

INTRODUCTION

Epidemiologic research in the field of psychiatry has shown that gender is a crucial variable influencing the distribution and features of mental disorders. Thus, the incorporation of a gender-related perspective into psychiatric research may have important implications for clinical practice, public health policy and theory. On a clinical ground, since gender can be considered an ‘immutable sociodemographic variable’ that cannot be influenced by the disease, it would be useful to identify gender-specific factors that may influence and predict symptom patterns, co-morbidity, course, outcome and response to treatment. From the public health perspective, gender differences in rates of a disorder and its characteristics may help to understand the health needs and treatment expectations of a population, to set the techniques for the monitoring of the health of a population, to plan treatment decisions and health services, to allocate the existing resources or to search for new resources. Theoretically, since gender is a proxy for a complex of psychosocial and biological variables, much can be known about specific etiologic or risk factors that underlie the association, interactions or processes by which vulnerability progresses into clinical disorders (Goodwin & Blehar, 1993).

The aim of this work is to report the main epidemiologic findings on gender differences in affective disorders and schizophrenia. In addition, methodologic difficulties that might bias research findings and be artifactual determinants of gender differences are discussed. Finally, the main factors that can best explain the gender differences in affective disorders and schizophrenia will be outlined as suggested by different theoretical approaches.

METHOD

Papers were collected using four strategies. First, the Medline data base was searched for the years 1985 to January 1995, using the keyword 'mental disorders' matched with any of the following in the title or abstract: 'sex', 'gender', 'male', 'female', 'men', 'women', 'boys', and 'girls'. This search was originally aimed at locating papers on gender differences in any type of mental disorders, not only in affective disorders and schizophrenia. Thus, we integrated this search by consulting the Medline CD ROM database for the period 1985 to May 1996, using the keywords 'schizophrenia', 'affective disorder', 'bipolar disorder', 'manic disorder', 'depressive disorder', 'mania', 'depression' and 'dysthymia' matched with the keywords 'sex' and 'gender' in the title or abstract. Second, literature was supplemented by hand-searching psychiatric journals and by consulting recent published reviews and bibliographies. Third, secondary references were identified by examining the reference sections of the papers located with the methods listed above. Finally, seven renowned experts in the field (see Acknowledgements) were contacted by mail and asked to select the most relevant papers recently published on gender differences in affective disorders and schizophrenia.

AFFECTIVE DISORDERS

INTRODUCTION

A wide variety of principles and criteria are being used in the assessment and classification of affective disorders. For example, Angold (1988) has suggested that at least eight distinctions can be made in the use of the term 'depression', some of which consider 'depression' as referring to a single item or state or trait and others refer to a mood state plus its biological, cognitive and/or behavioural concomitants. Similarly, Farmer & McGuffin (1989) have reviewed the 'contemporary confusion' surrounding the classification of depression and noted how the approaches advocated by nosologists varied considerably, some subdividing depression into one or two subtypes, while others proposing multiple different categories. Although different diagnostic concepts may serve specific research or clinical purposes, few of them have been sufficiently explored and validated (Jablensky, 1987). Indeed, research in the field of affective disorders faces specific methodological difficulties. Since biological or trait markers are relatively few and of unclear specificity, they are of little help in the identification of homogeneous diagnostic subtypes. Moreover, given the dimensional nature of much psychiatric morbidity and the absence of rarity points or interruptions in the symptom distribution, the definition of psychiatric 'caseness' is often arbitrary (Kendell, 1988). In light of these limitations, we have decided to present findings according to well-known systems for classifying mental disorders (e.g., the Research Diagnostic Criteria, the International Classification of Diseases, and the Diagnostic and Statistical Manual of Mental Disorders), since they have tried to reduce heterogeneity of diagnostic categories and limit criterion variance to an acceptable degree.

Although several forms of affective disorders have been investigated and described, epidemiologic evidence pertains mainly to bipolar disorder and major depression. Thus, these disorders are the main focus of the present work. In addition, epidemiologic findings on dysthymia (DSM-III - APA, 1980 - and DSM-III-R - APA, 1987), intermittent depression (Research Diagnostic Criteria - Spitzer et al., 1978) and brief recurrent depression (ICD-10 - WHO, 1992) are presented together under the heading 'persistent and recurrent depression'. A detailed review of similarities and differences between these diagnostic categories has been recently published after a working party of renowned experts summoned by the World Psychiatric Association (Costa e Silva & Freeman, 1994).

In the present work, epidemiologic findings are reported on gender differences in affective disorders among the adults, since published research suggests that gender differences are limited or even absent in either childhood or old age (Jorm, 1987; Burvill, 1995). Possible reasons for this age-related effect will be discussed alongside with the factors accounting for gender differences in affective disorders.

PREVALENCE OF MAJOR DEPRESSION

Current prevalence rates

Table I sets out the results of general population studies investigating current prevalence rates of major depression by sex of respondents. Current major depression was defined as criteria for major depression being fulfilled at the time of the examination or during the previous month.

Three studies were carried out in the United States. Weissman & Myers (1978) reported rates of affective disorders from the first epidemiologic survey applying Research Diagnostic Criteria to a community sample; the findings refer to a 1975 follow-up of a probability sample initially selected and studied in 1967. Prevalence estimates reported by Regier et al. (1993) were derived from the NIMH-Epidemiologic Catchment Area Study, a large-scale survey in which five sites geographically distributed throughout the United States were each randomly sampled to yield an estimate of rates of psychiatric disorders among noninstitutionalized adults aged 18 years and older. Finally, Blazer et al. (1994) presented one-month prevalence rates of major depression from the National Comorbidity Survey, in which a structured psychiatric interview was administered to a representative national sample of community residents in the United States aged 15 to 54 years. Although direct comparison between these studies is limited by a number of methodological factors (including differences in research instruments, diagnostic criteria, age range and size of the samples), a common finding was that rates of major depression were higher among females, with a female-to-male sex ratio ranging between 1.6 and 1.8.

In Florence, Italy, Faravelli et al. (1990) reported the highest female-to-male sex ratio (3.2) in a sample of subjects randomly selected from the lists of seven primary care physicians with post-graduate training in psychiatry; respondents were interviewed by primary care physicians, using a flow-chart that included the hierarchical system of diagnosis for affective disorders drawn from DSM-III.

Females were at increased risk for major depression compared to males also in Hollifield et al.'s (1990) study, in which a random sample of the families living in a small lowland town in Leshoto was examined. Moreover, when rates of current psychiatric disorders were corrected for gender and for alcohol abuse (due to a higher proportion of males being missed at interview and a sizeable proportion of individuals with major depression abusing alcohol), the same female-to-male sex ratio (1.6) was found, estimated rates for major depression being 7.1% in males and 11.7% in females.

Finally, Stefánsson et al. (1994) investigated rates of psychiatric disorders in a cohort of subjects including half of the population born in Iceland in 1931 and living there on December 1986.

