Intrauterine Hormonal Environment and Risk of Developing Anorexia Nervosa

Marco Procopio, MD, MRCPsych; Paul Marriott, PhD

Context: Anorexia nervosa (AN) is approximately 10 times more common in females than in males. The reasons for this difference are not yet understood. Several mechanisms have been hypothesized as possible causes.

Objective: To determine whether the different hormonal environments to which male and female fetuses are exposed in utero might contribute to the increased risk of developing AN in females.

Design, Setting, and Participants: The study is based on a large population-based cohort of Swedish twins. The strategy used is to compare the prevalence for AN between same-sex and opposite-sex twins.

Results: The study shows that the risk of developing AN in female twins is higher than in male twins, as expected. The only exception is male members of opposite-sex pairs, who have a higher risk of developing the illness when compared with other males (P = .62 for narrow diagnostic criteria and P = .60 for broad diagnostic criteria). In fact, their risk is at a level that is not statistically significantly different from that of females from such a pair. A plausible explanation for this phenomenon is that in pregnancies bearing a female fetus, a substance is produced, probably hormonal, that increases the risk of having AN in adulthood. Because the male half of an opposite-sex twin pair would also be exposed to this substance, it could account for the observed elevated risk in males with female twins. The most likely candidates are sex steroid hormones.

Conclusions: The results of our study are compatible with the hypothesis that intrauterine exposure to sex hormones might influence neurodevelopment, affecting the risk of developing AN in adult life. This might be a factor contributing to the higher risk of developing AN in females.

Arch Gen Psychiatry. 2007;64(12):1402-1408
affected to seek help and of their families to recognize the disorder. The hypothesis tested in this study is that the different hormonal environments to which male and female fetuses are exposed influence neurodevelopment in utero, contributing to the diversity in the risk of developing AN between the 2 sexes. The design used to test the hypothesis is a comparison of the prevalence of AN in members of monzygotic (MZ) and dizygotic (DZ) twin pairs, stratified by the possible different sex permutations. The analysis of the prevalence of AN in members of opposite-sex pairs of DZ twins is essential for testing the hypothesis.

Opposite-sex twins constitute a real-life laboratory for studying the influence of sex hormones on intrauterine development. The literature in humans and in animals shows that when a female and a male fetus are contiguous in utero, there is often a masculinization of the female fetus and a feminization of the male fetus. Furthermore, this can have permanent lifetime consequences. There are indications from the literature that the cause of this phenomenon might be the change of the hormonal environment in utero because of the presence of an opposite-sex twin.

A difference in the risk of developing AN in twins who are part of opposite-sex pairs, when compared with members of same-sex pairs, would be consistent with a prenatal influence of gonadal hormones on the likelihood of developing AN in adult life. This could explain, at least in part, the difference in the prevalence of AN between the sexes. An increased prevalence of AN in males who are members of opposite-sex pairs would be suggestive that the presence of a female fetus influences the intrauterine environment, favoring the development of AN. If there were a decreased prevalence of the disease in female members of opposite-sex pairs, this would be indicative that the presence of a male fetus in utero is a protective factor against the development of AN. The 2 previously stated outcomes are not mutually exclusive.

METHODS

SUBJECTS

We have used, for testing the hypothesis, the population identified by Bulik et al in their study on the epidemiological features and heritability of AN. We chose this study because the researchers identified the largest series of twin subjects diagnosed as having AN using rigorous diagnostic criteria and because, unlike most twin research, the data set includes opposite-sex twins, which is essential for our analysis.

The population studied by Bulik et al is represented by all members of the Swedish Twin Registry who were born between January 1, 1935, and December 31, 1958, and who consented to being interviewed on the telephone. The subjects in this population are those who, at the interview, satisfied the DSM-IV definition of AN regardless of the criterion that needed amenorrhea to make a diagnosis. A subpopulation of the previously described group that satisfied the full DSM-IV criteria for AN, including amenorrhea, was defined as fulfilling a narrow definition of AN. Therefore, all subjects who satisfied the narrow definition were also part of the group that satisfied the broader definition of AN, but not necessarily vice versa. This distinction applies only for females because the amenorrhea criterion is not relevant in males, in whom the broad and narrow groups coincide.

