Psychological Medicine, 2001, **31**, 1347–1359. © 2001 Cambridge University Press
DOI: 10.1017/S0033291701004743 Printed in the United Kingdom

A 25-year longitudinal, comparison study of the outcome of depression

H. BRODATY,[1] G. LUSCOMBE, C. PEISAH, K. ANSTEY AND G. ANDREWS

From the School of Psychiatry, University of New South Wales, Academic Department for Old Age Psychiatry, Prince of Wales Hospital and Clinical Research Unit for Anxiety Disorders, St Vincent's Hospital, Sydney, New South Wales, Australia

ABSTRACT

Background. There is still a relative paucity of information about the long-term course of depression.

Methods. Consecutive patients admitted to a teaching hospital psychiatry unit with symptoms of depression, previously assessed at 6 months and 2, 5 and 15 years after index admission, were reviewed at 25 years ($N = 49$, including eight informants of deceased probands, of an original 145 with major depression (DEPs)). Prospective psychiatric ($N = 22$) and retrospective surgical ($N = 50$) control groups assessed after 25 years were used for comparison.

Results. A further decade of follow-up confirmed the chronicity of depression. Of depressed patients (DEPs) followed for the full 25-year-period only 12% of the 49 original DEPs recovered and remained continuously well, 84% experienced recurrences, 2% experienced an unremitting course and another 2% died by suicide. Note that in the first 15-year-period 6% (9/145 DEPs) committed suicide, a further 38 died and 32 were lost to follow-up. They experienced an average of three episodes of depression over the 25 years. In the decade since the 15-year follow-up, 27% improved in clinical outcome (including four of five previously chronically depressed patients), 55% remained unchanged and 18% worsened; and the number of episodes per year declined. Patients initially diagnosed with neurotic or endogenous depression had similar long-term outcomes. The criteria for a current DSM-III-R disorder were met by 37% of DEPs, including 11% with depression or dysthymia. On the global assessment of functioning scale 78% of the DEPs had some impairment compared to 62% of psychiatric controls and 40% of surgical controls.

Conclusion. Even after 25 years, severe depressive disorders appear to have poor long-term outcomes. Patients with chronic outcomes over 15 years can improve when followed over longer periods.

INTRODUCTION

There has been a 'paradigmatic shift', with unipolar major depressive disorders now viewed as chronic illnesses, with episode recurrence the norm (Judd, 1997; Andrews, 2001). Long-term studies, beyond 10 years, of unipolar affective disorders indicate rates of repeat episodes of ~75% (Stephens & McHugh, 1991; Piccinelli & Wilkinson, 1994; O'Leary & Lee, 1996; Mueller *et al.* 1999), readmission of 35 to 62% (Lee & Murray, 1988; Smith & North, 1988; Stephens & McHugh, 1991; Thornicroft & Sartorius, 1993) and chronicity or persistence of 5 to 25% (Winokur & Morrison, 1973; Angst, 1986, 1997; Thornicroft & Sartorius, 1993; Judd, 1997; Judd *et al.* 1998). All subjects in these studies were in-patients of psychiatric units attached to universities, excepting two studies in which 23% (Judd *et al.* 1998; Mueller *et al.* 1999) and 37% (Thornicroft & Sartorius, 1993) were out-

[1] Address for correspondence: Professor Henry Brodaty, Academic Department for Old Age Psychiatry, Euroa Centre, Prince of Wales Hospital, Avoca Street, Randwick, Sydney, NSW 2031, Australia.

patients. The weighted average proportion of patients from psychiatric settings recovering and remaining continuously well at least 10 years following index depressive disorder has been reported to be 24%, with a range of 18 to 30% (Piccinelli & Wilkinson, 1994). Two more recent studies have reported slightly higher rates of 34 and 41% (Thornicroft & Sartorius, 1993; Surtees & Barkley, 1994).

For obvious logistical reasons, studies of depression outcome extending beyond 25 years are rare. The longest reported periods of follow-up are the 40-year Iowa 500 study (Winokur & Tsuang, 1996) the 39-year University of Oslo study (Opjordsmoen, 1989) the 37-year Henry Phipps Psychiatric Clinic study (Stephens & McHugh, 1991) and the 25-year Zurich study (Angst, 1986). Long-term follow-up studies such as these are critical to understanding the natural history of diseases and the implications for continuation of treatment.

