Sex Differences in the Risk of Schizophrenia

Evidence From Meta-analysis

André Aleman, PhD; René S. Kahn, MD, PhD; Jean-Paul Selten, MD, PhD

Background: Sex differences in the risk of a particular disorder can yield important clues regarding its pathogenesis. The evidence for a sex difference in the risk of schizophrenia is inconclusive. The purpose of this study was to integrate results from the published literature and to provide a quantitative index of the male-female ratio for the incidence of schizophrenia.

Methods: The MEDLINE and PsychLIT databases were searched for English-language publications on “incidence and schizophrenia” that appeared during the period between January 1980 and September 2001. Population-based incidence studies using standard clinical diagnostic criteria were included if they reported sex-specific incidence rates. Sex-specific incidence figures were extracted directly from each study. Categorical analyses were conducted on a subset of studies that met specific methodological criteria (to minimize criterion bias, hospital bias, and age bias). Study categorization and data extraction were performed independently by 2 of us (A. A. and J.-P.S.).

Results: Log risk ratio meta-analysis was conducted using a random-effects model. The incidence risk ratios for men to develop schizophrenia relative to women were 1.42 (95% confidence interval [CI], 1.30-1.56) when all studies were included in the analysis (49 effect sizes), 1.31 (95% CI, 1.13-1.51) when studies that minimized selection biases were analyzed separately (23 effect sizes), and 1.39 (95% CI, 1.15-1.68) when only high-quality studies were included (11 effect sizes). The sex difference was significantly smaller in studies with sample years before 1980 than those with sample years after 1980. No significant sex differences were reported in studies from developing countries. A final analysis, limited to studies with an age cutoff of 64 years or older (16 effect sizes), yielded a mean risk ratio of 1.32 (95% CI, 1.13-1.55).

Conclusion: This meta-analysis provides evidence for a sex difference in the risk of developing schizophrenia, as reported in the published literature from the last 2 decades.

Arch Gen Psychiatry. 2003;60:565-571

The male sex appears in general to suffer somewhat more frequently from dementia praecox than the female.

Emil Kraepelin, 1919

Despite Kraepelin’s early assertion, it is still unclear whether schizophrenia affects one sex more often than the other. This is unfortunate because sex differences in the risk of a particular disorder can yield important clues about that disorder’s pathogenesis. Sex differences allow a unidirectional interpretation; sex cannot be influenced by the disorder. In the case of schizophrenia, important differences between men and women have been shown regarding variables such as age of onset, premorbid functioning, symptomatologic characteristics, and course of illness. It is also widely acknowledged that in younger age groups, the risk is higher for men and that beyond age 40 years and perhaps even earlier, the risk is higher for women. However, the general view is that there are no sex differences in the lifetime risk of developing the disorder or that the evidence is inconclusive.

It has been argued repeatedly that any findings of a male excess are confounded by several important factors including criterion bias, age bias, and hospital bias. The role of criterion bias is illustrated by the observation that the application of narrow criteria for schizophrenia excludes more women than men. The criteria according to the DSM-III, DSM-III-R, and DSM-IV, for example, which imply that signs of the disturbance should be present continuously for at least 6 months, are narrower than those according to the International Classification of Diseases, Ninth Revision (ICD-9) or ICD-10. A second
point of concern is the use of an age cutoff because women in older age groups are at higher risk for developing the disorder than men.8 Thus, to obtain a reliable index of the male-female incidence ratio, studies should ideally include patients of all ages. Third, studies that are limited to inpatients may also promote sampling bias. Because the male sex is associated with a less favorable course of the disorder (eg, violence and aggression are more common in men), men may be overrepresented in samples based on first hospital admissions.10 The inclusion of patients recruited by outpatient services is an important factor in the reduction of methodological artifact.

In this article, we use meta-analytic methods to integrate results from the published literature and to provide a quantitative index of the male-female ratio for schizophrenia incidence. To minimize the selection biases described previously, we selected for separate meta-analysis several studies that met specific criteria. Moreover, we conducted an extra analysis on high-quality studies that administered semistructured diagnostic interviews to large numbers of patients.

STUDY SELECTION

MEDLINE and PsychLIT were systematically searched for articles on “incidence and schizophrenia” published between January 1980 and September 2001. Reference lists from identified articles were reviewed. A database of all schizophrenia incidence studies published during the targeted period was also searched.

