle et al. 1994) implemented through the SAS macro GLIMMIX, as in our previous work on this subject (Angold et al. 1998).

There was no association between time of day and depression status (OR = 1.006, $P = 0.95$), so there was no point in controlling for time of day of sample collection in the analyses.

**RESULTS**

Univariate distributions of FSH, LH, testosterone and estradiol

Table 2 gives the means and standard deviations for the distributions of each of the hormones considered here by Tanner stage. Predictably, the mean levels of each hormone increase with increasing Tanner stage.

**Correlations among hormone levels and Tanner stage**

As can be seen from Table 2, the correlations among hormone levels and pubertal stage were often high. This was expected, since puberty consists of a physiologically ordered set of endocrine changes with certain hormones actually driving the production of others (e.g. LH controlling estradiol output). These levels of correlation provide support for the validity of our measure of Tanner stage, because the hormones with direct end-organ effects on the components of the Tanner staging system (testosterone and estradiol) have substantially higher correlations with Tanner stage than do the trophic hormones (FSH and LH), the effects of which are manifested indirectly through hormones such as estradiol and oestrogen.

These correlation strengths are similar to, or somewhat higher than, those reported in other studies involving directly observed Tanner stage (Lee & Migeon, 1975; Lee et al. 1975; Apfel, 1980). Our rather stronger correlations may reflect greater assay sensitivity in this study than in earlier reports.

**Recalling continuous measures of Tanner stage and hormone levels**

Our key question concerns the relative sizes of effects of Tanner stage and hormone levels on the probability of being depressed. To enhance the comparability of the effects of the various hormones and Tanner stage we recalled each of these predictors to a minimum value of 0 and a

---

**Table 2: Mean levels of FSH, LH, testosterone and estradiol by Tanner stage (in parentheses)**

<table>
<thead>
<tr>
<th>Tanner Stage</th>
<th>FSH Mean (±SD)</th>
<th>LH Mean (±SD)</th>
<th>Testosterone Mean (±SD)</th>
<th>Estradiol Mean (±SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>24.4 (±6.0)</td>
<td>6.5 (±3.5)</td>
<td>17.6 (±9.5)</td>
<td>2.8 (±1.2)</td>
</tr>
<tr>
<td>II</td>
<td>32.1 (±4.1)</td>
<td>9.3 (±4.5)</td>
<td>23.5 (±12.7)</td>
<td>4.3 (±2.1)</td>
</tr>
<tr>
<td>III</td>
<td>35.6 (±5.8)</td>
<td>12.0 (±5.6)</td>
<td>27.0 (±11.5)</td>
<td>4.8 (±2.4)</td>
</tr>
<tr>
<td>IV</td>
<td>51.2 (±8.1)</td>
<td>18.5 (±8.1)</td>
<td>34.6 (±11.9)</td>
<td>6.9 (±3.1)</td>
</tr>
<tr>
<td>V</td>
<td>52.4 (±10.2)</td>
<td>23.1 (±10.2)</td>
<td>40.3 (±17.5)</td>
<td>8.6 (±5.9)</td>
</tr>
</tbody>
</table>
Table 3. Correlations among Tanner stage and hormone measures

<table>
<thead>
<tr>
<th>Measurement</th>
<th>FSH</th>
<th>LH</th>
<th>Oestradiol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tanner stage</td>
<td>0.44</td>
<td>0.34</td>
<td>0.35</td>
</tr>
<tr>
<td>Testosterone</td>
<td>0.31</td>
<td>0.24</td>
<td>0.28</td>
</tr>
<tr>
<td>FSH</td>
<td>—</td>
<td>0.21</td>
<td>0.18</td>
</tr>
<tr>
<td>LH</td>
<td>—</td>
<td>—</td>
<td>0.48</td>
</tr>
</tbody>
</table>