Depression has a profound impact on social function (Hall & Wise; 1995; Judd *et al*. 1996; Hirschfeld *et al*. 1998) and is ranked as the leading cause of disability in the world (Murray & Lopez, 1996). In the 16-year 'Stirling County' longitudinal community study of depression, low socio-economic status (SES) was linked to depression – incident cases over the follow-up being higher in the low SES group; also there was a trend for depression to be associated with downward changes in SES (Murphy *et al*. 1991). Follow-up studies of depressed persons extending 10 years and beyond report rates of severe social impairment both longitudinally and cross-sectionally. Over 10 years, Thornicroft & Sartorius (1993) found that 67% of patients experienced impaired social functioning for more than 5% of the follow-up period. At the 40-year follow-up of the Iowa 500 study Tsuang *et al*. (1979) found current rates of 19% with fair to poor marital status, 30% fair to poor residential status and 33% fair to poor occupational status. The 12-year follow-up of the Edinburgh series revealed 25% of patients had global assessment of functioning scores indicating moderate to severe impairment (Surtees & Barkley, 1994).

While the consensus opinion is that long-term outcome of affective disorders is bleak, minor inconsistencies in studies relate to different methods of determining outcome of depressive disorders pertaining to sampling, length of follow-up, definition of a depressive disorder (not just in terms of categorical *versus* dimensional definitions, but even in the use of diagnostic criteria), and the significance of subtyping e.g. melancholic *versus* non-melancholic. Other discrepancies between studies are attributable to the use of cross-sectional proportional outcomes *versus* cumulative risk over time. Control groups have rarely been used and only one has included non-affective psychotic and surgical controls (Winokur & Tsuang, 1996).

The sample in our study is an Australian cohort of consecutive in-patients admitted to a teaching hospital psychiatry ward during 1966 to 1970 with symptoms of depression. They were reassessed at 6 months and 2, 5 and 15 years (Kiloh *et al*. 1972*a*, 1988; Andrews *et al*. 1990). Our baseline sample comprised the 133 patients with adequate data at 15 years of the 145 patients originally diagnosed with either endogenous or neurotic depression (Kiloh *et al*. 1988). About one in five of these patients had recovered from the index episode and remained continuously well, whereas 12% remained incapacitated and 7% had committed suicide. Lee & Murray's (1988) 18-year assessment found a similar pattern of clinical outcome in that one in eight of their group recovered and remained continuously well and 28% (25/88) remained incapacitated or committed suicide.

The aims of this study were: (*i*) to provide further data on the long-term outcomes of depression over 25 years in terms of clinical outcome (recurrence and chronicity of depression), suicide, proportion of follow-up spent depressed, number of episodes of depression, rehospitalizations and other treatments for depression and psychosocial functioning; (*ii*) to provide cross-sectional outcome in terms of current depression, other psychiatric morbidity and psychosocial functioning; (*iii*) to compare outcomes with those of patients with other psychiatric diagnoses presenting with depressive symptoms and with surgical controls; (*iv*) to compare outcome at the 15 year follow-up with outcome at the 25 year follow-up, with particular interest in whether patients with a chronic course over 15 years would remain depressed over the next 10 years; and (*v*) to explore outcome by diagnostic type to see if the outcome of those

with an original diagnosis of endogenous depression differed from those with an original diagnosis of neurotic depression.

Our hypotheses were: (*i*) that we would confirm the significant long-term incapacity that results from depression; (*ii*) that the global clinical outcome at 25 years would be similar to that found at the 15 year follow-up or worse; and (*iii*) that there would be a greater rate of readmission in the endogenous group than in the neurotic group, i.e. replicating the finding from the 15 year follow-up.

