Puberty and depression: the roles of age, pubertal status and pubertal timing

A. ANGOLD, E. J. COSTELLO AND C. M. WORTMANN
From the Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, NC and Department of Anthropology, Emory University, GA, USA.

ABSTRACT

Background: Previous work has indicated that the 2:1 female:male sex ratio in unipolar depressive disorders does not emerge until some time between ages 10 and 15.

Methods: Data from four annual waves of data collection from the Great Smoky Mountains Study (GSMS) involving children aged nine to 16 were employed.

Results: Pubertal status better predicted the emergence of the expected sex ratio than did age. Only after the transition to mid-puberty (Tanner Stage III and above) were girls more likely than boys to be depressed. The timing of this transition had no effect on depression rates. Before Tanner Stage III, boys had higher rates of depression than girls, and the prevalence of depression appeared to fall in boys at an earlier pubertal stage than that at which it began to rise in girls. In addition, early transition to Tanner Stage III or earlier had a transient effect in reducing the prevalence of depression in boys.

Conclusions: The period of emergence of increased risk for depression in adolescent girls appears to be a relatively sharply demarcated developmental transition occurring in mid-puberty. Previously reported effects of the timing of puberty which have tended to be transient, appeared less important in decrease of risk for depression than pubertal status.

INTRODUCTION

Studies of adults suggest that the gender difference in rates of depression emerges during adolescence.

Studies of adults from several countries have emphasized the difference between men (Westman & Kleinman, 1977; Bebbington et al., 1981; Canino et al., 1987; Lee et al., 1987; Blum et al., 1984a,b; Cheng, 1989; Boz et al., 1989; Wells et al., 1989; Wittchen et al., 1992; Kesler et al., 1993, 1994; Westman et al., 1993; 1994; Blazer et al., 1994). Using retrospective data, some of these have pointed to adolescence as the time when this gender difference first appears. For example, the Epidemiological Catchment Area (ECA) studies (Burke et al., 1990) suggested that unipolar depression onset rates were equal in males and females until age 13-19, while the National Comorbidity Survey (NCS) (Kessler et al., 1993), provided evidence for the emergence of a gender difference by age 10-14.

Uncertainty about the timing of effects of depression

The child and adolescent epidemiological literature generally agrees that rates of depression are similar in prepubertal boys and girls, and that rates of depressive disorders begin to rise in girls at some time between childhood and age 15 (Rutter et al., 1976; Anderson et al., 1981; Cohen & Brooks, 1987; Kuchma et al., 1987, 1989; Bird et al., 1988; McGuire & Williams, 1984; Guze et al., 1988; Veliz et al., 1989; Fleming & Offord, 1990; McGeer et al., 1996; Nolen-Hoeksema et al., 1991; Angold & Rutter, 1992; Romer & Levitt, 1993; Lewinsohn et al., 1995). But this age range covers a lot of developmental ground. Two
longitudinal community studies suggested that female excess did not emerge until after age 13; the Dunedin Longitudinal Study (Anderson et al. 1987; McGee et al. 1992), and the New York Study (Cohot & Brooks, 1967; Vellet et al. 1988; Cohen et al. 1989a,b). On the other hand, Angold and Rutter's (1992) study of a large clinical population found that the presence or absence of referred girls with depressive disorders began to emerge at around the age of 10.

The timing of the change in sex ratios has important implications for theories about the relationship between depression and puberty. For example, if such a change actually begins by age 10, it may be that adolescence (which is in progress by middle childhood) is more relevant than puberty (Brooks-Gunn & Warren, 1992; Eccles & Bachman, 1992). It is also possible that these 'biological' changes are irrelevant, and that some other correlates of age or pubertal development are responsible for the apparent effects of puberty. For example, one large clinical study (Angold & Rutter, 1992), found that age continued to predict the presence of depression over and above effects of pubertal status. Only one of the studies considered so far involved any measures of pubertal status per se, rather than using age as a marker of development status. On the other hand, none of the work using one or more measures of pubertal status which we review next, has employed diagnostic measures of depression.