Table I - Current prevalence rates of major depression from general population studies

| Author Country, time | Sample (N) | Age range (years) | Instruments Diagnostic Criteria | Rates (%) | | | Female-to-Male Sex Ratio |
|--|---------------|----------------------|------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|-----------------------------|
| | | | | Males | Females | Total | |
| Weissman & Myers (1978)* USA, 1975-76 | 511 | 26 and over | SADS RDC | 3.2 | 5.2 | 4.3 | 1.6 : 1 |
| Regier et al. (1993)** USA, 1980-83 | 18,571 | 18 and over | DIS DSM-III | 1.6 | 2.9 | 2.2 | 1.8 : 1 |
| Blazer et al. (1994)** USA, 1990-92 | 8,098 | 15 - 54 | CIDI DSM-III-R | 3.8 | 5.9 | 4.9 | 1.6 : 1 |
| Faravelli et al. (1990)* Italy, 1984 | 1,000 | 15 and over | Flow chart DSM-III | 1.3 | 4.1 | 2.8 | 3.2 : 1 |
| Hollifield et al. (1990)** Leshoto, 1986-1987 | 356 | 19 - 93 | DIS DSM-III | 8.8 | 14.5 | 12.4 | 1.6 : 1 |
| Stefansson et al. (1994)** Iceland, 1987-88 | 862 | 55-57 | DIS DSM-III | 0.9 ^a 0.5 ^b | 2.9 ^a 0.5 ^b | 1.9 ^a 0.5 ^b | 3.2 : 1 1.0 : 1 |
| | | | | 0.7 ^c | 1.7 ^c | 1.2 ^c | 2.4 : 1 |

* Major depression present at interview

** 1-month prevalence rates for major depression

^a Major depressive episode (total)

^b Major depression, single episode

^c Major depression, recurrent

Rates of major depression were low compared to the other studies. The preponderance of major depression among females was accounted for by recurrent major depression, whereas a female-to-male sex ratio equal to 1.0 was obtained for subjects suffering from a single episode of major depression over the month preceding the examination.

For comparison, Table II shows the results of studies which investigated rates of depressive disorders in the general population, using the Present State Examination. The Present State Examination is a semistructured psychiatric interview for eliciting and rating psychiatric symptoms present during the month preceding the evaluation; the interview is supplemented by a computer program (CATEGO), which allows the allocation of patients into classes according to their current symptoms (Wing & Sturt, 1978). The decision to consider these studies separately was based on two reasons. First, the Present State Examination by itself does not generate diagnoses, although it provides a symptom profile that allows the approximate allocation of patients into one of the ICD-

Table II - One-month prevalence rates of depressive disorders according to PSE-ID-CATEGO and ICD-8

| Author Country, time | Sample (N) | Age range (years) | Rates (%) | | | Female-to-Male Sex Ratio |
|--|---------------|----------------------|-------------------|-------------------|------------------|-----------------------------|
| | | | Males | Females | Total | |
| Orley & Wing (1979) Uganda, 1972 | 206 | 18 - 65 | 14.3 ^a | 22.6 ^a | — | 1.6 : 1 |
| Henderson et al. (1979) Australia, 1977 | 756 | 18 and over | 2.6 ^b | 6.7 ^b | — | 2.6 : 1 |
| Bebbington et al. (1981) UK, n.r. | 310 | 18-64 | 4.8 ^b | 9.0 ^b | 7.0 ^b | 1.9 : 1 |
| Mavreas et al. (1986) Greece, n.r. | 489 | 18-74 | 4.3 ^b | 10.1 ^b | 7.4 ^b | 2.1 : 1 |
| Mavreas & Bebbington (1987) UK*, n.r. | 219 | 18-64 | 4.1 ^b | 7.0 ^b | 5.5 ^b | 1.7 : 1 |
| Lethinen et al. (1990) Finland, 1985-87 | 747 | 30-80 | 2.4 ^b | 6.5 ^b | 4.6 ^b | 2.7 : 1 |

* Study conducted among Greek Cypriot immigrants in London

^a Depressive disorders not otherwise specified

^b Including ICD-8 categories 296.2 (manic-depressive psychosis, depressed type) and 300.4 (depressive neurosis)
n.r. = not reported

8 categories. Second, the investigators using the Present State Examination in surveys of the general population reported overall prevalence rates of depressive disorders without distinction between 'depressive neurosis (300.4)' and 'manic-depressive psychosis, depressed type (296.2)', whereas recent epidemiologic findings on sex and age distribution of bipolar disorder, major depression and dysthymia support the separation of bipolar disorder from other types of affective disorders (Weissman et al., 1988). As a result of differences in geographical location, survey designs and socio-demographic characteristics of populations where surveys were carried out, only gross comparisons can be made among these studies. In spite of these limitations, all the studies pointed to a preponderance of depressive disorders among females, with a female-to-male sex ratio ranging between 1.6 and 2.7.

Six-month prevalence rates

Table III sets out the results of general population studies investigating six-month prevalence rates of major depression by sex of respondents. The six-month prevalence rates reported here

included those subjects who met lifetime criteria for major depression and experienced relevant symptoms or an episode of the disorder during the six months preceding the examination.

Six studies used the Diagnostic Interview Schedule, a structured psychiatric interview designed for administration by lay interviewers in large-scale epidemiologic surveys to detect psychiatric disorders according to the DSM-III criteria. Four of these studies relied on a two-stage probability sample design: at the first stage, households were systematically sampled from lists of residential addresses; at the second stage, a listing of residents was obtained for each sample household and one household member was randomly selected among those eligible. Canino et al. (1987) selected individuals living in households throughout Puerto Rico as well as household members temporarily away and those in institutions. Bland et al. (1988a) examined a sample of community residents in Edmonton, Canada, aged 18 years and older in addition to a systematic sample of residents in a nursing home/auxiliary hospital group; the findings presented here refer to the noninstitutionalized individuals only. In New Zealand, Oakley-Browne et al. (1989) selected household members aged 18 to 64 years resident in the Christchurch Urban Area, including the city itself, suburbs and the semi-rural margins. Since affective disorders and eating disorders were of particular interest, females aged 18 to 44 years were oversampled to increase the yield of these disorders. In France, Lepine et al. (1989) conducted their survey in Savigny, a newly built town located near Paris. The rates of major depression in Christchurch were clearly higher than at the other three sites; there was a predominance of females as compared to males in the rates of major depression at all sites, with a female-to-male sex ratio ranging between 1.3 and 2.4.

Higher rates of major depression among females were also found in the other two studies using the Diagnostic Interview Schedule and DSM-III criteria. In the National Survey of Deviant Behavior conducted in the United States only young adults aged 18 to 24 years were assessed for psychiatric status (Elliott et al., 1985). In the Icelandic sample studied by Stefánsson et al. (1994) recurrent major depression was responsible for the predominance of females in the overall rates of major depression, whereas the two sexes did not differ in reporting a single episode of major depression.

Finally, Levav et al. (1993) selected a probability sample of first generation Jewish Israelis of either European or North-African parents using a two-stage design, in which final diagnoses were made by psychiatrists on the basis of the Schedule for Affective Disorders and Schizophrenia-Research Diagnostic Criteria. The female-to-male sex ratio for major depression was slightly greater than 1.0 when major depression was assessed both at the probable and at the definite level of diagnostic confidence.

