

EDITORIAL

Epidemiology and depression¹

To examine the epidemiology of depression would appear to be simple. These are two familiar words and their connection might be expected to be readily available from a library search. Inviting though this may seem to the would-be researcher or examination candidate, the area is fraught with hazard. The terms covered by these rubrics are few but cannot be said to have been used accurately or equivalently. In the lexicon of epidemiology are such simple, technical terms as incidence, point prevalence, period prevalence, treated prevalence and lifetime expectancy or morbid risk. For depression there are the descriptive expressions of neurotic and psychotic depression, with the synonyms reactive and endogenous depression, for students of ICD and DSM-II. The term manic-depressive illness also appears to have some accepted meaning. Supplanting or parallel expressions, however, are those of bipolar and unipolar affective disorder which, although of European origin, have been embedded in recent American classifications. The Americans have chosen here to be idiosyncratic and are outside the international psychiatric lexicon. This trend is extended to recent epidemiological work where illnesses are classed as 'major' and 'minor' affective disorders. Finally, there is the particular difficulty of defining a 'case' of depression in community epidemiological surveys. The concept of depression has evolved from experience with severely ill patients in psychiatric hospitals, for the most part, of the western world. Determining appropriate cut-off points for case identification to detect the milder disorders in community surveys is a formidable task.

All of this creates a point of embarkation problem for the examination candidate and the reviewer of the studies on epidemiology and depression. The abiding principle in the search would seem to be the heuristic value of the work that has been done. It seems useful to review the findings to 1968 when Silverman produced her work; then subsequent surveys reported during the seventies using diverse methods to detect depressive symptomatology; and, finally, surveys using standard instruments and fixed diagnostic criteria.

In her small volume Silverman (1968*a*) drew together antecedent work of the epidemiology of depression. She reiterated that depression implies a mood, a symptom, a syndrome and a disease and is best understood at the level of clinical severity. At the time of her publication up to 40 surveys had been undertaken but, in her estimation, lacked comparability and tended to look at all mental illness rather than specific entities. She was, however, able to recalculate prevalence rates for depression. For point prevalence 7 community studies conducted between 1938 and 1960 could be used. The range of findings for depressive disorders was enormous. Ranges could, however, be stated for manic-depressive psychosis: these were from 0.5 to 2 per 1000 of population. For period prevalence 6 community studies conducted between 1933 and 1960 were used. The periods of investigation varied from 6 months to several years and the estimated range was 0.1 to 0.9 for manic-depressive illness alone, but 0.8 to 3 per 1000 for all depressive psychoses. Depressive symptoms, not discrete syndromes, in general population groups were reported to range from 7.2% to 23.6% of the surveyed populations. In the community surveys reviewed, Silverman (1968*b*) estimated that the prevalence rate for depressive psychosis was less than 1 per 1000 population and that of depressive neurosis was 2 or 3 times higher. She concluded that knowledge and theory about depression had not advanced much beyond the level of clinical description. She felt that the epidemiology of depressive states had been only slightly explored and advocated that it be studied separately from that of mental disorders as a group.

Since Silverman's publications many epidemiological studies specifically concerned with the affective disorders have appeared. For the most part, however, these have lacked comparability. Only a

¹ Address for correspondence: Dr M. R. Eastwood, Clarke Institute of Psychiatry, Epidemiology and Evaluation Section, 250 College Street, Toronto, Ontario, Canada M5T 1R8.

small number of these studies need be reviewed here to illustrate the problems. A number of typical epidemiological studies of depression in the general population reported during the 1970s are presented.

Blumenthal & Dielman (1975) used the Zung Self-rated Depression Scale (SDS) (Zung, 1965) on 160 non-random couples in 9 locations in 6 US states. They report that 13 % of the sample was over a cut-off of 56 on the Zung Scale. When a more stringent index was used, 5 % of the men and 14 % of the women reported depressed mood. Väisänen (1975) used a semi-structured psychiatric interview for case-finding in 1000 adults in urban and rural Finland. Diagnoses were established on the basis of impressions gained during the interviews, using criteria to which he was accustomed in his hospital practice. The prevalence for depressive neurosis was 7.6 %. Brunetti (1975) used a psychiatric interview on a small sample of 102 adults in rural France and reported a 1-year period prevalence of 5.9 % for neurotic depression. An indirect survey of the same area was conducted 10 years later. The entire population of 683 was surveyed, through local general practitioners and hospital records, and the rate of neurotic depression was 0.58 %. The higher rate of the first survey was explained by an adult-only sample and differences in methodology. Hällström (1970) conducted a population study among women in Gothenburg, Sweden: of the 884 females undergoing a comprehensive examination, 8 % were classified as suffering from depression. Nandi *et al.* (1979) used a psychiatric interview for case identification in a survey in rural India. Diagnostic criteria were formulated and each probable case was examined by 2 psychiatrists separately before being included as a case. They report a point prevalence for depression of 3.7 %, 2.5 % for males and 5.1 % for females.

