Accruing evidence indicates that mate selection is promoted by similarity in body fatness. Assortative mating for obesity may contribute genetically to the obesity epidemic by increasing the risk in subsequent generations. To test this hypothesis, the authors analyzed measured and validated questionnaire data on family members, obtained between 1987 and 2000 from 7,834 obese probands and from 829 subjects randomly ascertained from the general Swedish population. Spouse correlations in body mass index were strongest among couples with the shortest duration of cohabitation. Obesity concordance in parents was associated with an obesity prevalence of 20.1% in adult offspring compared with 1.4% if parents were concordantly nonobese (odds ratio = 18.3, 95% confidence interval: 9.0, 37.4). The prevalence was 8.2% if parents were obesity discordant (odds ratio = 6.5, 95% confidence interval: 3.2, 13.2). No association was found between rearing parents’ and nonbiologic offspring’s body mass index. These results agree with the hypothesis that assortative mating for obesity confers a higher risk of obesity in the offspring generation and thus contributes to the obesity epidemic. Parental obesity concordance is a strong, easily identifiable genetic risk factor that should be considered in the complex network of risk factors for obesity in designing primary prevention programs.

body mass index; genetic screening; genetics, population; marriage; obesity; spouses

Abbreviations: BMI, body mass index; SOS, Swedish Obese Subjects; XENDOS, Xenical in the Prevention of Diabetes in Obese Subjects.

Over the last decades, obesity prevalence has increased significantly in developed countries (1, 2). Although body fatness is influenced by both environmental and genetic determinants (3), the very recent history of the obesity epidemic has naturally singled out lifestyles and environmental factors, associated with positive energy balance, as the immediate culprits. On the other hand, the effect of any lifestyle factor appears to be modulated by genes, often referred to as gene-by-environment interaction. Support for this notion as applied to obesity was given by controlled experiments on monozygotic twins (4, 5), showing that persons with identical genotypes respond similarly regarding body mass and body composition to changes in energy balance, whereas the variability is larger among those whose genotypes differ. Thus, the current obesity epidemic may be viewed as an environmentally and lifestyle-driven increase in the penetrance of obesity susceptibility genes.

The prevalence and heterozygosity of susceptibility genes are expected to remain constant across generations if (and only if) mate choice occurs randomly with respect to associated phenotypes. Data from several studies, however, suggest that a portion of all unions is characterized by positive assortment for various measures of body mass and fatness (6–14).

Recently, Hebebrand et al. (11) hypothesized that the current obesity epidemic has a genetic component mediated...
by increased rates of assortative mating for body fatness. Theoretically, the genetic consequences of assortative mating for complex traits, such as obesity, are expected to become more significant as the correlation between genotype and phenotype (penetrance) increases, even if rates of assortative mating remain constant across generations.

In the present study, we investigated a large Swedish cohort for evidence of assortative mating for body mass index (BMI; calculated as body weight in kilograms divided by body height in meters squared). Furthermore, we tested the data for consistency with the hypothesis that spouse resemblance in BMI affects the prevalence of adult obesity in the subsequent generation.

**MATERIALS AND METHODS**

**Study participants**

The study probands were either obese patients, screened for inclusion eligibility for either of two Swedish nationwide obesity intervention studies—the Swedish Obese Subjects (SOS registry, n = 6,328) and the Xenical in the Prevention of Diabetes in Obese Subjects (XENDOS, n = 4,423)—or participants selected randomly from the general population (The SOS reference study, n = 1,135). Participants in the SOS registry and XENDOS studies were ascertained through advertisements in the media. Details of the study cohorts have been described previously, as referenced below. Briefly, the SOS study (15) included primarily severely obese subjects living in 18 of the 24 counties of Sweden. Recruitment occurred over the 1987–2000 period. Among inclusion criteria were age 37–60 years and BMI above 34 kg/m² for males and 38 kg/m² for females. Measurements took place at 480 primary health care centers. Mean age was 47.4 (standard deviation, 6.0) years, and mean BMI was 39.7 (standard deviation, 4.5) kg/m².