Table III - Six-month prevalence rates of major depression from general population studies

| Author Country, time | Sample (N) | Age range (years) | Instruments Diagnostic Criteria | Rates (%) | | | Female-to-Male Sex Ratio |
|--|---------------|----------------------|------------------------------------|--|--|--|-------------------------------|
| | | | | Males | Females | Total | |
| Canino et al. (1987) Puerto Rico, 1984 | 1,513 | 18 - 64 | DIS DSM-III | 2.4 | 3.3 | 3.0 | 1.3 : 1 |
| Bland et al. (1988a) Canada, 1983-86 | 3,258 | 18 and over | DIS DSM-III | 2.5 | 3.9 | 3.2 | 1.6 : 1 |
| Oakley-Browne et al. (1989) New Zealand, 1986 | 1,498 | 18 - 64 | DIS DSM-III | 3.4 | 7.1 | 5.3 | 2.1 : 1 |
| Lepine et al. (1989) France, n.r. | 749 | 18 and over | DIS DSM-III | 1.5 | 3.6 | — | 2.4 : 1 |
| Elliott et al. (1985) USA, 1983 | 1,496 | 18 - 24 | DIS DSM - III | 3.1* | 8.0* | 5.5* | 2.6 : 1 |
| Stefansson et al. (1994) Iceland, 1987-88 | 862 | 55 - 57 | DIS DSM-III | 1.1 ^a 0.9 ^b 0.7 ^c | 3.6 ^a 0.7 ^b 1.7 ^c | 2.3 ^a 0.8 ^b 1.2 ^c | 3.3 : 1 0.8 : 1 2.4 : 1 |
| Levav et al. (1993) Israel, 1982-88 | 2,741 | 24 - 33 | SADS RDC | 3.8 ^d 2.6 ^e | 4.5 ^d 3.4 ^e | 4.2 ^d 3.0 ^e | 1.2 : 1 1.3 : 1 |

* Data derived from Weissman et al. (1988)

^a Major depressive episode (total)

^b Major depression, single episode

^c Major depression, recurrent

^d Either probable or definite level of diagnostic accuracy

^e Only definite level of diagnostic accuracy

n.r. = not reported

Twelve-month prevalence rates

Table IV sets out the results of general population studies investigating 12-month prevalence rates of major depression by sex of respondents. The 12-month prevalence rates reported here included those subjects who met lifetime criteria for major depression and had experienced relevant symptoms or an episode of the disorder during the 12 months preceding the examination.

In five studies, rates of major depression were expressed according to the DSM-III criteria. In the United States, the nationwide survey carried out by Uhlenhuth et al. (1983) primarily to determine the prevalence and patterns of psychotropic drug prescriptions and the NIMH-Epidemiologic Catchment Area Study (Robins & Regier, 1991) were based on a stratified multistage probability

sample of subjects in the noninstitutionalized civilian population. In Florence, Italy, Faravelli et al. (1990) selected a random sample of the general population from the lists of primary care physicians. Ernst & Angst (1992) selected a cohort of young males and females from the Kanton Zürich in Switzerland, who were born in 1957 and 1958. Finally, Stefánsson et al. (1994) selected half of the population born in Iceland in 1931.

On the other hand, prevalence rates of major depression were expressed according to the DSM-III-R criteria in the survey conducted in Savigny, France (Lepine et al., 1993) and in the National Comorbidity Survey in the United States (Kessler et al., 1994).

Despite wide variation in rates of major depression across the studies (between 2.7% and 10.3%), rates were higher in females compared to males and this trend was consistent across all the studies, with a female-to-male sex ratio ranging between 1.7 and 2.9.

Table IV - Twelve-month prevalence rates of major depression from general population studies

| Author Country, time | Sample (N) | Age range (years) | Instruments Diagnostic Criteria | Rates (%) | | | Female-to-Male Sex Ratio |
|--|---------------|----------------------|-------------------------------------|--|--|--|-------------------------------|
| | | | | Males | Females | Total | |
| Uhlenhuth et al. (1983) USA, 1979 | 3,161 | 18 - 79 | Symptom checklist DSM - III | 2.8 | 6.9 | 5.1 | 2.5 : 1 |
| Robins & Regier (1991) USA, 1980-83 | 18,571 | 18 and over | DIS DSM-III | 1.4 | 4.0 | 2.7 | 2.9 : 1 |
| Faravelli et al. (1990) Italy, 1984 | 1,000 | 15 and over | Flow chart DSM-III | 3.5 | 8.8 | 6.3 | 2.5 : 1 |
| Ernst & Angst (1992) Switzerland, 1979 | 591 | 20 - 21 | Semistructured interview DSM-III | 4.8 | 10.8 | — | 2.3 : 1 |
| Stefansson et al. (1994) Iceland, 1987-88 | 862 | 55 - 57 | DIS DSM-III | 1.8 ^a 0.9 ^b 0.9 ^c | 4.0 ^a 1.0 ^b 1.9 ^c | 2.9 ^a 0.9 ^b 1.4 ^c | 2.2 : 1 1.1 : 1 2.1 : 1 |
| Lepine et al. (1993) France, n.r. | 1,746 | 18 and over | DIS DSM-III-R | 3.4 | 6.0 | — | 1.8 : 1 |
| Kessler et al. (1994) USA, 1990-92 | 8,098 | 15 - 54 | CIDI DSM-III-R | 7.7 | 12.9 | 10.3 | 1.7 : 1 |

^a Major depressive episode (total)

^b Major depression, single episode

^c Major depression, recurrent

n.r. = not reported

Lifetime prevalence rates

Table V sets out the results of general population studies investigating lifetime prevalence rates of major depression by sex of respondents. The lifetime prevalence rates of major depression reported here included those subjects who ever met the criteria for the disorder during the entire lifespan prior to the examination.

Nine studies used the Diagnostic Interview Schedule and DSM-III criteria. The surveys carried out in Puerto Rico (Canino et al., 1987), Edmonton (Bland et al., 1988a) and Christchurch (Wells et al., 1989) and the NIMH-Epidemiologic Catchment Area Study (Robins & Regier, 1991) were based on household probability samples, with estimated rates for major depression being obtained by weighting the collected data to correct for the selection procedure. The studies carried out in Savigny (Lepine et al., 1989) and in Iceland (Stefánsson et al., 1991) were drawn as random samples and, thus, the weighting procedure was not required. The Taiwan Psychiatric Epidemiology Project sampled three populations in metropolitan, township and rural areas, using a multistage random sampling method that did not create skewed sampling weights in a specific sampling area (Hwu et al., 1989). The Korean Epidemiologic Study of Mental Disorders was a nationwide survey including households in Seoul and in rural locations scattered over the country; in each household all family members aged 18 to 65 years were interviewed, provided that they had lived there for more than three months (Lee et al., 1990a, b). The Munich Follow-up Study (Wittchen et al., 1992) was a seven-year prospective and retrospective follow-up investigation of a stratified random sample of the general population of former West Germany; the stratification method used at follow-up included all the individuals reporting high scores on the clinical rating scales at the baseline evaluation plus a 39.8% random sample of those with low scores. Finally, the Shatin Community Mental Health Survey randomly selected households in Hong Kong, with one member between 18 and 64 years of age being randomly interviewed from each selected household (Chen et al., 1993). The lifetime rates of major depression varied widely by site, ranging between 3.3% and 16.8%. However, there was a predominance of females as compared to males in the rates of major depression at all sites, with a female-to-male sex ratio ranging between 1.4 and 3.4.

Two studies were based on different research instruments and diagnostic criteria. Weissman & Myers (1978) relied on the Schedule for Affective Disorders and Schizophrenia-Research Diagnostic Criteria. Kessler et al. (1994) used the Composite International Diagnostic Interview, a structured interview that was derived from the Diagnostic Interview Schedule and generates diagnoses according to the DSM-III-R criteria. Both studies showed a predominance of females as compared to males in the rates of major depression.