METHOD

Sample at index

Patients were 212 consecutive in-patients suffering either from a depressive illness (DEP) (145/212 or 68·4%), or a non-psychotic psychiatric illness presenting with depressive symptoms (PSYCON) (67/212 or 31·6%) admitted to Prince Henry Hospital, Sydney, Australia between April 1966 and April 1970. Patients with schizophrenia/paraphrenia or an organic brain syndrome were excluded. Details regarding the identification, interview, diagnosis and follow-up of this cohort have been reported (Kiloh *et al.* 1972*a*, *b*, 1988; Andrews *et al.* 1973, 1990).

Consensus ICD-8 diagnoses (WHO, 1967) had been reached by three psychiatrists, based on information from a structured psychiatric interview (Kiloh *et al.* 1972*b*). Of the 212 (143 females and 69 males), 69 (32·6%) were diagnosed, consensually using established criteria, with psychotic or endogenous depression (ED), 76 (35·9%) with neurotic depression (ND) and 67 (31·6%) with other non-psychotic disorders (PSYCON) presenting with depressive symptomatology (50 patients with various anxiety and somatoform disorders, 11 with substance abuse, and six with other disorders). All patients were quite ill having been referred to and requiring admission in a university teaching hospital. Those with primary depression would meet criteria for major depressive disorder in current terminology. The depressive diagnoses and their subtypes appear to be reliable as they remained stable at the 15-year follow-up (Andrews *et al.* 1990) and they proved consistent with DSM-III-R diagnoses generated by recoding original structured interview records (Andrews *et al.* 1999). Particular attention was paid to subtyping depression as the study originated as part of the debate about the unitary *versus* the endogenous–neurotic dichotomy formulation of depression (Kiloh & Garside, 1963). Baseline ICD-8 diagnoses can be translated into current ICD-10 diagnoses. Endogenous depression was classified in ICD-8 as 296.2 manic depressive psychosis, depressed type which translates into ICD-10 as F32 depressive episode (single episode) and F33 recurrent depressive disorder; or 296.0 involutional melancholia, which in ICD-10 is severe depressive episode with (F32.2) or without (F32.3) psychotic symptoms and F33.2 or 3 respectively if it is recurrent. Note that depressive psychosis in the ICD-8 did not mean psychotic as we now understand it but rather endogenous or melancholic depression. Neurotic depression was classed in ICD-8 as 300.4 depressive neurosis in ICD-8, which translates in ICD-10 to the following; F34.1 dysthymia; F32.0 mild depressive disorder, current episode mild; F41.2 mixed anxiety and depressive disorder; and F43.21 adjustment reaction, prolonged depressive reaction.

Over the 25 years, bipolar affective disorder occurred in none of the depressive group, one person in the psychiatric control and one in the surgical control groups.

DEP and PSYCON sample at 25 years

By the 15-year follow-up 61 (28·8%) of the 212 probands were deceased, 29 (13·7%) refused reassessment and 21 (9·9%) were not located (see Table 1 for attrition rates). Of the 61 deceased probands, 50 (82·0%) died naturally (29 EDs, 9 NDs and 12 PSYCONs) and 11 (18·0%) committed suicide (4 EDs, 5 NDs and 2 PSYCONs). This left 101 probands from the 15-year follow-up available for assessment at 25 years. A further proband, who had previously refused involvement, rejoined the study. Of the 102 probands potentially available for follow-up, 61 (59·8%) were interviewed, two (2·0%) had committed suicide in the last decade (1 ED and 1 PSYCON) 17 (16·7%) had died naturally (8 EDs, 5 NDs and 4 PSYCON), 13 (12·8%) could not be located and nine (8·8%) refused to participate. Informants were interviewed for ten of the 19

Table 1. *Attrition rates*

	ED	ND	PSYCON	Total
Index	69	76	67	212
At 15 years				
Refused/not located	−9	−23	−18	−50
Deceased	−33	−14	−14	−61
Sample remaining > 15 years	27	39	35	101
At 25 years				
Refused/not located	−5	−7	−10	−22
Deceased, no oucome	−4	−2	−3	−9
Previously refused and rejoined	1	0	0	+1
Sample at 25 years*	19	30	22	71

ED, endogenous depression; ND, neurotic depression; PSYCON, other neurosis controls.
* Includes deceased outcome: 5 EDs, 3 NDs, 2 PSYCONs.

recently deceased probands. Detailed outcome data were available for a total of 71 probands (49 DEPs and 22 PSYCONs).