**Morphological stage and pubertal timing**

Previous studies of menarche or morphological development (secondary sex characteristics, usually measured by Tanner stages (Tanner, 1962)) have suggested that levels of these pubertal status markers per se may not be significantly related to mood or other disturbances (Olswang et al. 1988; Sosman et al. 1987; Brooks-Gunn & Warren, 1989; Paukoff et al. 1991, Angold & Rutter, 1992; Ge et al. 1996), but that their timing may be. Early puberty has been associated with problem behaviours in girls, but with good adjustment in boys (see Stattin & Magnusson 1990) for a review. Stattin & Magnusson (1990) have argued from their longitudinal study that the negative effects of early development in girls were generated by the impact of early maturity on girls' social lives. However, these effects had largely disappeared by the time the girls were in mid-adolescence, whereas the female excess of depressive disorders continues throughout adulthood (Weissman & Klerman, 1977). In the Carolina Longitudinal Study, Carstes & Canino (1994) did not find any effect of early puberty on behavioural deviance.

**Hormonal studies of puberty and psychopathology**

Studies with direct hormonal measures lend further weight to the idea that puberty is 'bad' for girls, but 'good' for boys as far as depression is concerned, but these studies have mainly (though not exclusively) focused on hormone concentrations (the analytical equivalent of morphological or menarchal status), rather than on the timing of changes in hormonal concentrations. The NIMH study of puberty and psychopathology (Nestellmann et al. 1987a,b; Sosman et al. 1987a,b) found negative associations between testosterone/estradiol ratio, sex hormone binding globulin, androstenedione concentration and negative emotional tone in boys. These workers also reported an association of early menarche (measured by oestradiol and testosterone/estradiol ratio) with reduced negative emotional tone in boys, but more negative emotional tone in girls. Adrenal androgens concentrations correlated with negative emotional tone in boys, but not girls, while the opposite was the case for FSH (Sosman et al. 1985). Overall, hormone concentrations were more strongly associated with affective behaviour in boys than in girls (Nestellmann et al. 1990). Brooks-Gunn & Warren (1989) found that negative effect increased in 10-14-year-old girls during rapid oestrogen rise, but saw no effect of the age at which increased oestrogen secretion occurred. This work has been extended to include a 1-year follow-up of 72 girls from the original sample (Paukoff et al. 1991). The authors found a significant (linear effect of oestrogen level at time 1 on self-reported symptoms of depression; 1 year later, using the Center for Epidemiologic Studies Depression Scale (CES-D) (Radloff, 1977)), but no such effect on the Depression Withdrawal Scale of the Youth Self-Report (Achenbach & Edelbrock, 1991) or on the mothers' reports of their daughters' depressive affect (Brooks-Gunn et al. 1994).
Problems with the literature to date

Three lines of work, taken together, implicate the morphological and physiological developments of puberty in the increasing depression of adolescent girls. On the other hand, it is far from clear which of the many aspects of adrenarcheal and pubertal development are really important. In particular, studies based on morphological status (Tanner stages) or timing of menarche have tended to show different results have been more important in relation to pubertal stages. There are several reasons (listed below) for this state of affairs (see Angold & Wetherall, 1991) for a more detailed discussion of these and further points.

1. Few general population studies have explicitly considered the multiple dimensions of puberty and none of the diagnostic or non-diagnostic general population studies has included adequate measures of pubertal status.

2. The general population studies with good measures of morphological development have not used detailed or diagnostic measures of depression status.

3. The endocrine studies have highly selected samples that are representative of the general population, and have been characterized by non-diagnostic (questionnaire) measures of depression and the use of many statistical tests (mostly non-significant) in relatively small samples.

4. Few studies have included an age range beginning before puberty and extending to a point where most subjects would be expected to have begun puberty.