In Table V are also reported (when available) the female-to-male sex ratios resulting from rates of major depression at each site being standardized to the Epidemiologic Catchment Area five site household sample. This work was undertaken by the Cross National Collaborative Group, formed in 1990 by investigators having epidemiologic community data based on the Diagnostic Interview

Table V - Lifetime prevalence rates of major depression from general population studies

| Author Country, time | Sample (N) | Age range (years) | Instruments Diagnostic Criteria | Rates (%) | | | Female-to-Male | Female-to-Male |
|--|----------------|----------------------|------------------------------------|--|---|---|-------------------------------|---------------------------|
| | | | | Males | Females | Total | Sex Ratio | Sex Ratio* |
| Canino et al. (1987) Puerto Rico, 1984 | 1,513 | 18 - 64 | DIS DSM-III | 3.5 | 5.5 | 4.6 | 1.6 : 1 | 1.8 : 1 |
| Bland et al. (1988b) Canada, 1983-86 | 3,258 | 18 and over | DIS DSM-III | 5.9 | 11.4 | 8.6 | 1.9 : 1 | 1.9 : 1 |
| Wells et al. (1989) New Zealand, 1986 | 1,498 | 18 - 64 | DIS DSM-III | 8.8 | 16.3 | 12.6 | 1.9 : 1 | 2.1 : 1 |
| Robins & Regier (1991) USA, 1980-83 | 18,571 | 18 and over | DIS DSM-III | 2.6 | 7.0 | 4.9 | 2.7 : 1 | 2.6 : 1 |
| Lepine et al. (1993) France, n.r. | 749 | 18 and over | DIS DSM-III | 8.5 | 21.9 | — | 2.6 : 1 | 2.1 : 1 |
| Stefansson et al. (1991) Iceland, 1987-88 | 862 | 55 - 57 | DIS DSM-III | 2.9 ^a 1.4 ^b 1.4 ^c | 7.8 ^a 2.4 ^b 4.6 ^c | 5.3 ^a 1.9 ^b 2.9 ^c | 2.7 : 1 1.7 : 1 3.3 : 1 | |
| Hwu et al. (1989) Taiwan, 1982-85 | 11,004 | 18 and over | DIS DSM-III | 7.3 ^d 9.6 ^e 6.1 ^f | 10.2 ^d 24.7 ^e 14.1 ^f | 8.8 ^d 16.8 ^e 9.7 ^f | 1.4 : 1 2.6 : 1 2.3 : 1 | 1.6 : 1 (total sample) |
| Lee et al. (1990a,b) Korea, n.r. | 3,134 1,966 | 18 - 65 | DIS DSM-III | 2.4 ^g 2.9 ^h | 4.1 ^g 4.1 ^h | 3.3 ^g 3.5 ^h | 1.7 : 1 1.4 : 1 | 2.0 : 1 (total sample) |
| Wittchen et al. (1992) Germany, 1981 | 483 | 25 - 64 | DIS DSM-III | 4.0 | 13.6 | 9.0 | 3.4 : 1 | 3.5 : 1 |
| Chen et al. (1993) Hong Kong, 1984-86 | 7,229 | 18 - 64 | DIS DSM-III | 1.3 ^d 0.7 ^e 0.6 ^f | 2.4 ^a 1.3 ^b 1.1 ^c | — — — | 1.9 : 1 1.9 : 1 1.9 : 1 | |
| Weissman & Myers (1978) USA, 1975-76 | 511 | 26 and over | SADS RDC | 12.3 | 25.8 | 20.0 | 2.1 : 1 | |
| Kessler et al. (1994) USA, 1990-92 | 8,098 | 15 - 54 | CIDI DSM-III-R | 12.7 | 21.3 | 17.1 | 1.6 : 1 | |

* Rates of major depression at each site were standardized to the Epidemiologic Catchment Area five site household sample for the age group 18-64 years

a Major depressive episode (total)

b Major depression, single episode

c Major depression, recurrent

d Metropolitan Taipei

e Small towns

f Rural villages

g Urban area

h Rural area

n.r. = not reported

Schedule and DSM-III criteria, in order to analyze the data and compare the findings using a common data plan and the same definitions (Weissman et al., 1993). The Epidemiologic Catchment Area Study was designated as the reference sample since it had the largest sample size and a sampling scheme robust enough to ensure stability of findings. Prevalence rates standardized in this way produced estimates as if the population at each site had the same sex and age distribution of the Epidemiologic Catchment Area Study sample. Since the age ranges sampled at each site differed, analyses were limited to individuals aged 18 to 64 years. Females retained their predominance over males in rates of major depression at all sites, the female-to-male sex ratio ranging between 1.6 and 2.6.

PREVALENCE OF PERSISTENT AND RECURRENT DEPRESSION

Table VI sets out the results of general population studies investigating lifetime prevalence rates of dysthymia by sex of respondents.

Nine studies used the Diagnostic Interview Schedule and DSM-III criteria; since dysthymia is a chronic disorder, the Diagnostic Interview Schedule does not attempt to define onset or remission and only lifetime rates of the disorder are assessed. The surveys carried out in Puerto Rico (Canino et al., 1987), Edmonton (Bland et al., 1988b), Christchurch (Wells et al., 1989), Korea (Lee et al., 1990a, b), the United States (Robins & Regier, 1991) and Hong Kong (Chen et al., 1993) were based on household probability samples of the general population. The Taiwan Psychiatric Epidemiology Project sampled three populations in metropolitan, township and rural areas (Hwu et al., 1989). The study carried out in Iceland (Stefánsson et al., 1991) included half of the population born in Iceland in 1931 and living there in December 1986. Finally, the Munich Follow-up Study (Wittchen et al., 1992) was a seven-year prospective and retrospective follow-up investigation of a stratified random sample of the general population of former West Germany. The lifetime rates of dysthymia varied widely by site, ranging between 1.9% and 15.1%. There was a predominance of females as compared to males in the rates for dysthymia at all sites, with a female-to-male sex ratio ranging between 1.2 and 4.8.

On the other hand, in the National Comorbidity Survey the Composite International Diagnostic Interview was used to generate diagnoses according to the DSM-III-R criteria. A predominance of females was observed in lifetime prevalence rates of dysthymia (Table VI). The same study provided also twelve-month prevalence rates of dysthymia, these being 2.1% in males and 3.0% in females, with a female-to-male sex ratio of 1.4 (Kessler et al., 1994).

Intermittent depression, defined by the Schedule for Affective Disorders and Schizophrenia-Research Diagnostic Criteria, was investigated in a sample of first generation Jewish Israelis. At the

definite level of diagnostic confidence, six-month prevalence rates were 2.2% in males and 3.7% in females, with a female-to-male sex ratio of 1.7. The female-to-male sex ratio was 1.8, when intermittent depression was assessed according to both the definite and probable level of diagnostic confidence (Levav et al., 1993).

Finally, brief recurrent depression was investigated in a sample of young adults aged 23-24 years. Twelve-month prevalence rates were 3.9% in males and 4.9% in females, with a female-to-male sex ratio of 1.3. However, these rates were based on different thresholds for case-definition in males and females (3 criterion symptoms for males and 5 for females); using the same threshold in the two sexes, the female-to-male sex ratio was equal to 2.0 (Angst & Dobler-Mikola, 1985).