To be included in the meta-analysis, studies had to fulfill the following criteria: (1) use of standardized clinical diagnostic criteria (eg, DSM-IV or ICD-9); (2) reporting of sex-specific incidence figures; (3) provision or use of a denominator in computing incidence rates; and (4) publication in an English-language, peer-reviewed, indexed scientific journal. Meeting abstracts were not included.

META-ANALYSIS PROCEDURE

For each study a male-female risk ratio (RR) was calculated. Risk ratios were preferred to odds ratios because they are more intuitive to physicians and more robust to heterogeneity. Given the optimal statistical properties associated with the log RR, we used this method of meta-analysis.16,17 Data extraction and calculation of RRs were performed independently by 2 of us (A. A. and J.-P.S.), who reached a consensus in cases of discrepancy.

From the individual study effect sizes, a variance-weighted mean log RR was computed that could be transformed back into a normal RR. Many studies lacked sufficient information for 2×2 contingency tables (numerators and denominators divided by sex) because they did not provide sex-specific numbers of the general population in the investigated area. Moreover, when such numbers were given, there were large differences between studies, with some referring to the total population, others to the age-adjusted population, and still others to person-years at risk. To include the same metric for all studies, we therefore estimated the study variance using the following formula:

\[ V = \frac{1}{N_m} + \frac{1}{N_f}, \]

where \( N_m \) is the number of male cases and \( N_f \) is the number of female cases (derived from the formula for the variance of odds ratios).18 To prevent studies with very large samples from dominating the analyses, the size of those with more than 500 patients was set at 500, as suggested by Shadish and Haddock.17

A homogeneity statistic, \( Q_W \), was calculated to test whether the studies within an analysis could be considered to share a common population effect size.17 A significant \( Q_W \) statistic indicates heterogeneity of the individual study effect sizes, which poses a threat to a reliable interpretation of the results. In several analyses, 2 groups of studies were compared using categorical analysis (studies with an age cutoff < 64 years vs those with a cutoff ≥64 years). The \( Q_W \) statistic (a between-category homogeneity statistic)17 was used to test whether such groups differed significantly in their mean effect sizes. This statistic has a \( \chi^2 \) distribution and is analogous to an F test. All analyses were carried out using the random-effects model with MetaWin 2.0 statistical software (Sunderland, Mass).

When the same study reported results based on more than 1 classification system (eg, DSM-III-R and ICD-9), those based on the system that used the broadest schizophrenia definition were selected (in this example, the ICD-9). Studies that used CATEGO criteria19 (based on the CATEGO computer program) were included. However, when these studies also reported rates based on the ICD-9, the latter were preferred because CATEGO is not technically a system of diagnostic criteria. When studies published results based on DSM-III, DSM-III-R, or DSM-IV criteria, the categories “schizophreniform disorder” and “schizoaffective disorder” were included when possible.

Seven different analyses were carried out. First, an initial analysis was performed in which all studies were included (analysis 1). Second, a separate analysis (analysis 2) was conducted on a subset of studies that met the following requirements: (1) age cutoff of 54 years or older; (2) case findings included both inpatient and outpatient services; and (3) ICD classification of schizophrenia, all subtypes (section 295.x), or CATEGO “broad schizophrenia,” which includes the CATEGO classes S (schizophrenic psychosis), P (paranoid psychosis), and O(+) (other psychosis). The choice of requirement 3 was motivated by the observation that most studies used these criteria. The selection of these studies, which are relatively homogeneous in terms of diagnostic criteria, could contribute to the reduction of methodological variance. This subset of relatively unbiased studies was contrasted with those that did not fulfill these criteria (“biased” studies). Analysis 3 concerned high-quality studies that met all of the criteria from analysis 2 but also fulfilled the following requirements: (1) use of semistructured diagnostic interviews and (2) inclusion of at least 50 cases of schizophrenia.

To examine any effect of the inclusion of older subjects, an analysis was performed on all studies with an age cutoff of 64 years or older (analysis 4). In the next analysis, we evaluated the studies that used DSM-III-R and DSM-IV criteria separately (analysis 5). Studies based on DSM-III criteria were not included because these criteria demand that symptoms manifest before age 45 years. One study22 included patients with an ICD-8 or DSM-III-R diagnosis but reported sex-specific incidence rates only for the (mixed) group as a whole. Therefore, this study was omitted from analysis 5. Furthermore, we explored possible time trends for the male-female RR by comparing studies with sample years before 1980 with studies that included only sample years after 1980 (analysis 6). When the range of sample years within 1 study included the years before and after 1980, we included the study in the first group (very few studies had sample years that were exclusively before 1980). When a study included multiple samples22-23 with sample years before 1980 and independent groups with sample years after 1980, data from these samples were included separately in the corresponding analysis. Finally, findings from developed and developing countries were contrasted (analysis 7). The World Health Organization human development index of countries24 was used to classify the studies according to this variable. Spe-
results of meta-analyses of sex differences in schizophrenia incidence and categorical moderator analyses