The participation rate was 76%. The test-retest reliability was assessed by Bulik et al, who selected at random 105 individual twins who completed the telephone screening. They reinterviewed these subjects within 2 weeks from the first interview and assessed the short-term reliability for 2 key variables associated with eating disorders: low weight and dieting behavior. The Cohen's κ value was 0.63 for AN criterion A (“Did you ever weigh less than other people thought you should weigh?”) and 0.48 for the broader variable querying about dieting history. The researchers judged that the κ variables suggest moderate to substantial test-retest reliability. The reliability of the dieting history variable may be defined as low.

The data used for our analysis were taken from the summary statistics of Table 4 in the article by Bulik et al. We calculated, for each phenotype, the number of individual twins identified as having AN on the telephone interviews. Adding up, for each phenotype, the twins affected by AN and those not affected by the illness, we obtained the total number of twins who participated in the study for each phenotype. From this, we were able to calculate the prevalence of AN for each of the phenotypes. The same operation was done for narrow and broad diagnostic criteria.

STATISTICAL ANALYSIS

Table 1 and Table 2 show that there are marked differences in proportions between males and females in the prevalence of AN, except in the case of a member of an opposite-sex DZ twin pair. In the analysis of this article, the null hypotheses of no differences in proportions are compared using tests for equality of proportions for both dependent and independent samples, as implemented by a statistics package. Unless otherwise stated, all tests are 2-sided, and a continuity correction was used only if it did not exceed the difference of the sample proportions in absolute value. When the test used is appropriate for independent samples, a large sample (χ²) test for equal proportions was used, which is appropriate because the smallest sample was greater than 2000. For the case of dependent samples (eg, different sex twins), the McNemar test is used. The McNemar test is a test of equality of proportions for matched binary responses. It is equivalent to other tests based on linear models, such as the logistic or nonparametric tests, like the Cochran-Mantel-Haenszel test.

---

**Table 1. Prevalence of Anorexia Nervosa Using Narrow Diagnostic Criteria**

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Total No. of Twins</th>
<th>Twins Diagnosed as Having Anorexia Nervosa</th>
<th>Prevalence (95% Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female-female</td>
<td>3634</td>
<td>16</td>
<td>0.44 (0.26-0.73)</td>
</tr>
<tr>
<td>DZ</td>
<td>4818</td>
<td>35</td>
<td>0.73 (0.51-1.02)</td>
</tr>
<tr>
<td>Male-male</td>
<td>2854</td>
<td>2</td>
<td>0.07 (0.01-0.28)</td>
</tr>
<tr>
<td>MZ</td>
<td>4048</td>
<td>1</td>
<td>0.02 (0.001-0.16)</td>
</tr>
<tr>
<td>DZ opposite-sex twin</td>
<td>4478</td>
<td>20</td>
<td>0.45 (0.28-0.70)</td>
</tr>
<tr>
<td>Male</td>
<td>4478</td>
<td>16</td>
<td>0.36 (0.21-0.59)</td>
</tr>
</tbody>
</table>

Abbreviations: DZ, dizygotic; MZ, monozygotic.
Table 2. Prevalence of Anorexia Nervosa Using Broad Diagnostic Criteria

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Total No. of Twins</th>
<th>Twins Diagnosed as Having Anorexia Nervosa</th>
<th>Prevalence (95% Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male-male</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MZ</td>
<td>3634</td>
<td>36</td>
<td>0.99 (0.70-1.38)</td>
</tr>
<tr>
<td>DZ</td>
<td>4818</td>
<td>54</td>
<td>1.12 (0.85-1.47)</td>
</tr>
<tr>
<td>Male-female</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MZ</td>
<td>2854</td>
<td>2</td>
<td>0.07 (0.01-0.28)</td>
</tr>
<tr>
<td>DZ</td>
<td>4048</td>
<td>1</td>
<td>0.02 (0.001-0.16)</td>
</tr>
<tr>
<td>DZ opposite-sex twin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>4478</td>
<td>32</td>
<td>0.71 (0.50-1.02)</td>
</tr>
<tr>
<td>Male</td>
<td>4478</td>
<td>27</td>
<td>0.60 (0.41-0.89)</td>
</tr>
</tbody>
</table>

Abbreviations: See Table 1.

Because the McNemar test is based on the $\chi^2$, it is required that the data are not too sparse. In general, the total number of positive cases (males and females) needs to be greater than 10. In our data, the smallest value was 36; thus, the test was appropriate.