Table VI - Lifetime prevalence rates of dysthymia from general population studies

| Author Country, time | Sample (N) | Age range (years) | Instruments Diagnostic Criteria | Rates (%) | | | Female-to-Male Sex Ratio |
|--|----------------|----------------------|------------------------------------|---|---|---|-------------------------------|
| | | | | Males | Females | Total | |
| Canino et al. (1987) Puerto Rico, 1984 | 1,513 | 18 - 64 | DIS | 1.6 | 7.6 | 4.7 DSM-III | 4.8 : 1 |
| Bland et al. (1988b) Canada, 1983-86 | 3,258 | 18 and over | DIS DSM-III | 2.2 | 5.2 | 3.7 | 2.4 : 1 |
| Wells et al. (1989) New Zealand, 1986 | 1,498 | 18 - 64 | DIS DSM-III | 3.8 | 9.0 | 6.4 | 2.4 : 1 |
| Lee et al. (1990a,b) Korea, n.r. | 3,134 1,966 | 18 - 65 | DIS DSM-III | 1.8 ^a 1.3 ^b | 3.0 ^a 2.5 ^b | 2.4 ^a 1.9 ^b | 1.7 : 1 1.9 : 1 |
| Robins & Regier (1991) USA, 1980-83 | 18,571 | 18 and over | DIS DSM-III | 2.2 | 4.1 | 3.2 | 1.9 : 1 |
| Chen et al. (1993) Hong Kong, 1984-86 | 7,229 | 18 - 64 | DIS DSM-III | 1.1 | 2.8 | — | 2.6 : 1 |
| Hwu et al. (1989) Taiwan, 1982-85 | 11,004 | 18 and over | DIS DSM-III | 6.9 ^c 14.1 ^d 5.5 ^e | 11.4 ^c 16.2 ^d 14.1 ^e | 9.2 ^c 15.1 ^d 9.4 ^e | 1.7 : 1 1.2 : 1 2.6 : 1 |
| Stefansson et al. (1991) Iceland, 1987-88 | 862 | 55 - 57 | DIS DSM-III | 2.3 | 10.7 | 6.4 | 4.7 : 1 |
| Wittchen et al. (1992) Germany, 1981 | 483 | 25 - 64 | DIS DSM-III | 2.5 | 5.4 | 4.0 | 2.2 : 1 |
| Kessler et al. (1994) USA, 1990-92 | 8,098 | 15 - 54 | CIDI DSM-III-R | 4.8 | 8.0 | 6.4 | 1.7 : 1 |

^a Urban area

^b Rural area

^c Metropolitan Taipei

^d Small towns

^e Rural villages

n.r. = not reported

PREVALENCE OF BIPOLAR DISORDER

Since prevalence of bipolar disorder is expected to be low in the general population, only lifetime prevalence rates by sex of respondents are reported in Table VII. Even in this case, gender differences should be interpreted with caution, due to the small number of individuals that satisfied lifetime criteria for bipolar disorder in most of the studies. The lifetime prevalence rates reported here included those subjects who ever met the criteria for bipolar I disorder during the entire lifespan prior to the examination.

Nine studies were based on the Diagnostic Interview Schedule and DSM-III criteria. The surveys carried out in Puerto Rico (Canino et al., 1987), Edmonton (Bland et al., 1988b), Christchurch (Wells et al., 1989), Korea (Lee et al., 1990a, b), the United States (Robins & Regier, 1991) and Hong Kong (Chen et al., 1993) were based on household probability samples of the general population. The Taiwan Psychiatric Epidemiology Project was carried out in three distinct populations in metropolitan, township and rural areas (Hwu et al., 1989). The study carried out in Iceland (Stefánsson et al., 1991) included half of the population born in Iceland in 1931 and living there in December 1986. Finally, the Munich Follow-up Study (Wittchen et al., 1992) was a seven-year prospective and retrospective follow-up investigation of a stratified random sample of the general population of former West Germany. The lifetime rates of bipolar disorder were generally lower than 1% (range between 0.2% and 1.6%); the female-to-male sex ratio ranged between 0.1 and 3.5.

Two studies were based on different research instruments and diagnostic criteria. Levav et al. (1993) assessed a sample of first generation Jewish Israelis, using the Schedule for Affective Disorders and Schizophrenia-Research Diagnostic Criteria. Rates of bipolar I disorder tended to be higher in females at both the definite and probable level of diagnostic confidence (Table VII). For bipolar II disorder, lifetime prevalence rates were 0.8% in males and 0.3% in females at the definite level of diagnostic confidence, with a female-to-male sex ratio of 0.4; the female-to-male sex ratio was 0.5, when bipolar II disorder was assessed according to both the definite and probable level of diagnostic confidence. On the other hand, in the National Comorbidity Survey the Composite International Diagnostic Interview was used to generate diagnoses according to the DSM-III-R criteria, with lifetime rates for bipolar disorder being similar in males and females (Kessler et al., 1994).

Table VII - Lifetime prevalence rates of bipolar disorder from general population studies

| Author Country, time | Sample (N) | Age range (years) | Instruments Diagnostic Criteria | Rates (%) | | | Female-to-Male Sex Ratio |
|--|----------------|----------------------|------------------------------------|--|--|--|-----------------------------|
| | | | | Males | Females | Total | |
| Canino et al. (1987) Puerto Rico, 1984 | 1,513 | 18 - 64 | DIS DSM-III | 0.7 | 0.4 | 0.5 | 0.6 : 1 |
| Bland et al. (1988b) Canada, 1983-86 | 3,258 | 18 and over | DIS DSM-III | 0.7 | 0.4 | 0.6 | 0.6 : 1 |
| Wells et al. (1989) New Zealand, 1986 | 1,498 | 18 - 64 | DIS DSM-III | 0.5 | 0.9 | 0.7 | 1.8 : 1 |
| Lee et al. (1990a,b) Korea, n.r. | 3,134 1,966 | 18 - 65 | DIS DSM-III | 0.6 ^a 0.8 ^b | 0.3 ^a 0.1 ^b | 0.4 ^a 0.4 ^b | 0.5 : 1 0.1 : 1 |
| Robins & Regier (1991) USA, 1980-83 | 18,571 | 18 and over | DIS DSM-III | 0.7 | 0.9 | 0.8 | 1.3 : 1 |
| Chen et al. (1993) Hong Kong, 1984-86 | 7,229 | 18 - 64 | DIS DSM-III | 0.2 | 0.2 | 0.2 | 1.0 : 1 |
| Hwu et al. (1989) Taiwan, 1982-85 | 11,004 | 18 and over | DIS DSM-III | 1.6 ^c 1.2 ^d 1.2 ^e | 1.6 ^c 0.0 ^d 0.7 ^e | 1.6 ^c 0.7 ^d 1.0 ^e | 1.0 : 1 — 0.6 : 1 |
| Stefansson et al. (1991) Iceland, 1987-88 | 862 | 55 - 57 | DIS DSM-III | 0.2 | 0.2 | 0.2 | 1.0 : 1 |
| Wittchen et al. (1992) Germany, 1981 | 483 | 25 - 64 | DIS DSM-III | 0.0 | 0.5 | 0.2 | — |
| Levav et al. (1993) Israel, 1982-88 | 2,741 | 24 - 33 | SADS | 0.5* 0.2** | 0.8* 0.7** | 0.7* 0.5** | 1.6 : 1 3.5 : 1 |
| Kessler et al. (1994) USA, 1990-92 | 8,098 | 15 - 54 | CIDI DSM-III-R | 1.6 | 1.7 | 1.6 | 1.1 : 1 |

* Either probable or definite level of diagnostic confidence

** Only definite level of diagnostic confidence

a Urban area

b Rural area

c Metropolitan Taipei

d Small towns

f Rural villages

n.r. = not reported

FACTORS THAT MAY INFLUENCE GENDER DIFFERENCES IN PREVALENCE RATES OF AFFECTIVE DISORDERS

A consistent finding from general population surveys of affective disorders concerns the higher rates of depressive disorders in females compared to males, whereas no or inconsistent gender differences were detected for bipolar disorder. However, it is still controversial whether gender differences in depressive disorders are real or an artifact. Four main issues should be considered that might affect gender differences in rates of depressive disorders: I) the definition of caseness and measurement procedures; ii) the recall bias; iii) the course of disorder and mortality rates; iv) the geographical mobility.

i) Definition of caseness and measurement procedures

A few studies examined the hypothesis that gender differences in rates of major depression depends on the criteria used to make the diagnosis. The diagnostic criteria for major depression set out by Spitzer et al. (1978) (Research Diagnostic Criteria), the American Psychiatric Association (DSM-III, DSM-III-R, DSM-IV) and the World Health Organization (ICD-10) are based on a number of criterion symptoms that are associated with depressed mood. Any general tendency for females to report more criterion symptoms than males may contribute to females being more likely to meet diagnostic criteria for major depression, if the same number of symptoms is required to make a diagnosis in both sexes.