Table 4. Effects of Tanner stage and individual pubertal hormones on depression

<table>
<thead>
<tr>
<th></th>
<th>OR (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tanner stage (optimal)</td>
<td>1.0 (0.02-24)</td>
<td>0.04</td>
</tr>
<tr>
<td>FSH</td>
<td>1.0 (0.04-2.4)</td>
<td>0.2</td>
</tr>
<tr>
<td>LH</td>
<td>1.0 (0.01-2.0)</td>
<td>0.5</td>
</tr>
<tr>
<td>Testosterone</td>
<td>0.9 (0.19-4.2)</td>
<td>0.8</td>
</tr>
<tr>
<td>Oestradiol</td>
<td>2.1 (0.1-40)</td>
<td>0.006</td>
</tr>
</tbody>
</table>

maximum value of 4 (since Tanner stage has only five available scores). The results of this procedure is that, in every case, the odds ratio (OR) represents the increase in odds of depression for an increase of one-quarter of the observed range of the predictor. Note that this recoding has no effect whatsoever on measures of the significance of effects, it simply renders the ORs for the different effects directly comparable.

Effects of individual hormone levels and Tanner stage on depression status

Table 4 shows the odds ratios (OR) associated with the effects of Tanner stage and each hormone entered as a continuous variable in bivariate regressions. As we showed before (Angold et al., 1998), there was a substantial effect of Tanner stage. However, there were larger effects of testosterone and oestradiol (particularly the former). On the other hand, the OR for FSH, and LH were smaller and not statistically significant.

Effects of Tanner stage and hormone levels when both were included in models

In the next set of models (see Table 5) we included both Tanner stage and one of the hormones that had been significantly related to depression in the preceding analyses (testosterone or oestradiol). We split Tanner stage into two groups - those below Tanner stage III and those at Tanner stage III and above, because our previous analysis had indicated that this provided the best prediction of depression status (Angold et al., 1998). The inclusion of either testosterone or oestradiol in the models caused the OR for Tanner stage to fall substantially to non-significant levels. When both were included together with Tanner stage, their effects remained significant while the OR for Tanner stage fell still further to 1.2 (Table 5 row 6). The final row of Table 5 shows that the effects of testosterone and oestradiol with Tanner stage removed from the model remained almost unchanged.

<table>
<thead>
<tr>
<th></th>
<th>OR for Tanner stage</th>
<th>OR for Hormone</th>
<th>P for Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tanner stage (optimal)</td>
<td>0.05</td>
<td>3.11 (0.78-12)</td>
<td>0.02</td>
</tr>
<tr>
<td>Tanner plus testosterone</td>
<td>1.4 (0.2-23)</td>
<td>0.01</td>
<td>1.0 (0.4-2.8)</td>
</tr>
<tr>
<td>Tanner plus FSH</td>
<td>5.5 (2.8-11)</td>
<td>0.002</td>
<td>2.8 (0.4-22)</td>
</tr>
<tr>
<td>Tanner plus LH</td>
<td>4.7 (2.8-7.7)</td>
<td>0.002</td>
<td>4.7 (4.5-23)</td>
</tr>
<tr>
<td>Tanner plus oestradiol</td>
<td>4.0 (1.9-8.6)</td>
<td>0.001</td>
<td>2.1 (1.4-3.1)</td>
</tr>
<tr>
<td>Tanner plus testosterone and oestradiol</td>
<td>3.0 (1-9.4)</td>
<td>0.05</td>
<td>3.0 (1-9.4)</td>
</tr>
<tr>
<td>(depressed not uncontrolled)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

As mentioned above, our previous work on effects of pubertal status in girls indicated that the relationship between maturation (Tanner stage) and depression was not linear, but better described by binomializing the range with a cut-off at Tanner stage III. To examine the possibility that hormone thresholds might be present, we divided the ranges of testosterone

Table 5. Models including Tanner stage testosterone and oestradiol
and estrogen into quintiles and plotted rates of depression for each quintile (see Fig. 1). Where more than 20% of the observations had values below the sensitivity of the assay, the value presented represents all those with such values plus those with values above the floor sensitivity of the assay up to the next quintile. So, for instance the point at 3rd quintile on the x axis for estradiol includes all those whose values were below the sensitivity threshold of the estradiol assay (48%) plus 12% who were below the next quintile cut-point.