Table 1 and Table 2 represent the prevalence of AN for the different twin phenotypes using narrow and broad diagnostic criteria, respectively. There were no examples, in the sample analyzed, of couples of mixed-sex twins in which both members were affected by AN.

This study confirms, in a twin population, that females have a much higher prevalence of AN than males ($P < .001$ for broad and narrow criteria for comparing all males and females). The only exception, which is the main focus of our study, is observed when a man is the member of an opposite-sex DZ twin pair: in this case, his risk of developing AN increases to much higher levels than in same-sex male twins, levels that are not significantly different ($P = .62$ for narrow diagnostic criteria and $P = .60$ for broad diagnostic criteria) from that of the females in that pair. The risk of developing AN in females who are part of opposite-sex DZ twin pairs is not significantly different from that of other females. These results are obtained using narrow and broad criteria for the diagnosis of AN ($P = .48$ for narrow diagnostic criteria and $P = .18$ for broad diagnostic criteria).

The main question of interest in this investigation is whether different hormonal environments contribute to an increased risk of developing AN. Because this is fundamentally a question about a causal relationship, no definitive conclusions can be drawn from the data, which are observational and archival. The possibility that the observed differences between males in opposite-sex twin pairs compared with those in same-sex pairs could be because of unobserved confounding factors cannot be ruled out and has not been controlled for by any type of randomization.

One possible factor of interest is a cohort effect. This is of particular interest because Bulik et al noted a difference in the rate of AN in males before and after 1945. However, for such an effect to be important for this study, there would have to be a difference effect between same- and mixed-sex twin pairs and not just a time change within a sex group. The trend in time in males and females is clear, but we have no information about a differential effect across type of twin pair.

Other limitations include the fact that the data came from retrospective interviews and the change in detection rate for AN. Also, because of the rareness of reported male AN, the sample sizes are small, although, as previously discussed, they are adequate for the implementation of the statistical tests used.

Members of twin pairs are exposed to the same milieu in utero and usually share a similar environment during at least the first 2 decades of their postnatal lives. Therefore, the increased prevalence of AN in males who are members of opposite-sex DZ pairs could be explained either by the effect, on the prenatal environment, of the contemporaneous presence of a female fetus in utero or the effect of growing up with a female twin. It is unlikely that the increased risk observed in male members of opposite-sex twin pairs is because of obstetric complications connected to twin births, because otherwise the risk would be increased also in MZ and DZ pairs composed of 2 males.

As already mentioned, a hypothesis compatible with the results obtained in our study is that the presence of a twin sister in an opposite-sex DZ couple increases the likelihood that the twin brother is exposed to an environment, during childhood and/or adolescence, that increases his risk of developing AN. The literature on family structure in AN does not seem to support this hypothesis, showing that the risk of developing AN is not altered by the sex of siblings, certainly not to the extent of explaining the dramatic phenomenon observed in our study. These findings are confirmed by the literature on twins and AN, in which the role of the shared family environment does not seem to be significantly different from 0.

The results obtained are consistent with the hypothesis tested in our study: there exists an intrauterine influence because of gonadal hormones, possibly on neurodevelopment, that affects the risk of developing AN. The research on opposite-sex twins, described later, offers a framework that allows an understanding of the potential mechanisms involved in such a hypothesis.

Both animals and humans have characteristics that differ between sexes. These sexually dimorphic aspects include anatomical and psychological/behavioral features. There is extensive literature on the reciprocal influence of fetuses that are members of opposite-sex twin pairs in humans, and of mixed-sex litters in animals, on these sexually dimorphic characteristics. The indication is that female fetuses can be masculinized and male fetuses feminized by the presence of an adjacent opposite-sex twin in utero, in humans and in animals, with permanent consequences. This means that subjects who developed in utero in the proximity of an opposite-sex fetus can show some of the sexually dimorphic features that are usually present in the opposite sex.
The first studies to unravel the previously mentioned phenomenon showed a masculinized pattern of sexually dimorphic features in female mice that developed in utero in proximity to opposite-sex mice, when compared with other female mice. This phenomenon was explained by the influence on the female fetus of the testosterone produced by the adjacent male fetus(es). The described findings are consistent with the observation that the differentiation of the reproductive system in mammals depends mainly on the presence or absence of androgens during critical periods of intrauterine life.