In Zürich, Angst & Dobler-Mikola (1984) examined the 12-month prevalence rates of depressive episodes in a cohort of young adults aged 22-23 years. Among probands that reported episodes of depressed mood lasting two weeks or longer, females reported more symptoms compared to males, the median being five symptoms for females and three for males. Had the same number of depressive symptoms been required for both sexes to be assigned a diagnosis, then different rates of depressive disorder would be expected in males and females. Indeed, for the three months preceding examination, the female-to-male sex ratio of 0.6 for episodes of depressed mood turned to 1.9 and 1.8 when the Research Diagnostic Criteria and DSM-III criteria for major depression were applied, respectively. Moreover, males and females suffering from an episode of depressed mood of at least two-week duration gave similar ratings of subjective impairment and were equally impaired socially and occupationally. Thus, when the social impairment criteria were included in the diagnostic decision tree, an equal prevalence of depression was found in males and females. The Authors concluded that episodes of depression occurred at similar rates in males and females and any actual female preponderance resulted from excess reporting of depressive symptoms.

Different results were obtained by Fennig et al. (1994), who investigated the 12-month prevalence rates of DSM-III-R major depression in a relatively homogeneous population of managers and professionals from a large corporation. When the female-to-male sex ratio in rates of depression was examined by progressively increasing the number of symptoms required for a diagnosis of depressive episode, the female preponderance existed at each level, even though it became more pronounced as the number of symptoms increased. For each possible cut-off at five or more criterion symptoms (as required in DSM-III-R) the female-to-male sex ratio was near or above 2.0. The female preponderance persisted even if different cut-offs were used for assigning a diagnosis to males and females, that is five criterion symptoms for males and six for females. Finally, the pattern of association between depressive symptomatology and occupational impairment was very similar for males and females and the female-to-male sex ratio exceeded unity at all levels of impairment beyond the mildest. Thus, the pattern of higher female-to-male sex ratios was similar for the two systems of case identification (i.e., that based on symptoms and that based on impairment).

Similar findings were reported also by Kessler et al. (1993), using the data from the National Comorbidity Survey. About 46% of the males and 58% of the females reported the lifetime occurrence of at least one period of depressed mood or diminished interest in most of the normal activities lasting two weeks or longer. If endorsement of these stem questions was the only requirement for a diagnosis of depression, the female-to-male sex ratio was 1.3. The female-to-male sex ratio became progressively higher as the required number of criterion symptoms was increased: it was 1.7 when four or more criterion symptoms were required, as in the DSM-III-R criteria, and 2.5 when all the eight depressive symptoms included in DSM-III-R had to be present.

A true gender difference in lifetime rates of major depression rather than a general trend for females to report more criterion symptoms was suggested also by Young et al. (1990a) in a sample of first-degree relatives of probands who participated in the NIMH-Collaborative Study of the Psychobiology of Depression. Among those individuals who experienced depressed mood or loss of interest or pleasure of at least one week duration (criterion I of the Research Diagnostic Criteria) and impaired functioning or treatment-seeking (criterion II of the Research Diagnostic Criteria), similar proportions of males and females reported no, one and two criterion symptoms associated to depression. A gender difference in symptom count was found once three or more criterion symptoms were present (i.e., the cut-off used by the Research Diagnostic Criteria to make a lifetime diagnosis of past probable major depressive episode). The great bulk of the gender difference was based on only the depressed mood and impairment criteria, while using the criterion of three or more symptoms resulted in an increase in the female-to-male sex ratio of only 9%. Since for both males and females the discontinuity in the curve of symptom frequency occurred at the threshold level for a diagnosis of major depression

according to the Research Diagnostic Criteria, there was no support to the use of gender-specific threshold criteria.

In summary, although there is a tendency for females to report more criterion symptoms associated to depression compared to males, the findings available so far suggest that this does not seem to account entirely for gender differences in rates of major depression.

A related issue is whether males and females differ in the clinical manifestations of depression during a diagnosed depressive episode. Frank et al. (1988) reported on gender differences in a sample of 230 patients with recurrent depression. Females were more likely to report increased appetite and weight gain and less likely to report weight loss than their male counterparts; females also experienced more somatic anxiety, expressed anger and hypochondriasis. Similar results were found by Young et al. (1990b), who examined a sample of 498 probands seeking treatment for current nonpsychotic unipolar major depression according to the Research Diagnostic Criteria. Statistically significant differences were detected for two of the 41 clinical features considered, with increased appetite and weight gain occurring more frequently in females compared to males. Analysis of severity for those features that applied to all individuals showed no significant differences between males and females. Finally, Angst & Dobler-Mikola (1984) showed that females were more likely to report disturbance of appetite and sleep compared to males; in addition, reporting of symptoms by males tended to decrease and to be less reliable over a one-year period.

One approach to interpreting the results of these studies is to accept gender differences in clinical manifestations of depression as real; as an alternative interpretation, these findings might simply reflect sociocultural differences in the way males and females perceive their symptoms and respond to them (e.g., males might be engaged in more denial of symptoms than females, according to their socially desirable self-image). Whatever interpretation is endorsed, it might be expected that differences between males and females in experiencing or reporting depressive symptoms might lead to gender-specific response patterns depending on the individual items included in the rating scales for depression. Building on recent advances in structural equation models, Stommel et al. (1993) tested the degree to which the 20-item Center for Epidemiologic Studies Depression Scale (Radloff, 1977) was 'factorially invariant' across groups of male and female cancer patients. Two items were identified as producing biased responses according to sex of respondents: males, who otherwise had the same level of depressive symptoms as females, were less likely to have 'crying spells', but more likely to reduce their verbal communication ('talked less') compared to equally depressed females. When three additional items were excluded on the basis of other psychometric deficiencies, a subset of 15 items was left that captured almost all of the information of the original 20-item scale, but were free of any gender bias. Nevertheless, gender differences in mean levels of depressive symp-

tomatology, although significantly reduced, were not eliminated when the shortened version of the scale was used.

The psychometric properties of another well-known rating scale for depression (i.e., the Beck Depression Inventory - Beck et al., 1979) were compared in male and female psychiatric outpatients with affective disorders according to the DSM-III criteria. Only one of the 21 items differentiated males and females, with males reporting more severe self-dissatisfaction compared to females. Overall, the findings supported the comparability of the psychometric properties of the scale for both males and females and separate sex norms were not necessary for interpreting the scores reported on the scale (Steer et al., 1989).

ii) Recall bias

It has been suggested that the preponderance of females in rates of major depression might be the result of a gender difference in recalling past depressive episodes. The introduction of structured psychiatric interviews to generate standardized diagnostic decisions has raised the issues of test-retest reliability and inter-rater agreement as measures of the instrument error. A related paradigm is that of temporal stability, with agreement between assessments given at widely separated points in time (e.g., successive hospital admissions) reflecting 'true' lifetime clinical state and the validity of the underlying constructs (Rice et al., 1992).