Immediately apparent in Fig 1 is the strong suggestion of non-linearity in the relationship between testosterone and depression. There is a sharp jump between the third and fourth quintiles. On the other hand, the effects of estradiol appear to be reasonably linear. We then re-fit logistic regressions of Tanner stage and testosterone and estrogen on depression status, with testosterone entered as a binomial split at the 60th percentile. The results are shown in Table 6. Now the effect of Tanner stage fell to an OR of 1.0, while both testosterone and estradiol continued to have substantial significant effects. The last two rows of Table 6 shows the results of this model with the term for Tanner stage removed.

**DISCUSSION**

Having previously described a substantial effect of pubertal morphological status on depression rates in girls (Angold et al. 1998), we sought to determine whether this effect could be accounted for by changes in pubertal hormone levels alone, or whether physical status itself made an independent contribution. Obviously, the physical changes of puberty are dependent upon hormonal changes, but the nature of one's theories of why puberty leads to increased levels of depression in females depends on the nature of the most probable proximate causes of those changes. For instance, theories involving reactions to observable physical changes and negative emotions attached to changes in body image posit that morphological change is a key ingredient in the equation. In such a case, we would expect to see stronger associations with morphological status than hormone levels, since hormones are postulated to act via morphology. Conversely, theories relating depression to direct central hormonal effects indicate that the relationship between hormone levels and depression should be stronger than those between morphological development and depression because the latter is simply functioning as a marker for the underlying hormonal changes.

In this context, the first notable feature of our results is that the effect of pubertal morphological status on depression was dominated by
the addition of hormone variables to models of depression. The OR associated with pubertal status was reduced from 2.9 to 1.0 by the addition of oestrogen and testosterone levels to the model. On the other hand, the effects of testosterone and oestrogen remained consistently significant, even when Tanner stage was controlled.

The second major question concerns which pubertal hormones might be responsible. There was no evidence that the trophic hormones (FSH and LH) were directly involved. On the other hand, there was consistent support for effects of both testosterone and oestrogen in the increasing prevalence of depression. The effect of testosterone was particularly marked. Even more striking was the fact that the effects of testosterone on depression were only observed above the 60th percentiles of testosterone levels in this sample (corresponding to a level of 24.7 ng/dl).

These findings agree with previous weak evidence that negative affect is associated with higher levels of androgens and oestrogen in adolescent girls (see the Introduction to this paper), but move us forward in fulfilling such findings to the diagnosis of depression and distinguishing effects of hormone levels versus those of morphological status. Our data imply that morphological status is not the 'active ingredient' in the effects of puberty on depression in girls.

How then might puberty exert its effects? Our data suggest that more direct effects of testosterone and oestrogen are involved. One possible mechanism involves specific CNS actions by these hormones; indeed, our observed dissociation of the peripheral (Tanner stage) from the central (depression) action of oestradiol fits current understandings of steroid hormone action. A combination of tissue-specific metabolic pathways and receptor populations allows highly differentiated steroid effects both outside and within the CNS (Alonso-Solís et al. 1996; Fink et al. 1996). The two known oestrogen receptors are differentially regulated and expressed (Murphy et al. 1997) and show distinct patterns of distribution within the brain, particularly the limbic system (Laflamme et al. 1993; Oosterlaak et al. 1998). High co-localization of oestrogen receptor beta with CRF-bearing cells in the paraventricular nucleus (Laflamme et al. 1993) provides anatomical support for the stress diathesis model of depression (Heim et al. 1997, Plotsky et al. 1999).

It is also possible that, at the intracellular level, these apparent effects of both testosterone and oestradiol represent only an oestrogenic effect, since when behavioural effects of testosterone in animals have been investigated at the level of the brain receptors involved, most have proved to occur via oestrogen receptors following intracellular aromatization of testosterone to oestrogen (Hatchison et al. 1990, Rasnitsyn et al. 1990). Since oestrogen is the latest of the hormones studied here to begin to rise in puberty (see Fig. 1), it makes sense that an intracellular oestrogen effect could appear in early puberty as an effect of peripheral testosterone and only later manifest directly as an effect of peripheral oestrogen.