<table>
<thead>
<tr>
<th>Variable</th>
<th>No. of Effect Sizes</th>
<th>RR</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
<th>Qw</th>
<th>Qb</th>
</tr>
</thead>
<tbody>
<tr>
<td>All studies</td>
<td>49</td>
<td>1.42</td>
<td>1.30</td>
<td>1.56</td>
<td>50.7</td>
<td>1.0</td>
</tr>
<tr>
<td>Sex-related sample bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unbiased*</td>
<td>24</td>
<td>1.35</td>
<td>1.17</td>
<td>1.56</td>
<td>26.8</td>
<td>1.7</td>
</tr>
<tr>
<td>Biased</td>
<td>25</td>
<td>1.48</td>
<td>1.30</td>
<td>1.69</td>
<td>24.5</td>
<td></td>
</tr>
<tr>
<td>High quality</td>
<td>11</td>
<td>1.39</td>
<td>1.15</td>
<td>1.68</td>
<td>9.0</td>
<td></td>
</tr>
<tr>
<td>Age cutoff, y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;64</td>
<td>32</td>
<td>1.50</td>
<td>1.33</td>
<td>1.68</td>
<td>29.4</td>
<td></td>
</tr>
<tr>
<td>≥64</td>
<td>16</td>
<td>1.32</td>
<td>1.13</td>
<td>1.55</td>
<td>19.7</td>
<td></td>
</tr>
<tr>
<td>Diagnostic system</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICD-8, ICD-9, ICD-10</td>
<td>29</td>
<td>1.45</td>
<td>1.29</td>
<td>1.63</td>
<td>32.1</td>
<td></td>
</tr>
<tr>
<td>DSM-III-R, DSM-IV</td>
<td>8</td>
<td>1.58</td>
<td>1.21</td>
<td>1.98</td>
<td>6.1</td>
<td></td>
</tr>
<tr>
<td>Sample years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predominantly before 1980</td>
<td>25</td>
<td>1.27</td>
<td>1.15</td>
<td>1.41</td>
<td>34.9</td>
<td></td>
</tr>
<tr>
<td>Only after 1980</td>
<td>24</td>
<td>1.54</td>
<td>1.39</td>
<td>1.71</td>
<td>36.4</td>
<td></td>
</tr>
<tr>
<td>Development index</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>42</td>
<td>1.48</td>
<td>1.34</td>
<td>1.63</td>
<td>40.7</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>7</td>
<td>1.09</td>
<td>0.79</td>
<td>1.49</td>
<td>5.1</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; ICD-8, International Classification of Diseases, Eighth Revision; ICD-9, ICD, Ninth Revision; ICD-10, ICD, 10th Revision, Qh, between-category homogeneity statistic (df = 1; not significant unless otherwise indicated); Qw, within-category homogeneity statistic (df = number of effect sizes minus 1); RR, weighted mean risk ratio.

*Studies were denoted “unbiased” when they fulfilled the following criteria: age cutoff ≥54 years and case finding includes inpatient and outpatient services. The use of ICD or CATEGO criteria was an additional criterion to enhance diagnostic homogeneity.

†P < .01.
‡P < .05.

Thirty-eight articles8,20-23,25-57 met our criteria for inclusion in the initial analysis, yielding 49 effect sizes (several studies reported on multiple independent groups). These included register studies, prospective first-contact studies (for a defined period and location), and cohort studies.

Results of the meta-analyses are detailed in the Table. Analysis 1, which concerned all studies (49 effect sizes), yielded a mean RR of 1.42 (95% confidence interval [CI], 1.30-1.56). Analysis 2 compared studies that minimized sex-related sampling bias with those that did not. The results are listed in the Table. The magnitude of the mean RR is somewhat smaller for the studies that minimized sampling bias than for studies that did not control for sampling bias (1.35 vs 1.48), but this difference was not significant (Q = 1.0; P = .32). The studies that minimized sampling bias still show a considerable sex difference in schizophrenia incidence. Analysis 3, with the methodologically more rigorous studies (11 effect sizes), yielded a mean RR of 1.39 (95% CI, 1.15-1.68).