Comparison of short-term and lifetime prevalence rates of major depression may provide indirect evidence of instability of results. In the NIMH-Epidemiologic Catchment Area Study the ratio of lifetime to six-month prevalence rates of major depression was 2.0, indicating that half of the subjects reporting a lifetime episode of major depression had it in the six months preceding the examination (Weissman et al., 1988). Similar findings were reported by other general population studies using the Diagnostic Interview Schedule and DSM-III criteria: the ratios of lifetime to six-month prevalence rates of major depression were 1.5 in Puerto Rico (Canino et al., 1987), 2.3 in Iceland (Stefánsson et al., 1991; 1994), 2.4 in Christchurch (Wells et al., 1989; Oakley-Browne et al., 1989), 2.7 in Edmonton (Bland et al., 1988 a, b) and 3.0 in München (Wittchen et al., 1992). Although it cannot be excluded that subjects reporting major depression have depressive episodes that are long and/or highly recurrent, these findings suggest that remote episodes might be forgotten, resulting in an under-estimate of lifetime rates of depression. Indeed, Parker (1987) raised doubts about the validity of the lifetime prevalence estimates generated in the reports of the Epidemiologic Catchment Area Study and identified three reasons to suspect their accuracy: the ratio of lifetime and six-month prevalence data, the discordance with previous estimates of lifetime morbidity, and the curvilinear association of lifetime prevalence estimates with increasing age.

The issue of temporal stability in the assignment of psychiatric diagnoses has received direct attention by a number of studies. Mazure & Gershon (1979) interviewed a sample of adult patients hospitalized for a major affective disorder, their first-degree relatives and controls on two occasions separated by a mean interim time of 6.7 months. Overall, 18 of the 21 subjects with major affective disorders on any interview were consistently diagnosed and 10 of the 11 subjects receiving a diagnosis of unipolar depression at the first interview were diagnosed so also at the second. Although high reliability was found for the number of depressive symptom groups, there was only modest reliability between the number of affective episodes reported each time.

Similar results were found by Andreasen et al. (1981), who examined the six-month reliability of lifetime diagnoses in a sample of 50 relatives of affectively ill probands and control subjects. High reliability was shown for the diagnosis of major depressive episode and for the items concerning depressive symptoms (with the exception of agitation-retardation and tiredness), whereas interviewers reported low levels of agreement in the number of depressive episodes that occurred in the past.

Finally, Rice et al. (1992) collected data on 1,629 relatives of probands with affective disorders that were interviewed twice, six years apart. Of those with a lifetime diagnosis of major depression at initial assessment (24% of the sample), 74% were positive also on second interview. The misclassification decreased as disorder severity increased, since individuals reporting more depressive symptoms, more episodes of depression or some type of treatment at initial assessment as well as those who attempted suicide or were incapacitated had higher stability in reporting; instead, sex of respondents was not significantly related to stability.

Although these studies provide important clues on the issue of stability of diagnosis over time, they were not conducted in general population samples. This may be a possible limitation in the generalisation of their results, since diagnostic stability tended to be strongly influenced by such covariates at initial evaluation as the number of symptoms reported and treatment-seeking behaviour. Therefore, different levels of diagnostic stability may be expected in samples including milder cases, like those drawn from the general population. Thus, Bromet et al. (1986) assessed the stability of lifetime episodes of major depression in a community sample of women interviewed on two occasions, 18 months apart. Overall, stability was poor with 'k' for definite major depression being 0.41. Of the 144 women that were diagnosed at either interview as depressed for the period preceding first interview, only 38% consistently reported episodes of major depression on both occasions, 42% of those initially reporting an episode subsequently 'forgot' that episode, whereas 20% of those who initially denied a lifetime episode subsequently 'remembered' an episode. Individuals receiving a diagnosis of major depression at both interviews reported more depressive symptom groups than those diagnosed as depressed

at either first or second interview. Finally, women who reliably reported lifetime episodes of depression were consistent about details such as medication use, but were inconsistent about number of depressive episodes, length of longest episode and age at first episode. A similar analysis of the data collected during the NIMH-Epidemiologic Catchment Area study revealed that 61% of the 529 respondents with a lifetime diagnosis of major depression at baseline failed to receive that diagnosis when interviewed again one year later (Dohrenwend, 1989).

The findings reported above strongly suggest that passage of time may affect recall of previous episodes of depression. Other studies have specifically investigated possible gender differences in temporal stability which might be artifactual determinants of a female preponderance in rates of major depression. In their cohort of young adults, Angst & Dobler-Mikola (1984) found that 7.4% of males and 4.8% of females reported depressive episodes during the three months preceding examination, whereas only 1.8% of males as opposed to 10.8% of females reported depressive episodes between four and 12 months prior to examination. The Authors argued that the preponderance of females in 12-month rates of major depression could be ascribed to males reporting previous depressive episodes in a less reliable way.

These findings were not replicated by other studies. When two-year stability of recall of DSM-III diagnoses was assessed in a sample of individuals aged six to 23 years, a trend toward better recall of diagnoses was found in boys compared to girls, for all disorders except anxiety. Boys' recall of DSM-III major depression was excellent ($k' = 0.76$) and significantly better than girls' recall ($k' = 0.53$); moreover, boys showed better recall for major depression even when stricter criteria were used (duration of at least four weeks and impairment in a major social role) (Fendrich et al., 1990).

A slight trend for females to report more distant episodes of major depression was found in a highly-educated white collar sample, even though there was no significant gender difference in the distribution of episodes over a 12-month period. In addition, a regression analysis found no significant gender-by-time interaction in the prediction of number of symptom criteria, indicating that males did not report relatively fewer symptoms in the remote episodes as compared to females (Fennig et al., 1994). Similarly, when spouses, relatives and controls of affectively ill probands were examined on two occasions, six years apart, females developed major depression during the six-year period at a rate twice that for males; however, male and female subjects did not differ by the year in which they reported their first episode (Coryell et al., 1992).

A more sophisticated approach based on longitudinal design and corroborative witness reports was implemented by Wilhelm & Parker (1994) in a sample of teacher trainees who were assessed for socio-demographic and historical issues at intake and again five and 10 years later.

Although males and females showed similar overall stability in reporting episodes of depression five years apart (mean 'k' of 0.49 and 0.55 for males and females, respectively), a closer inspection revealed a differential 'passage of time' effect: less immediate or threshold episodes were 'forgotten' especially by males, whereas sub-clinical or threshold episodes were subsequently 'remembered' or reported as more severe especially by females, thus contributing to a female preponderance in rates of depression. In other words, contrasting phenomena (i.e., males more likely to 'forget' over time and females more likely to 'remember' previously unreported episodes) suggested similar overall consistency of recall for the two sexes, allowing for a possible artifact in gender differences of depression to be overlooked. After correction for 'passage of time' effects on stability of prevalence estimates of depression, a clear evidence of any lifetime gender difference was no longer found.

On the other hand, the higher rate of 'forgetting' episodes by males compared to females was not found in other studies. In the sample of young offspring that were assessed for stability of recall by Fendrich et al. (1990), 10 males met caseness criteria on each occasion (and eight on both), whereas 32 females met caseness criteria on first occasion, 17 on the second, and only 16 on both. As a consequence of differential recall between the two sexes, the gender difference in rates of major depression, which was statistically significant at first occasion, was no longer so at the second. Moreover, when consistency in reporting episodes of major depression by relatives of affectively ill probands was examined separately in the two sexes over a 6-year interval, episodes of 'forgetting' and 'remembering' compensated for each other in males, whereas females 'forgot' previously nominated episodes at a higher rate than 'remembering' past episodes previously not nominated (Warshaw et al., 1991).