This is not to say that depression is simply caused by increased levels of androgens and oestrogen—that would scarcely be the case. Every physiologically normal adult female has reached and surpassed the levels of androgens and oestrogens that are associated with depression in this sample, but most will not become depressed. Even those who do will not remain depressed at all times. Our data indicate that pubertal increases in testosterone and oestrogen levels cause those hormones to surpass a threshold at which women are rendered more susceptible to depression, but there is ample evidence that, in most cases, other exogenous (such as life events or daily hassles) or endogenous (such as cognitive style) factors are required to explain the development of individual episodes of disorder.

We must consider the possibility that our results are a spurious effect of weaknesses in the study design or its measurements protocols. We have only annual single point measures of hormones uncontrolled for time of day or menstrual cycle phase. Ignoring effects of time of day on hormone levels will have served to increase the error in our models and reduce the apparent effects of the hormones. The issue of cyclicity also bears close scrutiny. The first point to note is that we are dealing here with transition from very low levels of oestrogen and testosterone, and that many of our subjects were not at the point where their cyclicity is at all marked. Secondly, insofar as our single point measures are relatively poor measures of the
overall oestrogen load manifested during any monthly cycle, or its peak level, or any other oestrogen parameter one cares to mention, that would militate against our finding any hormonal effects. In other words both of these potential problems with our measures result in conservative potential errors with respect to our findings.

A further question concerns the idea that our hormone measures could be seen as having better measurement properties than a self-report measure of Tanner stage. After all, the hormones are measured on ratio scales, while Tanner stage is measured on a five-point ordinal scale. But we must ask the question ‘better in relation to what?’ In relation to itself the coefficient of variation of the Tanner staging method we used is certainly reported to be poorer than that of the steroid assays, but this is not the point. We feel confident that this self-assessment is a better measure of physical development than are testosterone and oestrogen levels (especially in the light of possible critiques of our hormone measurement strategy mentioned above).

Theories of the increase of depression in adolescent girls posit a mechanism involving increased attention from boys who observe that a girl’s sexual maturity such as that pronounced by Salton & Magnan (1950) demand that these morphological manifestations should have more effect on depression than the hormonal changes that generate them. Our data suggest the opposite. It is also important to remember that our final models involved re-expression of the ratio scale range of testosterone to a binomial distribution — thus Tanner stage and testosterone were reduced to a 0/1 dichotomy. The result was that when both were included in models of depression, the OR for Tanner stage was 1.0, while that for testosterone was 4.3. It is hard to see how a larger difference could result from measurement artifacts.

This is a simple study, and none of its measures were by any means perfect. Nevertheless, our findings reveal no effect on depression rates related to central regulatory endocrinodisturbances (as indexed by FSH and LH levels) early or late in puberty, and contradict the hypothesis that increased depression in adolescent girls is linked to indirect psychosocial effects of body morphology. They support a link between increased rates of depression in girls and rising levels of testosterone and oestrogens in mid-to-late-puberty. We interpret these results as indicating that rising testosterone and oestrogen levels potentiate risk for depression through mechanisms that are unrelated to psychosocial effects of body morphology. We suggest that further work in this area will need to extend more closely to distinguishing possible effects of central morphological processes from those of the peripheral endocrine changes that they organize.

This work was supported by the William T Grant Foundation through faculty scholar awards to A.A., E.J.C. and C.W. and project awards 92-14499-92 and 94-14499-92. Subject interviewings was supported by grant MH49096 from the National Institute of Mental Health. Support for pilot work for this project was provided by a grant from the Locarno Foundation.

REFERENCES


Bennett, P. E. (1984). The origins of age differences in depressive symptoms: testing the exp. in the intensive research of Psychiatry & Development. Psychol, 1, 1-44.