Analysis 4, which involved studies with an age cutoff of 64 years or older, yielded a mean RR of 1.32 (16 effect sizes; 95% CI, 1.13-1.55). The difference with the other studies, which had an age cutoff of younger than 64 years (32 effect sizes; mean RR, 1.50), was not significant (Q = 1.7; P = .19). When studies using DSM-III-R or DSM-IV criteria were analyzed separately (9 studies; analysis 5), a mean RR of 1.58 was obtained. The difference from the mean effect size of the studies using ICD criteria (30 effect sizes; mean RR, 1.45) was not statistically significant.

The Table also shows the results of the analysis (analysis 6) in which studies including sample years before 1980 were compared with studies that had only sample years before 1980. The difference in the mean RR (1.27 and 1.54, respectively) was significant (Q = 7.3; P = .007). To examine the possibility that this result was due to the application of more restrictive criteria after 1980, we carried out an additional analysis that was confined to studies using ICD criteria. The difference (mean RR, 1.32; 95% CI, 1.18-1.48; 16 effect sizes; vs mean RR, 1.55; 95% CI, 1.37-1.74; 16 effect sizes) became somewhat smaller but remained statistically significant (Q = 4.1; P = .04). Thus, the sex difference was found to be smaller in studies that included samples diagnosed before 1980.

Finally, studies from countries with a high World Health Organization human development index were compared with those that had a medium development
The difference in mean RR was significant \((Q_B=4.9; P=.03)\); there appeared to be no significant sex difference in schizophrenia incidence in the countries with a medium development index. The CI for these countries included 1.0 (mean RR, 1.09; 95% CI, 0.79-1.49), whereas this was not the case for countries with a high development index (mean RR, 1.48; 95% CI, 1.34-1.63).

**Figure 1** plots the log RRs associated with the studies that were entered into either analysis 1, 2, or 3. Although analysis 3 yielded more negative effect sizes (ie, a female excess) for individual studies, the general pattern of a male excess is essentially the same, implying that the sex difference in incidence may be relatively insensitive to methodological factors.

The results from this systematic review of studies published during the period between January 1980 and September 2001 point to a sex difference in the risk for schizophrenia. The risk remained significantly higher in men after controlling for potentially confounding factors (ie, age bias, criterion bias, and hospital bias). A separate meta-analysis of studies that minimized selection bias used semi-structured diagnostic interviews, and included 50 cases or more also revealed a significantly higher risk of schizophrenia in men. The magnitude of the male excess reported in this article is somewhat smaller than the female preponderance in Alzheimer disease, for which the mean odds ratio in a recent meta-analysis\(^5\) was estimated at 1.56.

Do these findings indicate that men are definitively at higher risk for developing schizophrenia? To enter an incidence study of schizophrenia, the patient has to contact a psychiatric service (stage 1), the treating physician has to recognize the symptoms of psychosis (stage 2), and the physician has to recognize a certain pattern indicative of schizophrenia (stage 3). Errors inducing sex bias could occur at each of these stages.

As far as stage 1 is concerned, it is uncertain whether women are more or less likely than men to seek treatment for psychotic symptoms. Ideally, large door-to-door surveys to detect all cases of schizophrenia in a given area could solve the question of whether the higher treated incidence rates for men are compensated for by large numbers of untreated women. However, the major surveys of the past 2 decades relied on structured diagnostic interviews, and the results are of questionable validity.\(^5\) We found no evidence that the male excess in treated incidence studies could be explained by biases occurring at stage 2 or 3. The analysis in which studies with an age cutoff of 64 years were included did not reveal a significantly smaller sex difference. The male excess remained significant when broad diagnostic criteria were applied and when the case-finding methods included inpatient and outpatient services. The lack of evidence for hospital bias as an explanation for the male excess is consistent with the results of a study from England.\(^6\) This study...
examined factors associated with hospital admission in 484 patients who made a first contact for nonaffective psychosis; it failed to find evidence of a sex bias.

It has been argued that schizophrenia is under-diagnosed at first hospital admission, that this bias is most pronounced in women, and that it can be reduced by the use of semistructured diagnostic interviews. The results of our separate meta-analysis on the subset of studies that used such interviews still showed a male preponderance in the incidence of schizophrenia. It is possible that men are more frequently misdiagnosed as having schizophrenia in the first stage of the illness and that the diagnosis is adjusted at a later stage (to bipolar disorder, for example). However, this is at odds with the findings of a large and meticulous study that did not observe sex differences in changes of diagnosis after 2 years.