The SOS reference study (16) was completed during 1997 at 22 outpatient clinics. Mean age was 49.4 (standard deviation, 6.8) years, and mean BMI was 36.8 (standard deviation, 3.9) kg/m².

Inclusion criteria in the XENDOS study (17) included age 30–60 years and a BMI of 30 kg/m² or more, and recruitment took place during 1997 at 22 outpatient clinics. Mean age was 44.7 (standard deviation, 8.0) years, and mean BMI was 36.8 (standard deviation, 3.9) kg/m².

Body mass and stature measured prior to any intervention were used in the present study. The demographics of the study subjects are summarized in table 1.

Data for spouses and adult offspring regarding current body mass, stature, age, sex, duration of cohabitation, as well as biologic relatedness to offspring were obtained from questionnaires completed by the probands. After exclusion of probands who had not reported data on spouses, data for a total of 8,663 spouse pairs were available for analysis. The research projects were approved by all ethics committees concerned, and all participants gave written informed consent.

**Statistical methods**

BMI was adjusted by using a multiple regression procedure in which age, age squared, and age cubed were forwardly entered and, if significant at the 5 percent level, were retained as covariates in the model. The model was applied separately in six groups: male spouses, female spouses, biologic daughters, biologic sons, foster daughters, and foster sons. The means and variances of the residuals, which constituted the adjusted phenotype, were standardized to the original scale for the purpose of graphic display.

Spouse resemblance was assessed by using Spearman rank correlation. This nonparametric method was chosen to avoid confounding by nonnormal trait distributions. Ninety-five percent confidence intervals for the correlation coefficients were computed by using Fisher’s z transformation. Bivariate analyses of the SOS reference population were based on sex-specific quintile ranks obtained from the same population. A chi-squared test was used to investigate

---

**TABLE 1. Age and BMI by generation and sex among families ascertained either through an obese proband (Swedish Obese Subjects intervention study, 1987–2000, and the XENDOS* study, 1997) or randomly from the general population (The SOS reference study, 1994–1999)**

<table>
<thead>
<tr>
<th></th>
<th>Proband and spouse</th>
<th>Adult offspring</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD*)</td>
<td>Range</td>
</tr>
<tr>
<td><strong>Ascertainment for obesity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age in years Males (n = 7,834)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>47.8 (8.0)</td>
<td>22–87</td>
</tr>
<tr>
<td>BMI</td>
<td>31.2 (6.4)</td>
<td>17–75</td>
</tr>
<tr>
<td><strong>Random ascertainment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age in years Males (n = 829)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>50.2 (7.8)</td>
<td>24–81</td>
</tr>
<tr>
<td>BMI</td>
<td>25.5 (3.2)</td>
<td>18–42</td>
</tr>
</tbody>
</table>

* BMI, body mass index (weight in kilograms/height in meters squared); XENDOS, Xenical in the Prevention of Diabetes in Obese Subjects; SD, standard deviation.
the magnitude and statistical significance of deviations from the null hypothesis of random mating. Because the SOS reference, the SOS registry, and the XENDOS studies used different BMI eligibility criteria, these studies were pooled to cover a wide range of BMI values. All analyses were performed by using SAS version 8.2 software (SAS/STAT Users' Manual; SAS Institute, Inc., Cary, North Carolina).

RESULTS

Validity of reported anthropometric data

The validity of BMI, calculated from body mass and stature in spouses and adult offspring as reported by probands, was assessed in an ongoing family ancillary study within the SOS project. Subjects completed questionnaires identical to those used in the current study. The BMI of 135 spouses and 146 biologic adult offspring measured 1–30 days following questionnaire completion was compared with reported data. Results are shown in figure 1. Average difference between reported and measured BMI in spouses and offspring was –0.2 and –0.8 kg/m², respectively. The negative sign indicates a tendency toward underestimation of BMI, as has been seen in previous validation studies of data from self-reports (18–23) and reports for relatives (24).