More surprisingly, it has been observed, at times unintentionally, while analyzing data on the masculinization of females, that also the opposite effect occurs: a feminization in the male mice that were surrounded in utero by females. This phenomenon is difficult to explain when looking at the development of feminine characteristics as simply the passive result of absent or low concentrations of androgens, without involving the action of other hormones.

Observations in rats, mongolian gerbils, and pigs confirm that the position in utero in relation to opposite-sex littermates contributes to variability of dimorphic features within the same-sex group. It is not yet clear from the literature if the hormonal influence between opposite-sex littermates takes place through fetofetal diffusion or the local circulation of hormones.

The research conducted in humans is consistent with the previously described animal studies in suggesting that often twins of an opposite-sex pair show sexually dimorphic features discordant with their chromosomal sex in humans and mixed-sex litters in animals, therefore, supports a masculinizing effect on females caused by the presence of a contiguous male fetus in utero, possibly through hormonal percolation. This is not surprising because the evidence on sexual differentiation in mammals shows that a fetus, even if genetically female, develops the sexual characteristics of a male if exposed to androgens during a critical phase of the intrauterine development.

There is also evidence, although not as robust, for the symmetrical effect: the feminization of males who were contiguous to female fetuses in utero. This phenomenon is not consistent with the view of feminization as an “on-off” passive default process that takes place in the absence of androgens. It is instead in line with a more modern view of feminization as the result of the active interaction of several sex hormones, including estrogens, and not just of the presence or absence of androgens. This model is particularly useful for traits that are in a continuum between masculinity and femininity, even within members of the same sex, like in personality traits, and not just either male or female, like in primary sexual characteristics.

A model that takes the presence or absence of androgens as the only factor that affects the masculinization-feminization phenomenon would have predicted that the results in this article would be consistent with the action of androgens as a protective factor against the development of AN. This study is instead compatible with the existence of factors that increase the risk of developing AN in adult life in pregnancies that involve a female fetus. As previously stated, a model with a simple dichotomy between feminine and masculine characteristics, driven only by the presence or absence of androgens, is likely to be simplistic, especially for complex traits like the ones we are studying. Other hormones, including estrogens, are likely to be important in this process.
Evidence for an influence of prenatal sex hormones on the future development of eating disorders also comes from other lines of research. A recent study by Klump et al, opposite-sex twins with the female member having AN (OS Females), and same-sex male pairs (SS Males). The vertical bars represent the 95% confidence interval for the estimate of the prevalence.

and it has to be expected that their concentration in the body fluids at different stages during the intruterine development is relevant.

Based on the results in our study, estrogens are, therefore, tempting candidates for the substance that increases the risk of AN in females through its action during neurodevelopment. This is supported by the evidence in the literature that neuronal growth and proliferation are coregulated by gonadal hormones during neurodevelopment, with permanent and irreversible structural organizing effects on the central nervous system.

Evidence for an influence of prenatal sex hormones on the future development of eating disorders also comes from other lines of research. A recent study by Klump et al showed that the risk of developing an eating disorder is correlated with the individual's digits' length ratio, which is an indicator of the hormonal environment to which the subject was exposed in utero. In another study, Culbert et al examined disordered eating, not just AN, in same-sex and opposite-sex twins, showing evidence that was compatible with an influence of prenatal hormones on the development of eating disorders. The evidence in the latter study was in favor of testosterone being a protective factor, rather than estrogens a predisposing one. The results of this study are not incompatible with ours. One possible explanation for the difference is that the populations in the 2 studies are not homogeneous. The study by Culbert et al did not include only patients with AN but a broader group of subjects with a variety of symptoms of disordered eating. Also, in our study, there is a trend, albeit not reaching statistical significance, for a decreased prevalence of AN in the female members of opposite-sex twin pairs when compared with females in same-sex twin pairs. The trend was stronger when using broad diagnostic criteria, therefore analyzing a population more similar to the one in the study by Culbert et al, rather than using narrow criteria. It is, therefore, possible that both effects are present, a protective action of testosterone and a pathogenic action of estrogens, and each study was able to show just 1 of the 2 phenomena at a statistically significant level because of type II errors that are relatively large because of the rarity of AN. Another possible explanation is that both hormones are involved in the development of eating disorders and the proportion between the concentrations of estrogens and androgens in utero could play a part in shaping the symptoms of the eating disorder in adult life.