These studies clearly have methodological deficiencies. The rates vary considerably, the periods of observation are not specified, and diagnostic criteria are not consistent.

In an attempt to overcome these difficulties, studies were undertaken separately during the 1970s, on 4 continents, to survey disorders reliably with the intention of gaining a refined knowledge of the epidemiology of depression. Each investigator in these studies used a standard approach with close attention being given to the definition of 'caseness' (see Table 1).

In 1970, and later in 1974, Brown *et al.* (1977) conducted a survey of women from Camberwell in

Table 1. Estimates of the prevalence of depression using standardized instruments

Reference	Place	Period	Method	Type disorder	Sex	Rate (%)
Brown <i>et al.</i> (1977)	Camberwell, 1970, 1974; North Uist, 1975	3 months	PSE + clinical judgement	Depression	F	13.3 cases + 12.2 borderline cases
		1 year		Depression	F	14.8 cases + 13.3 borderline cases
Wing (1976)	London	?	PSE-ID-CATEGO	Depression	F	5.8 cases + 5.2 borderline cases
						8.4 cases + 7.1 borderline cases
Wing <i>et al.</i> (1978)	London	?	PSE-ID-CATEGO	Depression (mostly 300.4)	F	9.28 definite + borderline cases
Henderson <i>et al.</i> (1979)	Canberra, 1977	?	GHQ + PSE-ID-CATEGO	Depression (296.2 + 300.4)	M	2.6 definite + threshold cases
					F	6.7 definite + threshold cases
Orley <i>et al.</i> (1979)	Uganda, 1972	?	PSE-ID-CATEGO	Depression	M	14.3 definite + threshold cases
					F	22.6 definite + threshold cases
Weissman & Myers (1978)	New Haven, 1975-6	Point	SADS-L + RDC	Major depression	M	3.2 definite + probable cases
					F	5.2 definite + probable cases
					M	2.3 definite + probable cases
		Lifetime		Major depression	F	2.7 definite + probable cases
					M	12.3 definite + probable cases
					F	25.8 definite + probable cases
Minor depression	M	5.9 definite + probable cases				
	F	11.7 definite + probable cases				

south-east London. The total sample size was 458. The shortened form of the Present State Examination (PSE) (Wing *et al.* 1974), a semi-structured interview, was used. In 1975 a similar survey was conducted on 154 women in North Uist in the Outer Hebrides. The likelihood of being a 'case' was determined from the interview plus a consensus of clinical judgement. Women were classified as cases if they were considered to have experienced a definite psychiatric disorder in the 3 months preceding the interview and as borderline cases if they had definite symptoms which were less frequent or severe than those of the 'cases'. The 1-year period prevalence for definite 'case depression' was 8.4 % for Uist and 14.8 % for Camberwell, the respective 3-month figures were 5.8 % and 13.3 %. Within the preceding year 15.6 % of the Uist women and 28.2 % of the Camberwell women were considered to have suffered a depression or borderline depression.

Wing (1976) set out to overcome past difficulties with comparability of studies. Knowledge of psychiatric disorders gained on in-patients and out-patients was used to investigate significant psychiatric morbidity in individuals in the general population. The 54 obligatory questions in the 9th edition of the PSE were used to this end, and rules were laid down for an 8-point 'index of definition' (ID) of psychiatric disorder, ranging from no problems to definite cases. Using this index, a threshold level was determined above which it was possible to recognize and classify psychiatric disorders using the CATEGO program of clinical classification (Wing *et al.* 1974). Applying this technique, the distribution of disorders among a random sample of 237 females from a London suburb was compared with that for groups of in-patients and out-patients. Wing's general population sample was part of a sample previously surveyed by Brown's Bedford group. Including borderline and definite disorders, a figure of 8 % for depression in the community (2.11 % for ICD 296.2 and 5.91 % for ICD 300.4) *can be determined* from his data. It should be noted that 80 % of the cases of psychiatric disorder were at the level of borderline disorder, but *the comparable figure for depression is not stated*. In this particular study Wing was more concerned with the technique than in comparing the frequencies of the different disorders in the samples. Wing *et al.* (1978) elaborated on the earlier study by making comparisons between methods of case identification. Although the same sample of 237 females is used, the rate varies slightly, with the PSE detecting 22 depressive disorders (a prevalence rate of 9.28 %) as compared with 19 depressive disorders in the 1976 paper. This discrepancy is not explained. When the US Feighner criteria (Feighner *et al.* 1972) were applied to these cases, only 1 definite and 2 probable 'cases' were accepted, giving a rate of 0.42 % definite and 1.26 % for definite and probable 'cases' combined.