5. Few studies have taken a longitudinal approach or provide repeated measures of the same subjects across puberty.

6. Retrospective information from adults, the main source of most information until recently, does not provide the level of precision as to timing, or the information on pubertal status needed to address the key questions here. It is also subject to the danger of differential recall by currently depressed and non-depressed adults.

The results that have been obtained in these studies have provided insight into the impact of pubertal development on depression. We present such a study here addressing the following questions: (1) At what age does the female preponderance of depressions begin? (2) Is pubertal status more strongly associated with the emergence of the female preponderance than age? (3) Does pubertal timing have a greater effect than pubertal stage? (4) Does the amount of pubertal change over time have more effect on depression rates than the level of puberty reached by a certain age?

METHOD

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

Sampling frame

A representative sample of 4500 9-11, and 13 year-olds, recruited through the Student Information Management System (SIMS) of the public school system of 11 counties in western North Carolina, was selected using a household equal probability design. As 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 behavior (externalizing) problems, together with some basic demographic and service use questions. All children scoring above a predetermined cut-off score of 20 (design to include about 25% of the population) on the behavioral questions, plus a 1-to-10 random sample of those scoring below the cut-off, were recruited for the longitudinal study. Eighty per cent of eligible families agreed to participate in the interviews at least once, wave (1023 of 1346).

Shortly after being screened, eligible children and one of their parents were interviewed. They were re-interviewed 1, 2 and 3 years later. Interviews were conducted between 1992 and 1996. Between 10% and 24% of the sample participated in each wave; 3733 observations on 1073 children are included in the analysis.

Informed consent was obtained from both children and parents. Because sexual development is a sensitive topic, we showed the Tanner stage personal assessment instrument (see below) to parents before giving it to the
children, and specifically asked permission to use it. At each wave, between 6% and 13% of
parents refused to have the scale administered to their children. Surprisingly, the refusal rate
was twice as high for boys as it was for girls (11.5% vs. 5.9%; \( P = 3 \times 10^{-5} \)). Thus, girls were sig-
nificantly over-represented for analyses involving Tanner staging. However, refusal to complete
the Tanner stage assessment was not significantly related to age, prepubertal status, or depression
diagnosis, so it seems unlikely that this source of missing data was a source of bias in the results.

Table 1 shows the numbers of boys and girls interviewed at each wave of data collection and
the numbers for whom Tanner stage data were available. The total number of observations of
pubertal status over four waves was 1332.

The number of males was always greater than the number of females because the beha
vioral items in the screen used to select the interviewed sample selected more boys than girls, as expected (see e.g. Achenbach et al. 1987). This imbalance
was corrected using a weighting system in all the analyses to generate unbiased population
estimates. Table 2 shows the age distribution of the sample sizes by wave in the Great
Smoky Mountains Study (GSMSS).

<table>
<thead>
<tr>
<th>Wave</th>
<th>Gender</th>
<th>Total N</th>
<th>N with Tanner stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Female</td>
<td>44</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>322</td>
<td>317</td>
</tr>
<tr>
<td>2</td>
<td>Female</td>
<td>420</td>
<td>417</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>541</td>
<td>536</td>
</tr>
<tr>
<td>3</td>
<td>Female</td>
<td>417</td>
<td>414</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>513</td>
<td>511</td>
</tr>
<tr>
<td>4</td>
<td>Female</td>
<td>357</td>
<td>354</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>433</td>
<td>430</td>
</tr>
</tbody>
</table>

Table 3: Age and gender distribution of the observations over four waves

<table>
<thead>
<tr>
<th>Age</th>
<th>Gender</th>
<th>Boys</th>
<th>Girls</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td></td>
<td>1278</td>
<td>1208</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>1272</td>
<td>1207</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>1246</td>
<td>1218</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>1277</td>
<td>1219</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>1234</td>
<td>1213</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>1250</td>
<td>1211</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>1242</td>
<td>1213</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>1231</td>
<td>1217</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>4928</td>
<td>4612</td>
</tr>
</tbody>
</table>

(continued to the nearest birthday) of the observations by gender. At ages 11 to 14, the numbers are larger because individuals from two age
cohorts contributed observations; for instance, 11-year-olds came from wave 1 of the 11-year-
old’s cohort and wave 3 of the 9-year-old cohort.