Assortative mating

Spouse correlation in current age-adjusted BMI, based on all 8,663 pairs, was 0.18 (95 percent confidence interval: 0.16, 0.20). Spouse similarity does not necessarily indicate assortative mating because it might also be the effect of cohabitation. In contrast to assortative mating, a cohabitation effect should make spouses more similar over time. To investigate the nature of the spouse resemblance, spouse correlations were analyzed within five duration-of-cohabitation classes. As shown in figure 2, an assortative mating effect was strongly supported; the correlation was highest among pairs who had lived together 5 years or less, whereas correlation point estimates were lower for pairs with a longer duration of cohabitation. This decline in correlation was not related to interclass differences in age, BMI or year during which the proband was included in the study.

To investigate whether spouse resemblance was uniform across the entire BMI range, the 829 available pairs from the SOS reference study were analyzed categorically. (This analysis required that only the randomly ascertained pairs be used, since pairs from the obesity studies always included one obese proband, by definition.) Sex-specific BMI quintiles were computed, thus yielding equally large sex-by-quintile subgroups, with 4 percent frequency per pair-type expected under random mating. As illustrated in figure 3, deviations from this frequency were positive among quintile-concordant pairs and negative among discordant pairs (chi-square = 43.5, df = 16, p = 0.0002). Such deviations were evident across the whole range, but most prominently within the top and bottom quintiles. Thus, assortative mating appears to occur most frequently within the extremes of the BMI distribution in the general population.

Adult obesity in the second generation

The relation between current spouse resemblance in obesity status and age-adjusted BMI in adult offspring, conditional on parent-offspring biologic relatedness, was
investigated in families of the 4,118 probands who had reported weight and height of spouses and adult offspring. Offspring were classified as 1) reared by two biologic parents, 2) reared by one biologic and one foster parent, or 3) reared by two foster parents. Parental pairs were classified as 1) obesity concordant, 2) concordantly nonobese, or 3) obesity discordant. The interaction between the biologic relatedness vector and parental obesity status vector yielded nine offspring classes. Moreover, an additional offspring class was defined because cases for whom rearing parents were discordant for both obesity and biologic relatedness to the offspring were dichotomized depending on whether it was the biologic or nonbiologic parent who was obese.

Figure 4 shows means of age-adjusted BMI among the 10 classes of offspring. The effect of parental obesity on offspring BMI appeared to totally depend on whether there was biologic relatedness to the parents. In offspring reared by their biologic parents, obesity prevalence was 20.1 percent (227/1,131) when both parents were obese (odds ratio = 18.3, 95 percent confidence interval: 9.0, 37.4) and 8.2 percent (413/50,41) when pairs were of the obese/nonobese type (odds ratio = 6.5, 95 percent confidence interval: 3.2, 13.2). The prevalence was 1.4 percent (8/592) when both biologic parents were nonobese. Although representing 16 percent of parental pairs, obese/obese pairs contributed 35 percent of all obese adult offspring.

Simulation studies

Although the notion of a genetic component in the obesity epidemic was supported by the observation that assortative mating was most prominent in the upper BMI distribution and by the substantially higher obesity risk for biologic offspring from such matings, it is not evident that the observed deviations from random mating could affect the genotype at associated loci. We investigated this issue by simulating a normally distributed trait under additive influence of a biallelic quantitative trait locus in Hardy-Weinberg equilibrium, with the risk allele prevalence set to 30 percent. Trait decile-specific genotype probabilities were paired at frequencies identical to the decile mating-type frequencies observed in the population-based sample, and a new generation of trait values was iteratively simulated based on the updated set of genotype probabilities. A detailed account of the simulation procedure is given in the Appendix. The effect over multiple generations on the prevalence of trait values exceeding 1, 2, and 3 standard deviations above the population mean is shown in figure 5. The slopes of the regression lines suggest that the highest per-generation incidence occurs at the uppermost extreme of the distribution.

DISCUSSION

In the present study, we tested a large body of measured or validated questionnaire data for consistency with the hypothesis that assortative mating for obesity occurs and contributes to the obesity epidemic. Ideally, assessment of mate selection should be based on data obtained when selection took place. Because such information was unavailable, we used current spouse BMI resemblance by duration of cohabitation as a proxy. The
finding that BMI correlation was strongest among the couples with the shortest duration of cohabitation provides evidence that spouse resemblance was the consequence of assortative mating rather than the sharing of a common environment.