Henderson *et al.* (1979) carried out a comparable study in Canberra, Australia, using a sample obtained from the electoral roll. Employing a two-stage procedure, the General Health Questionnaire (GHQ) (Goldberg, 1972) and the PSE, the population was screened for mental disorder. First, scores of more than 2 on the 12-item GHQ showed 30.5 % to be disturbed and 22.4 % severely so. The PSE was administered to a stratified sample of phase one according to the probability of each respondent being a 'case'. Using the PSE, 1.8 % of the females and 1.1 % of the males were definite 'cases' of psychiatric disorder and 9.2 % of the females and 5.9 % of the males were threshold 'cases'. In terms of CATEGO analysis 2.6 % of males and 6.7 % of females suffered from depression (including both ICD 296.2 and ICD 300.4). No mention is made of the definite/threshold breakdown for these depressions.

Orley *et al.* (1979) conducted a survey in 2 rural Ugandan villages using the PSE-ID-CATEGO procedure. In all, 206 adults were interviewed and 4.9 % were found to be definite cases of psychiatric disorder and 20.4 % threshold 'cases'. Combining all cases, both definite and threshold, 27 % of the women and 24 % of the men had psychiatric disorders. The CATEGO program classified 14.3 % of the men and 22.6 % of the women as depressed (again, there is no separation of the definite and threshold cases), or approximately 18 % of the total sample. When the Feighner criteria for depressive disorders was applied to these PSE data, there were 6.3 % definite and 4.9 % probable cases.

In the US Weissman & Myers (1978) presented data on patients examined in the community of New Haven, Connecticut, which had a population of 72000. From a systematic sample, 86 % were interviewed in 1967, of which 77 % were re-interviewed in 1969. In 1975-6 there were further interviews of 46 % of the original population sample. This final sample comprised 515 subjects. Although

there were some racial and class differences, the samples of each interview were generally the same, as was the symptom status. However, due to the attrition of the original sample, there is the possibility that severe disorders may be underrated in the last survey. For the 1975–6 re-interviews the lifetime version of the Schedule for Affective Disorders and Schizophrenia (SADS-L) (Endicott & Spitzer, 1978) was used and the Research Diagnostic Criteria (RDC) (Spitzer *et al.* 1978) were applied. The current point prevalence for depression was 6.8 % (5.7 % definite), with 4.3 % being major (3.7 % definite) and 2.5 % minor depression (2.0 % definite). A major depression is an illness of dysphoric mood of neurotic or psychotic proportion lasting at least one week, accompanied by the depressive syndrome. A minor depression is a non-psychotic episode of illness in which a sustained mood of depression predominates without the full depressive syndrome that characterizes major depression. For lifetime rates, 20 % had major depression (18 % definite), 9.2 % minor depression (8.6 % definite), giving a total lifetime rate for depression of 26.7 % (24.7 % definite). The majority of major depressions (86 %) were primary, that is a first episode not preceded by another psychiatric condition. There were no current cases of mania or hypomania. Using the SADS and RDC, 7.9 % of the women were classified as depressed (5.2 % with major depression and 2.7 % with minor depression) and 5.5 % of the men (3.2 % with major and 2.3 % minor depression). These figures were not broken down into definite and probable cases.

How may these recent studies be drawn together? Only the female data will be reviewed because some of the investigations were limited to female samples. Brown *et al.*, using the PSE plus clinical judgement to determine 'caseness', reported very high rates. Thus, 28.2 % of the Camberwell women were classified as either cases or borderline cases of depression within a 1-year period and 25.5 % within 3 months. Wing's sample of 237 women was part of Brown's original Camberwell sample of 458. They were re-interviewed, using the PSE, from 5 to 84 days after the Bedford team. Wing's method of case identification (PSE-ID-CATEGO) gave a much lower rate of depression and his criteria, plus the time lag before reassessment when symptoms could change, may account for this. Henderson *et al.* and Orley *et al.* also used the PSE-ID-CATEGO procedure in their surveys. The rates for the women of Canberra and Camberwell were similar: 6.7 % and 8–9 % respectively. Orley's figures for Ugandan women were more than twice as high at 22.6 %. One reason for this greater morbidity might be that depressive disorders in rural Africa remain untreated. The Weissman & Myers figures, derived using the SADS-L and RDC on a US urban population, appear similar to the Wing and the Henderson data.