Measures: Psychiatric symptoms and disorders

Children and parents were interviewed using the Child and Adolescent Psychiatric Assessment
(CAPA) (Angold et al. 1996), which generates Diagnostic and Statistical Manual of Mental
Disorders (DSM-IV) (American Psychiatric Association 1994) diagnoses and a range of
symptom scale scores, measures of functional impairment and family burden. The interview
also includes measures of maternal depression, family psychiatric history, the child’s physical health
and development and mental health service use for details, see (Costello et al. 1996).

Computer algorithms that combined symptom information from parent and child were used
to determine whether a symptom was present. When symptoms were found to be present during
the interview, their dates of onset were also collected, thus allowing for determina-
tion of whether they met the symptom overlap and duration criteria for the various
DSM-IV depression diagnoses. If either parent or child reported a symptom as present in
the past 3 months, it was counted toward the relevant CAPA/DSM-IV scale score or di-
gnosis. This 3-month “primary period” was selected rather than, say, a 1-year or lifetime
period, because shorter recall periods are associated with more accurate recall (see e.g.
Angold et al. 1996a). We considered three depression diagnoses: DSM-IV major depressive
episode, dysthymia and depression—not otherwise specified (NOS). The last of these diagnostic
categories was made up of individuals who met the DSM-IV experiential criteria for minor

Pubertal morphology status

Ratings of pubertal morphology status were based on the standard Tanner staging system
(Tanner, 1962). Originally developed for clinical use, it involved a physical examination, which
is
impractical in non-clinical settings. Self-ratings performed with the aid of schematic drawings of body parts (breasts, pubic hair in girls; genitalia and pubic hair in boys) have yielded good correlations with physical examination based on Tanner stages (Morriss & Udry, 1980). Each child was provided with sex-appropriate schematic drawings and requested to rate her or himself on each dimension. Unless otherwise noted, both self-ratings were averaged to yield a single individual score ranging from 1 (prepubertal), to 6 (adult level of development). For some analyses, children were defined as 'immature' (Tanner stage I or I) or 'mature' (Tanner stages III or higher). Girls were also asked about their menstrual histories, and if post-menarcheal, questions about the date of onset of their menarcheal period.

Analytical strategy
The principal analytical technique was mixed effects hierarchical linear modelling (HLM) (Diggle et al. 1994) with the logistic link function (since the outcome variable was the categorical variable of depression diagnosis). A random effect was introduced to account for correlations between the scores of each individual across waves. In addition, sandwich variance corrections (Diggle et al. 1994; Pickles et al. 1994) were applied to adjust for the effects induced by the sampling stratification based on the use of a screen. The SAS macro GLIMMIX was used to compute the estimates.

RESULTS
1 Age and gender differences in depressive disorder
Fig. 1 shows the 3-month prevalence of depressive disorders by age. The age-specific rate curves were quite different for boys and girls. At or above age 13 girls had consistently higher rates of depression than boys, but this was not the case at earlier ages. A second notable point is that the boys showed a fall in depression rates over the age of 9, while the girls showed an increase after the age of 12. However, in neither case did we observe a smooth change in depression rates with age. Fitting a mixed effects logistic model to these data resulted in a significant effect only of age (OR = 1.07, P = 0.02), reflecting the small overall increase in depression prevalence across this age range.