In summary, the evidence collected so far strongly suggests instability of recall of depressive episodes over time, but does not support the notion that differential recall in the two sexes is entirely responsible for the gender difference in rates of depression.

iii) Course of disorder and mortality rates

Since prevalence is a product of incidence and length of illness, gender differences in the course of the disorder may be responsible for gender differences in rates of depression. For example, Stefánsson et al. (1994) showed that the preponderance of major depression among females was accounted for by recurrent major depression, whereas a female-to-male sex ratio close to 1.0 was obtained for subjects suffering from a single episode of major depression over the 12 months preceding examination. However, the course of disorder loses its relevance in computing lifetime prevalence rates and these were reported to be higher in females compared to males.

In addition, a gender difference in prevalence rates of major depression might be due to mortality rates being different in the two sexes. Since prevalence rates are based on the proportion of survivors experiencing the disorder, gender differences in mortality rates among those with the disorder may have an effect on the observed prevalence rates, with affected individuals of one sex being 'removed' from the older population at a higher rate.

Mortality has been reported to be higher in patients with psychiatric disorders compared to the general population, although less excessive since the introduction of modern psychiatric treatments, improved medical care and shorter duration of admissions (Sims, 1987; Amaddeo et al., 1995). However, the vast majority of data on this issue have been provided by studies carried out among patients contacting psychiatric services or in highly selective nonclinical samples (i.e., students or industrial workers). This is likely to be a bias, since individuals selected in terms of their help-seeking behaviour, prior mental morbidity or other characteristics (e.g., age, sex, occupation) do not constitute a representative sample of the individuals with mental disorders from the communities in which they live. For example, individuals receiving psychiatric care, especially on an inpatient basis, are usually more chronically and severely impaired than other members of the general population suffering from mental disorders but not seeking help and this is expected to have a negative effect on their mortality.

Only a few mortality studies were based on epidemiologic surveys of the general population. Most of these studies found no effect of mental health on mortality risk or, when an effect was found, it was entirely explained by the covariation between measures of mental health and other variables, primarily related to physical health problems (Roberts et al., 1990). On the other hand, three community studies found a relationship between mental health and mortality, even after controlling for possible covariates. In a 12-year follow-up of 610 individuals, psychological distress was predictive of mortality (hazard ratio = 1.7), when preexisting disease, smoking habits and a variety of other risk factors (including a social class and a social network index) were controlled for (Somervell et al., 1989). Moreover, in the Stirling County Study the presence of any type of depressive and anxiety disorder had a significant association with excess mortality over a 16-year interval, but did not significantly interact with sex. However, the increased risk (1.6) was mainly due to depression with or without anxiety and the relationship between depression and death was significantly more pronounced in males (increased risk = 2.1) than in females (increased risk = 1.2) (Murphy et al., 1987). Finally, Livingston Bruce et al. (1994) showed that respondents with recent episodes of major depression at initial interview were about two times more likely to die during a nine-year follow-up period compared to those without recent major depression. The increased risk was confirmed when the effect of other comorbid psychiatric disorders was controlled for. A significant interaction between recent major depression and sex indicated that, in spite of major depression being associated with increased mortality in both sexes, the effect was stronger in males (relative risk = 4.2) compared to females (relative risk = 1.7).

Overall, there appears to be controversial evidence from general population studies about the negative effect of mental disorders on mortality. The disparate findings reported so far may be attributed to study differences in period, place and design, including sample size, length of follow-up, efforts made to reduce attrition and case ascertainment procedures. The two studies showing a relationship between premature mortality and depression suggest that some portion of the preponderance of females in rates of depression is possibly due to differential mortality rates between the two sexes. It is not clear whether higher rates of mortality in males can account entirely for the gender difference in prevalence rates of depression.

iv) Geographical mobility

Gender differences in rates of depression may be influenced by selective migration, with females moving into the catchment areas of university-based research centres. This explanation seems unlikely, since gender differences have been detected in general population studies that sampled individuals living in different countries and areas, including urban, suburban and rural areas.

INCIDENCE OF MAJOR DEPRESSION

Despite the advantage of using incidence rates to study risk factors and to compute age-specific lifetime prevalence estimates on the basis of mortality rates (Kramer et al., 1980), incidence studies of psychiatric disorders are rare since they would require the repeated assessment of a well-defined population over a sufficient number of years. The Lundby Study started in 1947 with a psychiatric examination of all the individuals (N = 2,550) living in a geographically defined area near the town of Lund, in the south of Sweden. The Swedish national registration system allowed for these individuals to be followed-up for 25 years until 1972, in spite of about half moving to mainly urban areas. Diagnostic criteria for depressive illness remained similar throughout the whole study and were based on information elicited by psychiatrists during clinical interviews without reference to well-known diagnostic classifications of mental disorders. Incidence rates increased during the 25-year period in both sexes and were higher in females compared to males: in males, incidence rates were 0.14% during the period 1947-1957 and 0.37% during the period 1957-1972; the corresponding figures for females were 0.41% and 0.77% (Hagnell et al., 1982).

Annual first incidence rates of specific psychiatric disorders were also computed from the data collected at four sites (Baltimore, Durham, St. Louis and Los Angeles) of the NIMH-Epidemiologic Catchment Area Study (Eaton et al., 1989). Rates were based on about 80% of the respondents at baseline interview, that were successfully contacted and interviewed one year later. Although inci-

dence rates are to be considered with caution, due to small numbers of new cases being diagnosed during the follow-up period, females had higher incidence rates of DSM-III major depression in all age groups compared to males and the overall difference was statistically significant (annual incidence rates of 1.10% and 1.98% for males and females, respectively). Gender-specific predicted probabilities showed that for females there was a rise to the peak years of onset in the middle forties with a decline thereafter; for males, the probability of onset was a monotonically decreasing function of age. The relationship of age to incidence of major depression was statistically significant for females, but not for males.

A study based on medical records of a single general practice in London, England, showed that incidence rates of recognized depression increased for males, but hardly changed for females during the 20 years between 1957 and 1976. Despite the different trends in the two sexes, estimates of annual incidence rates revealed that both in 1962 and in 1972 incidence rates were higher in females compared to males (Dunn & Skuse, 1981).

Finally, using 18 years of data from the Camberwell Psychiatric Case Register, Der & Bebbington (1987) provided incidence of all affective disorders in Camberwell, England. Incidence rates of depression were higher in females compared to males, irrespective of the severity of the disorder. For severe depression, incidence rates were 0.29% and 0.52% in males and females, respectively; for moderate depression, the corresponding figures in males and females were, respectively, 0.85% and 1.69% and for depression not otherwise specified 0.30% and 0.49%.

INCIDENCE OF BIPOLAR DISORDER

Incidence of bipolar disorder was estimated using first admission statistics and ranged between 0.9 per 100,000 per year and 6.0 per 100,000 per year (Müller et al., 1968; Eagles & Whalley, 1985; Goodwin & Redfield Jemison, 1990; Sibisi, 1990; Hunt et al., 1993), with admission rates being similar for males and females (Eagles & Whalley, 1985; Goodwin & Redfield Jemison, 1990; Sibisi, 1990). In addition, Goodwin & Redfield Jemison (1990) reported annual incidence rates for people aged 15 years or older seeking treatment and diagnosed as having bipolar disorder. Incidence rates ranged between 11 per 100,000 and 21 per 100,000 per year, with the female-to-male ratio ranging between 0.5 and 3.7.