It might appear from these studies that the prevalence of depression has increased enormously, with the reported figures being anywhere from 25 to 50 times greater than those suggested by Silverman. However, the increase may reflect better methodology rather than any growth in actual prevalence. The recent investigators have made important contributions to the epidemiology of depression in the area of improved methodology. Their standard approach to case-finding makes their results more communicable and repeatable and their more stringent criteria have allowed separation of the momentary presence of depressive symptoms from the longer lasting depressive disorders. However, collectively, there are several methodologies involved in these studies, some essential data are unclear or missing and comparisons are difficult on close inspection.

The different diagnostic criteria used in the studies make comparison problematical. The criteria for determining a case, using standardized instruments, are usually based upon lists of symptoms occurring for variable periods of time. Questionnaire scores may be a function of totalling the number of symptoms, the duration of symptoms, some formula relating symptoms and duration or some weighting system whereby empirically a psychotic symptom would be assumed to be more important than a neurotic one. Due emphasis, therefore, has been placed upon the number and type of symptoms, although Craig & Van Natta (1976) report that duration of symptoms may be a more important factor than the sheer number of symptoms in assessing depression. Symptoms and duration appear to vary between the standardized instruments. The Present State Examination allots a weight of 1 for a symptom being present in moderate degree during the previous month or, if severe, present less than 50 % of the time, and a weight of 2 if present in severe form for more than 50 % of the time. It is hard to determine how duration is incorporated into the CATEGO classification. The major-

minor classification of the American Research Diagnostic Criteria requires at least a 1-week duration of symptoms (definite if more than 2 weeks, probable if 1–2 weeks). In both the PSE and SADS the question regarding depression is stated in the past tense, which could mean that the patient has been ill some time during the criteria period, sufficiently so for inclusion as a case, but not necessarily on the day of examination. It is a moot point whether the rates quoted are point or period prevalence figures. Superficially, the PSE data of the Wing and Henderson studies and the SADS data of Weissman resemble one another. However, when the Feighner criteria, identical to the RDC for symptoms with a minimal duration of 1 month, were applied by the English workers in London and Uganda to their PSE data, the rates dropped dramatically. Since Wing states that his cases were of a long duration, then the difference between the 2 methods must lie in the management of the list of symptoms and not their duration. A further difference is that the relative proportions of the major/minor depressions and psychotic/neurotic depressions are inverse in the Wing and Weissman studies. Thus, while superficially the English and American studies appear to validate one another consensually, this is not borne out on close inspection.

A further problem of comparison is the certainty with which a case is decided upon. It would appear that the certainty of diagnosis is much less using the PSE than with SADS, possibly because of a more tentative approach by the English workers. The majority of the cases defined by the SADS were definite, whereas those found by the PSE were largely borderline or threshold cases. For the diagnosis of depression the accuracy figures for the PSE data and by sex in the SADS data are not stated. Since certainty and severity are related, the number of severe cases of depression remains unclear. The studies described suggest that, according to some definition, 5–10 % of the population has some kind of depressive syndrome. The utility of this finding is questionable until it is known at what level of severity treatment is warranted and indeed likely to be available.

Another important area which is inconsistent between the studies is the type of prevalence rate reported. The rates will vary according to the survey period examined. Since period prevalence is a function of point prevalence plus incidence, then the longer the survey period the greater the prevalence rate. So 1-day point prevalence rates and 1-month period prevalence rates would be less than 1-year prevalence rates. The Brown, Wing, Henderson and Orley studies use the PSE to gather their data. This is intended to cover symptoms experienced during the 4 weeks prior to the interview. Brown *et al.*, however, report 3-months and 1-year period prevalence figures (it is assumed that they altered the PSE accordingly); and the other investigators do not identify clearly their rates as to whether they consider them to be point prevalence or 1-month period prevalence figures. Weissman reports 2 rates: current point prevalence and lifetime prevalence. She readily admits to the limitations of the latter rate in her study, as only the surviving members of the population were included in the cross-sectional survey. Another issue which goes hand in hand with the period of observation is the time at which the surveys were actually conducted. The term point prevalence is considered often preferable to period prevalence when generally used in epidemiology. The concept of period prevalence has been criticized as a meaningless mixture of 2 incommensurable concepts: point prevalence and incidence. However, it is believed that the term 1-year period prevalence has a unique contribution in the epidemiology of affective disorders which have been shown to be the only diagnoses, apart from alcoholism, to exhibit seasonal variations (Eastwood & Stiasny, 1978). As such, point prevalence figures, and perhaps incidence figures too, would vary at different times of the year in their presentation. On the other hand, unless the condition were increasing, being given more attention or waning, the 1-year period prevalence would not change much from year to year. If the point prevalence, or 1-month period prevalence, were taken in the spring or autumn for affective disorders, the results would be higher than if taken in the summer or winter. The increased rates in spring and autumn would be a mixture of new episodes in old cases and new episodes (incidence) in previously well people. As yet, the proportions of each are unknown. None of the studies outlined clarifies when the samples were surveyed.