Nonetheless, biases may have occurred at stage 2 or 3. The possibility remains that some physicians do not recognize psychotic symptoms and do not refer the patient for a semistructured diagnostic interview. Physicians may also recognize these phenomena but fail to draw the conclusion that the patient has a psychotic disorder and erroneously label these symptoms as dissociative phenomena or (transient) psychotic phenomena within the framework of a borderline personality disorder. The only way to counteract this kind of bias is to examine all patients who make a first contact for any psychiatric disorder by means of diagnostic interviews.

Authors who comment on methodological pitfalls of research concerning the sex difference in incidence of schizophrenia are almost exclusively concerned with bias that favors male predominance. However, bias might obscure a higher male risk as well. Specifically, mortality due to schizophrenia is more than twice as high in men as in women, and it is possible that this male excess also applies to subjects who commit suicide before they contact a psychiatric service. Finally, to enhance methodological homogeneity, we focused on studies using ICD criteria. Notably, there is no evidence that ICD criteria are more valid than DSM-III-R or DSM-IV criteria.

When summarizing the evidence in light of potential methodological limitations, we are left with 2 possibilities, which are illustrated graphically in Figure 2. The first possibility is that the real incidence of schizophrenia is identical for men and women but that it is more readily detected in men because of greater severity of the disorder (Figure 2A). Additional factors might be that women with psychosis are less likely to seek treatment and that their symptoms go unrecognized by physicians. The second possibility is that the real incidence is indeed larger for men (Figure 2B). As indicated in Figure 2, the detection threshold, which is influenced by severity of the disorder, is an important variable that complicates the interpretation of epidemiological studies.

Evidence shows that the proportion of men with deficit schizophrenia is even higher than that of men with schizophrenia in general. A recent meta-analysis of studies of patients with deficit schizophrenia found a mean odds ratio of 1.75, although this review was based on experimental rather than epidemiological studies. The Roscommon Family Study, a population-based study on the treated prevalence of schizophrenia, found that the proportion of men in the deficit group was 91% compared with 63% in the nondeficit group.

If there is a higher incidence of schizophrenia in men, has this always been the case? Ødegård estimated the lifetime risk for hospitalized schizophrenia on the basis of first admissions to Norwegian hospitals during the period from 1926 to 1935 and found a male excess of only 7%. It could be argued, therefore, that there was a slight excess during the first three quarters of the 20th century and that this excess has increased significantly since then. The results of our categorical analysis support this interpretation because RRs were higher in study samples from after 1980. Possible factors that could lead to an increase in male preponderance are the use of illicit drugs, which could precipitate schizophrenia in those who are genetically at risk, or the introduction of oral contraceptives because estrogens may have a protective effect in women. It has also been suggested that the incidence of schizophrenia is declining and that this decline is more prominent in women, but the evidence is inconclusive.

The fact that the sex difference was not apparent in countries with a medium development index may also point to effects of illicit drugs or oral contraceptives; the use of both is less widespread in developing countries. It may imply, however, that factors associated with industrialization...
CONCLUSIONS

This meta-analysis documents an apparently higher risk for men of developing schizophrenia, as reported in the published literature during the last 2 decades. However, given the methodological concerns, the existence of a true sex difference remains controversial. The possibility that schizophrenia may go unrecognized to a larger extent in women than in men has not been completely ruled out. We conclude that the male sex is a major risk factor for a more severe and therefore more easily recognizable form of schizophrenia.

Submitted for publication March 20, 2002; final revision received November 8, 2002; accepted November 15, 2002.

Dr Alenm was supported by Vernieuwingssimpuls Grant 016-026-027 from the Netherlands Organization for Scientific Research, The Hague.

We thank J. J. Hox, PhD, for statistical advice, and H. Schmack, PhD, for his help with Figure 2. We also thank El Saadi, MD, of the Schizophrenia Research Centre at Brisbane, Queensland, Australia, for providing the database of schizophrenia incidence studies.

Corresponding author and reprints: Jean-Paul Selten, MD, PhD, Department of Psychiatry, A00.241, University Medical Center, Heidelberglaan 100, NL-3584 CX Utrecht, The Netherlands (e-mail: j.p.selten@azu.nl).

REFERENCES

70. Susser E, Wanderling J. Epidemiology of nonaffective acute remitting psychosis vs schizophrenia: sex and sociocultural setting. Arch Gen Psychiatry. 1994;51:294-301.