2 Morphological development and gender differences in depression
Age and morphological development are highly correlated (Pearson r = 0.75 in this sample). However, if morphological development is really an 'active ingredient' in generating age and gender-related changes in rates of depression, the morphological development by gender interaction should be a better predictor of depressive disorders than the age by gender interaction. As shown in Fig. 2, this is exactly what we found. Using Tanner stage, rather than age, 'cleaned up' the developmental pattern quite considerably, with girls having consistently higher rates of depressive disorders than Tanner stage III and boys having higher rates before that. The expected significant interaction between developmental level and gender was seen (gender, OR = 7.2, P = 0.02; morphological development, OR = 0.3, P = 0.002; gender by morphological development, OR = 0.72 (male
coded 1, female coded 0). $P = 0.008$). Immature girls were less likely to be depressed than immature boys ($OR = 0.46, P = 0.02$), but more immature girls were more likely to be depressed than more mature boys ($OR = 2.1$, $P = 0.04$). More mature girls had a substantially higher rate of depression than immature girls ($OR = 3.4$, $P < 0.05$). In boys, the only notable Tanner stage effect lay in the difference between the rate of depression in those at Tanner Stage I and the rest. Dividing the boys in this way resulted in an $OR$ of $0.41$ ($P = 0.03$). These effects were consistent for each of the three depression diagnoses (major depression, dysthymia, and depression-NOS), and so these and other analyses consider these three as a single group (which approach also ensures sufficient power for the subsequent analyses).

3 Effects of the process of change in pubertal status

The idea that the process of change in pubertal status itself might be important in generating differences in depression rates (Brooks-Gunn & Warren, 1985) implies that among the more mature children, girls who had only recently reached this level of maturity should have higher levels of depression than those who had been at that level of maturity for some time. We tested this hypothesis by dividing observations on the more mature participants ($N = 1014$) into those involving individuals who at the previous observation had been classified as being immature ($N = 384$) versus those who had been classified as being more mature ($N = 1230$). In the former group, a change in maturity level had occurred in the past year, while in the latter it had not. We then looked at the rates of depression in boys and girls in these two groups (Fig. 3). There was a significant interaction between change in pubertal status and gender ($P = 0.009$), but the effect driving the interaction was the low rate of depression in boys who changed from the immature to the more mature group. The effect of change was not significant in girls ($OR = 1.25, P = 0.06$), but was highly significant in boys ($OR = 0.24, P = 0.0009$).

We examined rate of change in pubertal status over the previous year, defined as number of Tanner stages traversed between waves. In neither boys nor girls was the rate of change found to have an effect on depression rate in the more mature children.
4 Pubertal status or pubertal timing?

As mentioned in the introduction, several studies have suggested that pubertal timing may be important in generating changes in levels of psychopathology. A true timing effect could have generated the results seen in our data so far, because those with earlier puberty contributed more observations at later stages— in other words, timing of puberty and level of puberty were partially confounded. However, we can uncover these variables among more mature children (Tanner stage III and above), the timing hypothesis predicts that the atoms of depression in girls compared with boys will only be seen in those who reached Tanner stage III relatively early. Thus, we should expect to find a significant interaction between timing of puberty and gender in more mature children. On the other hand, the ‘level’ hypothesis, which argues that girls’ increased risk is associated with maturity per se, regardless of the age at which that maturity was reached, predicts that only a main effect of gender should be observed in more mature children.

To investigate these possibilities we divided the children into those who reached Tanner stage III before the age of 12, and those who did not reach Tanner stage III until later data from 15-year-olds at wave I, who were already more mature, were excluded because we did not know when they reached this level of maturities. The results failed to provide any support for the timing hypothesis. The interaction between early onset and gender was non-significant, and the trend was in the wrong direction (with the early onset group of girls having a lower rate of depression than the normative puberty group), but as we saw above, there was no main effect of gender.

Since one of the major markers of pubertal stage in a number of previous studies of pubertal timing-psychopathology links has been date of menarche, we dichotomized girls into those with menarche before age 12 (the definition of early puberty used by Knutson & Magnusson [1993]), and those with later menarche. Looking at the impact of menarchal timing on the rate of depression in observations conducted at Tanner stage III and above, we found no evidence for an effect of menarchal timing (OR = 0.91, P = 0.08). Again, even the minimal trend was in the opposite direction from that predicted by the timing hypothesis.