Bearing all these differences between the studies in mind, it is difficult to say that true prevalence rates for depression have been established. Furthermore, it is not known how many cases should

be seen by either general practitioners or psychiatrists for treatment. Although the studies reviewed have made considerable improvements in methodology, future investigations must take into consideration the period of observation, the seasonal variation of affective disorders, duration of symptoms and certainty of diagnosis in their surveys.

M. R. EASTWOOD AND P. M. KRAMER

REFERENCES

- Blumenthal, M. D. & Dielman, T. E. (1975). Depressive symptomatology and role function in a general population. *Archives of General Psychiatry* **32**, 985–991.
- Brown, G. W., Davidson, S., Harris, T., MacLean, U., Pollock, S. & Prudo, R. (1977). Psychiatric disorder in London and North Uist. *Social Science and Medicine* **11**, 367–377.
- Brunetti, P. M. (1975). Rural Vacluse: two surveys on the prevalence of mental disorders. Summary of data. *Acta psychiatrica scandinavica*, Supplement 263, 12–15.
- Craig, T. J. & Van Natta, P. A. (1976). Presence and persistence of depressive symptoms in patient and community populations. *American Journal of Psychiatry* **133**, 1426–1429.
- Eastwood, M. R. & Stiasny, S. (1978). Psychiatric disorder, hospital admission, and season. *Archives of General Psychiatry* **35**, 769–771.
- Endicott, J. & Spitzer, R. L. (1978). A diagnostic interview. The Schedule for Affective Disorders and Schizophrenia. *Archives of General Psychiatry* **35**, 837–844.
- Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A., Winokur, G. & Munoz, R. (1972). Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry* **26**, 57–63.
- Goldberg, D. P. (1972). *The Detection of Psychiatric Illness by Questionnaire*. Oxford University Press: London.
- Hällström, T. (1970). Depressions among women in Gothenburg. *Acta psychiatrica scandinavica*, Supplement 217, 25–26.
- Henderson, S., Duncan-Jones, P., Byrne, D. G., Scott, R. & Adcock, S. (1979). Psychiatric disorder in Canberra. A standardized study of prevalence. *Acta psychiatrica scandinavica* **60**, 355–374.
- Nandi, D. N., Banerjee, G., Boral, G. C., Ganguli, H., Ajmany (Sachdev), S., Ghosh, A. & Sarkar, S. (1979). Socio-economic status and prevalence of mental disorders in certain rural communities in India. *Acta psychiatrica scandinavica* **59**, 276–293.
- Orley, J., Blitt, D. M. & Wing, J. K. (1979). Psychiatric disorders in two African villages. *Archives of General Psychiatry* **36**, 513–520.
- Silverman, C. (1968a). *The Epidemiology of Depression*. The Johns Hopkins Press: Baltimore.
- Silverman, C. (1968b). The epidemiology of depression – a review. *American Journal of Psychiatry* **124**, 43–51.
- Spitzer, R. L., Endicott, J. & Robins, E. (1978). Research diagnostic criteria. Rationale and reliability. *Archives of General Psychiatry* **35**, 773–782.
- Väisänen, E. (1975). Psychiatric disorders in Finland. *Acta psychiatrica scandinavica*, Supplement 263, 22–33.
- Weissman, M. M. & Myers, J. K. (1978). Affective disorders in a US urban community. *Archives of General Psychiatry* **35**, 1304–1311.
- Wing, J. K. (1976). A technique for studying psychiatric morbidity in in-patient and out-patient series and in general population samples. *Psychological Medicine* **6**, 665–671.
- Wing, J. K., Cooper, J. E. & Sartorius, N. (1974). *The Measurement and Classification of Psychiatric Symptoms*. Cambridge University Press: Cambridge.
- Wing, J. K., Mann, S. A., Leff, J. P. & Nixon, J. M. (1978). The concept of a 'case' in psychiatric population surveys. *Psychological Medicine* **8**, 203–217.
- Zung, W. W. K. (1965). A Self-rated Depression Scale. *Archives of General Psychiatry* **12**, 63–70.