Although assortative mating for BMI occurred across the whole range, it was more pronounced at the extremes. This finding is in agreement with published data from the Canada Fitness Survey (13). Whereas an excess of obese-obese matings is consistent with the observed increase in obesity prevalence, the expected increase in leanness due to an excess of lean-lean matings has not been reported before to our

FIGURE 3. Associations between quintiles of body mass index (BMI) among 829 randomly ascertained spouse pairs. Positive (negative) deviations from 4% (the expected frequency given random mating) indicate positive (negative) nonrandom assortment. The Swedish Obese Subjects reference study, 1994–1999.

FIGURE 4. Body mass index (BMI) of 7,824 adults given obesity status in rearing parents and parent-offspring biologic relatedness. Offspring whose parents are discordant for both obesity and biologic relatedness are represented by paired bars: obese biologic parent (right), obese foster parent (left). Pooled data from the Swedish Obese Subjects intervention and reference studies, 1987–2000, and the XENDOS study, 1997. n, total number of offspring in each category along the horizontal axes. BMI was adjusted for age. Error bars represent 95% confidence intervals.
knowledge. This finding might be due to heterogeneity in the etiology of leanness, that is, the lean population constitutes a mix of yet-unexposed and exposed, but genetically obesity-resistant people, thus potentially diminishing the genetic effects of assortative mating for leanness to the point that it is not always detectable, particularly in smaller studies. While it is obvious that a lean phenotype does not absolutely require a particular genotype, it is also evident (from this study, among others) that an obesogenic environment does not inevitably lead to obesity. Consequently, one could hypothesize that assortative mating for obesity may have a greater genetic component than would assortative mating for leanness.

Direct evidence to the effect that assortative mating for obesity contributes genetically to the obesity epidemic will ultimately be obtained by molecular genetic studies over multiple generations, in which secular changes in allelic distributions at associated loci can be observed. The simulations, albeit simplified, indicated that modest deviations from random mating, as observed in this study, are sufficient to bring about appreciable changes at an associated locus. Indirect evidence to this effect was also provided; obese-obese pairs were more common than expected, and offspring reared by them had the highest mean BMI and obesity prevalence, but only if there was biologic kinship to the offspring. Since this analysis was based on current (rather than premarital) spouse resemblance, and since only adult offspring BMI was available, the minimal spouse duration of cohabitation was 17 years. Consequently, parental trait similarity did not necessarily equal premarital similarity. Given the tendency of BMI to increase with age, the most likely bias was introduced by the unrealistic assumption that all currently obese parents were already obese at the time of mate selection. However, it is unlikely that the results were driven by this bias, since that would imply that children whose parents became obese later in life have a higher obesity risk than children of parents with earlier-onset obesity. Conversely, inclusion of parents with late-onset obesity more likely resulted in underestimation of the obesity risk for offspring, because late-onset obesity appears to be less heritable than early-onset obesity (25).

Our results agree with previously reported evidence of assortative mating as a risk factor for obesity in offspring. In a study on German children hospitalized because of extreme obesity, it was observed that their parents’ BMI, whether contemporaneously measured or recalled at age 20 and 30 years, tended to cluster in the upper distribution (11). A population-based study in Canada showed stronger correlation in BMI among parents of children in the upper or lower extremes of the distribution (13). Recently, an English study found that parental obesity was the strongest of eight risk factors for obesity at age 7 years (26). None of those studies provided any data to indicate that assortative mating influences obesity prevalence, for example, that parental obesity concordance was more common than expected. Furthermore, we investigated the effects of parental obesity on an offspring’s risk in adulthood. Finally, our results highlight the fact that biologic inheritance is a requirement for the transfer of obesity risk across generations; that is, exposure to an obesity-promoting environment alone is not always sufficient to cause obesity.

In conclusion, these data support the hypothesis that positive assortative mating for obesity includes a genetic component that may partly explain the current obesity epidemic. The observation that spouse concordance for obesity was associated with a 20-fold higher obesity risk for biologic offspring compared with children of concordantly nonobese parents may provide a simple, yet powerful screening tool. Therefore, we propose that biologic mothers’ and fathers’ BMI be considered in early identification of persons who would benefit the most from targeted primary and secondary obesity prevention programs.