DISCUSSION

Previous work established that the female predominance observed in numerous adult studies of depression is present in childhood, but present by age 15-19. However, it has not been determined just where in the adolescent period the change occurs, or which of the many components of the adolescent transition are most strongly related to it. Most previous studies have relied either on adult recall of onset of onset in lifetime diagnostic interviews; or depression scale scores rather than diagnoses, or age or menarche as the only developmental markers. Our purpose was to examine these questions in a relatively large prospective study, with diagnostic assessments and specific measures of pubertal status.

The timing of the emergence of the female predominance in depression

In keeping with the two other gender population studies that provide some kind of age breakdown in early adolescence, the Dunedin Study (Anderson et al. 1987; McCall et al. 1992) and the New York Study (Cohen & Brooks, 1981; Velev et al. 1988; Cohen et al. 1992a, b), we found that the female predominance of depression did not emerge until age 13. Adult retrospective studies have pointed to some time between 10 and 15 (Burke et al. 1990; Kemler et al. 1993). If the difference were already present at age 10 (transitions at age 10-11 were also suggested by Angel & Rutter’s work on a large clinical sample (Angel & Rutter, 1992) and Weissman et al.’s [1988] family study of children at high and low risk of depression), then it would be unlikely that puberty could be invoked as the active factor in the generation of these differences, since at age 10 most girls are still pre-pubertal. By age 13, the great majority of girls are past menarche and into puberty. Timing of the change in depression rates around age 13 makes it unlikely that the increases in estrogen androgens that constitute adolescence can be seen as having any major responsibility since these changes occur in later childhood.
Some aspect of puberty itself is important in the appearance of the female preponderance in depression

An important question is whether the biological changes of puberty themselves are implicated in the increasing rates of depression in girls, or whether these might better be explained by other age-related factors. In relation to this issue we first observed that Tanner stage fitted the depression rate data better than age. The expected association between age and sex in predicting depression was not significant, but that involving Tanner stage was, suggesting that Tanner stage was closer to being an ‘active ingredient’ in the transition to the adult sex ratio than age or some other age-related change.

It also appeared that this transition was a mid-pubertal event (occurring in Tanner stage III), since the rate of depression in girls in the earliest pubertal stage (i.e. Tanner stage I) was no different from that in prepuberty (Tanner stage I). This finding is in accord with retrospective data from the National Comorbidity Survey, which placed the increase in rates of onset of depression in the age range 10-14 (Kessler et al. 1993), whereas the ECA studies (also retrospective) had suggested that the female excess of onset did not begin until ages 15-19 (Burke et al. 1990). Similarly, it agrees with findings from the Dunedin and New York child and adolescent epidemiological studies (Anderson et al. 1992; Cohen & Brodzinsky, 1987; Veliz et al. 1989; McGee et al. 1992; Cohen et al. 1993b, 1993c). However, it extends these findings in demonstrating that: (a) by this stage, statistically significant differences are present rather than just suggestive differences; (b) its effect is quite sharply localized in its developmental location or the transition to mid-puberty, regardless of its timing; and (c) that the sex ratio reverses, with boys being more likely than girls to be depressed before mid-puberty. The most important component of this reversal was a tripling of the prevalence of depression in girls at mid-puberty.

Pubertal status was more important than age at puberty

Although some writers have emphasized the importance of pubertal timing in the generation of psychopathology, it must be remembered that: (a) such timing effects have not, by any means, been consistently observed; (b) they have been based on very limited measures of depression; and (c) any observed differences have not been demonstrated to be stable beyond mid-adolescence. This last point is of particular relevance here, because the female preponderance of depression persists through at least middle age. Based on either age or reaching Tanner stage III or menarcheal age in girls, we found no evidence for a timing effect, and such trends as there were, were in the wrong direction.