It is interesting, with regard to the previous discussion, to consider a trend in the prevalence of AN, with same-sex female twins showing the highest rates, followed by opposite-sex female twins, opposite-sex male twins, and eventually same-sex male twins. This trend cannot be tested formally from the statistical point of view, because the explanatory variate is categorical, but it does offer a visual qualitative perspective of the potential influence on the prevalence of AN by male and female sex hormones (Figure 1 and Figure 2).

The prevalence of AN in the study by Bulik et al was significantly higher in DZ twins than in MZ twins. Before developing hypotheses on the possible relevance of this result, it will need to be replicated. We are not aware of any other published studies that could verify the previously described result. Twin studies are, in fact, usually engineered with the aim to compare the concordance between MZ and DZ twins and not to compare their prevalence. It would be of interest to try to replicate the previously described result in other twin series, also on the evidence that the literature seems to indicate the presence of a similar phenomenon in patients affected by schizophrenia.

In conclusion, the results obtained after comparing the prevalence of AN in opposite-sex and same-sex twins are compatible with the hypothesis that female fetuses produce a hormonal environment in utero that increases the risk of developing AN in adult life, possibly by influencing neurodevelopment. There is also evidence for a trend, albeit not at a statistically significant level, for a protective effect on the development of AN of hormones produced by male fetuses. The difference in the prevalence of AN between males and females might, therefore, originate, at least in part, in utero.
The finding of an influence of prenatal sex hormones on the cause of AN could be integrated in several ways with the evidence from the literature\(^2\)\(^5\)\(^2\)\(^6\)\(^3\) showing that genetic factors have a central role in the cause of the illness. One possibility would be a "two-hit model" in which the genetic endowment would predispose the individual to AN and the prenatal hormonal environment could offer the second hit. It could be hypothesized that individuals who are not exposed to the second hit might show only a partial phenotype or vulnerability to other "second hits" later in life. The 2 previously described explanations, no matter how attractive from the theoretical point of view, are in contrast with the literature\(^2\)\(^5\)\(^2\)\(^6\)\(^3\) that consistently shows that vulnerability to other "second hits" later in life. The 2

A third possibility, which is not mutually exclusive from the previously described possibilities, is that the genetic predisposition for AN coincides with the genetic predisposition to a certain prenatal hormonal setting, unifying the evidence for a genetic cause and our findings.

Submitted for Publication: October 8, 2006; final revision received June 13, 2007; accepted June 15, 2007.

Correspondence: Marco Procopio, MD, MRCPsych, Naper House, 62-64 Bridge St, Walton on Thames, Surrey KT12 1AP, England (marcopropioio00@hotmail.com).

Financial Disclosure: None reported.

REFERENCES

27. vorn Saal FS, Bronson FH. Sexual characteristics of adult female mice are correlated with their blood testosterone levels during prenatal development. Science. 1980;208(4444):597-599.
42. Rohde Parfet KA, Lambersen WR, Rieke AR, Cantile TG, Ganjam VK, vorn Saal


tional and structural MR image analysis and implementation as FSL. *Neuroimage*. 2004;23(suppl 1):S208-S219.


**Correction**

Notice of Retraction. The revision of Table 4 in the Original Article by Cynthia M. Bulik, PhD, et al, titled “Prevalence, Heritability, and Prospective Risk Factors for Anorexia Nervosa,” published in the March 2006 issue of the *Archives* (2006;63[3]:305-312), nullifies the conclusions about the role of intrauterine hormonal environment for male twins in the risk of anorexia nervosa presented in the Original Article by Marco Procopio, MD, MRCPsych, and Paul Marriott, PhD, titled “Intrauterine Hormonal Environment and Risk of Developing Anorexia Nervosa,” published in the December 2007 issue of the *Archives* (2007;64[12]:1402-1408). Hence, we retract the article by Procopio and Marriott from the scientific literature.

For editorial comment, see page 994.

See also page 1061.


Errors in Table. In the Original Article by Cynthia M. Bulik, PhD, et al, titled “Prevalence, Heritability, and Prospective Risk Factors for Anorexia Nervosa,” published in the March 2006 issue of the Archives (2006;63[3]:303-312), errors of presentation occurred in Table 4 on page 310. These were errors of presentation only and have no effect on the findings of the study. The corrected table is reproduced here in its entirety.