ACKNOWLEDGMENTS

Supported by a grant (05239) from the Swedish Medical Research Council and by grants from Hoffman-La Roche, Basel, Switzerland (to Dr. Sjöström), and by an unrestricted grant from Bristol-Myers Squibb and the George A. Bray Chair in Nutrition (to Dr. Bouchard).

Conflict of interest: none declared.
REFERENCES


APPENDIX

Simulation of Genetic Nonrandom Mating Effects on a Quantitative Trait Locus

Consider an additive quantitative trait locus with alleles A and a with probabilities p and q, respectively. Let the genotypes AA, Aa, and aa be in Hardy-Weinberg equilibrium \((p^2 + 2pq + q^2 = 1)\), with corresponding trait means \(-1, 0, \text{ and } 1\). Rank the pooled simulated trait values into \(b\) quantiles, and estimate for each quantile the genotype probabilities of AA, Aa, and aa. Let the vector

\[
\mathbf{m}_i = \begin{pmatrix}
\hat{m}_{aa} \\
\hat{m}_{Aa} \\
\hat{m}_{AA}
\end{pmatrix}
\]

denote genotype probabilities for males in the \(i\)th quantile and

\[
\mathbf{f}_j = \begin{pmatrix}
\hat{f}_{AA} \\
\hat{f}_{Aa} \\
\hat{f}_{aa}
\end{pmatrix}
\]

denote females in the \(j\)th quantile. The probability matrix of genotype unions from the \(ij\)th mating type, \(G_{ij}\), is defined by the outer product

\[
G_{ij} = \mathbf{m}_i \mathbf{f}_j
\]

For offspring of the \(ij\)th mating type, the probability of AA is given by

\[
P_{AA_j} = (1 1 1) \begin{pmatrix}
1 \\
1 \\
1
\end{pmatrix} \begin{pmatrix}
\hat{G}_{ij} & \mathbf{A}
\end{pmatrix}
\]

where \(\mathbf{A} = \begin{pmatrix}
0 & 0 & 0 \\
1 & 1 & 0 \\
1 & 1 & 0
\end{pmatrix}\).
whereas the offspring probability of $aa$ is given by

$$p_{aa_o} = (1 \ 1 \ 1) \left( G_{ij} \cdot A^T \right) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

(Note that $G_{ij} \cdot A$ represents element-wise multiplication.) To obtain the grand genotype probabilities ($k_{AA}$, $k_{aa}$, $k_{aA}$) in offspring, let $p_{AA_i}$ form the elements of a quadratic $i$ by $j$ matrix

$$P_{AA} = \begin{pmatrix} p_{AA_{11}} & p_{AA_{12}} & \cdots & p_{AA_{1k}} \\ p_{AA_{21}} & p_{AA_{22}} & \cdots & p_{AA_{2k}} \\ \vdots & \vdots & \ddots & \vdots \\ p_{AA_{k1}} & p_{AA_{k2}} & \cdots & p_{AA_{kk}} \end{pmatrix}$$

and use for each element a weight proportional to the mating probability $q_{ij}$ (as estimated from the observed mating frequencies in the population under study)

$$k_{AA} = (1 \cdots 1_b)(P_{AA} \cdot N)(1 \cdots 1_b)^T,$$

where

$$N = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1b} \\ q_{21} & q_{22} & \cdots & q_{2b} \\ \vdots & \vdots & \ddots & \vdots \\ q_{b1} & q_{b2} & \cdots & q_{bb} \end{pmatrix}.$$

(Similarly to above, $P_{AA} \cdot N$ is element-wise multiplication.) The corresponding probabilities of $aa$ and $aA$ are given by

$$k_{aa} = (1 \cdots 1_b)(P_{aa} \cdot N)(1 \cdots 1_b)^T$$

and

$$k_{aA} = 1 - (k_{AA} + k_{aa}).$$

Based on the new genotype probability estimates, a new data set is simulated, and the procedure is iterated by using the $N$ matrix.