Pubertal status and change in pubertal status

Another possible explanation for the apparent effects of pubertal status discussed so far was that they might result from a transient effect of changing physical development (Brooks-Gunn & Warren, 1985). The key difference between these interpretations is that the status hypothesis predicts that physical maturity in girls is associated with increased levels of depression whether or not it was achieved recently. The chronology hypothesis suggests that the psychological effects of hormonal and physical changes are responsible for increased rates of depression, and that once a steady state has been reached, depression should again become less common in girls—in other words we should observe a transient peak in the female prevalence of depression around puberty. First, we looked at whether having made the transition to Tanner stage III or above only during the preceding year, as opposed to earlier, increased the rate of depression in girls, and found that it did not. Secondly, we found that the degree of change in Tanner stage over the preceding year had no significant effect on the rate of depression in those who were at Tanner stage III and above. Thus, in girls, the identified effect on depression rates was reaching Tanner stage III and above—a marked difference when they got there ‘how fast’ they got there, or once there, ‘how fast’ they subsequently matured.

In boys, however, a different, and unexpected, picture emerged. The transition to stage III of puberty was associated with a significant reduction in the rate of depression. Those who had reached this stage only during the preceding year had a much lower rate (OR = 0.34) than those who had been at stage III or above for a year or more. Thus, the transition to mid-
puberty appeared transiently to protect boys from depression. On the other hand, whether they had recently made this transition, or not, boys had lower depression rates after Tanner stage III than girls, and the rate of change in pubertal status had no effect on the prevalence of depression.

What really happens to rates of depression in boys across puberty?

Depression was about twice as common in boys before puberty (Tanner stage I) than in boys who were experiencing the physical changes of puberty (around 6% vs. around 3%). This difference was significant (P = 0.03), but this comparison was only made post hoc because the data indicated that this was where any significant comparison would be. However, in combination with the substantial effect of the transition to stage III, just discussed, the possibility that puberty is associated with at least a transient reduction in the prevalence of depression in boys should not be dismissed. No such effects were observed in the Dunedin or New York epidemiological studies, but those studies included few cases of depression (fewer than 10 at each wave), and did not include measures of pubertal status. Work from the Pittsburgh Youth Study, using a depression questionnaire, (Angold et al. 1996) identified a substantial fall in depression scale scores in boys between first and sixth grades, with a levelling off after sixth grade. However, a later reduction in depression scores with age from first to sixth grade is not consistent with an explanation in terms of puberty, though it would fit with the fall in rates of depression observed after the age of 9 in our data. These data are also consistent with work linking higher levels of adrenal and gonadal androgens (androstenedione and testosterone) with lower levels of negative affectivity (Nestleman et al. 1990).

How are these changes generated?

The first point to make here is that the occurrence of changes in girls in mid-puberty rules out an explanation in terms of adrenarche, though the presence of adrenarchal effects remains open in boys. Secondly, the absence of a unifying effect could explain Slattum & Magness interpretation of effects on psychopathology in terms of early maturing girls become associated with older males with conduct problems. However, since these effects had largely dissipated by age 15 in their study (Slattum & Magness, 1980), and were not reproduced in Colman & Carins' (1994) work, they were never a good candidate for explaining the long-enduring phenomenon of women having a higher rate of depression than men.

While this study does not indicate what mechanisms are actually at work in generating the change in relative rates of depression in males and females (which could be anything from cognitive changes, the social impact of puberty, or changes in levels of life stress to direct effects of hormones on the brain (see Nolen-Hodges & Olin, 1994; Leadbeater et al. 1995) for extensive discussions of many possibilities), it does suggest that further work needs to be focused on identifying enduring changes in risk or protective factors that occur in relation to the transition to mid-puberty, rather than at other times. It also suggests that age is a poor marker for identifying such changes, and that we need more studies that include direct measures of the components of puberty itself.

This project was supported by grant MH 48095 from the National Institute of Mental Health. Additional support was provided through a project support award and Faculty Scholar awards from the William T. Grant Foundation to Dr. Angold, Dr. Costello and Dr. Worthman, and a Center Grant from the Leon Lowenstein Foundation to Dr. Angold.

REFERENCES