Surgical control sample

A retrospective pairwise control sample of 50 patients was recruited from people who were admitted consecutively to Prince Henry Hospital for routine surgery (appendectomy and cholecystectomy) between 1966 and 1972 and had no history of psychiatric illness prior to surgical admission (based on current assessment). Attempts were made to match for age, sex and year of admission. There were difficulties matching for very old probands and in tracing and recruiting control patients. Whereas the probands had had at least four contacts over the past 25 years, the surgical controls had not had any follow-up since their admission and contact with them took longer to establish. All probands or their informants and surgical controls provided written informed consent.

Follow-up procedures

All research staff were blind to the original psychiatric diagnoses but it proved impossible to maintain blindness as to whether patients' index admission was psychiatric or surgical. Researchers remained blind to the course of the illness. Assessment was undertaken by psychiatrists or research pscyhologists who had been trained in the follow-up procedures and had completed the WHO CIDI training course (see below). In 1994 and 1995, approximately 25 years after their index admission, we attempted to determine whether or not patients had died through State Death Registers and hospital records, and then to locate them by use of their previous address, current telephone directories, current electoral rolls or informants. (Voting in Australia is compulsory and electoral rolls are comprehensive.)

Not withstanding willingness to participate, the advanced age of many of the probands, the use of informants and the inherent difficulties of studying an ill (or previously ill) population meant that there was never a complete dataset for all information.

Outcome measures

The Kiloh *et al.* (1988) 'clinical criteria' for outcome of depression comprised: 'recovered and continuously well', 'recovered in part or in whole but had subsequent attacks', and 'remained incapacitated or suicided'. Subsequent attack was defined as depressed mood with accompanying symptoms sufficient to interfere with normal function for at least 2 weeks. Based on all information obtained, independent ratings were made by two or three assessors of whom at least one was a senior psychiatrist (H. B.) and consensus was reached at case conference.

'Timeline charts', following the methodology employed in the 15-year follow-up (Kiloh *et al.* 1988), were constructed to chart the 10-year period from the previous follow-up as follows.

Suicide, both attempted and completed

Patients were asked about suicide attempts during the 16–25 year period, and death certificates were requested for all patients known to have died during this period. Because of the known under-reporting of suicide on death certificates, informants were questioned when an unnatural cause of death was recorded. We previously reported that of the 108 DEPs with information on mortality at 25 years 10 had suicided (Brodaty *et al.* 1997).

Length of episodes of non-bereavement depression

Patients were specifically interviewed regarding length of individual episodes, which were defined

https://doi.org/10.1017/S0033291701004743 Published online by Cambridge University Press

as interfering with function. Because of inherent problems with precision of recall over a 10 year period, these data were recorded as percentage categories of the time spent depressed since last follow-up.

Number of episodes of depression persisting at least 2 weeks (*excluding bereavement reactions*)

The number of episodes of depression persisting for at least 2 weeks was determined by research psychologists using a semi-structured interview.

Treatments for depression

Patients and informants were specifically questioned regarding current and interval pharmaceutical therapies, other psychiatric and psychological therapies, ECT and hospitalizations (where possible hospitalizations were verified by requesting discharge summaries). Medications were categorized according to an Australian physician drug manual (MIMS Australia, 1995). It was not possible to ascertain drug doses reliably.

Instruments

Index and current ratings of depression severity were measured using 21-item Hamilton Rating Scale (HAM-D; Hamilton, 1960) and the 21-item Beck Depression Inventory (BDI; Beck et al. 1961).

The Composite International Diagnostic Interview (CIDI version 1.1; World Health Organization, 1993), used to determine current (defined as present within the month prior to interview) psychiatric morbidity using DSM-III-R (American Psychiatric Association, 1987) classification, was administered by psychologists who had completed a WHO regional centre training course. The CIDI has been designed to be administered by either a lay person or to be self-complete, and to allow the use of a proxy. Thus, where possible, informants of deceased subjects or those with dementia were interviewed. The CIDI sections covered somatoform and dissociative disorders, phobic and other anxiety disorders, depressive disorders and dysthymic disorders, manic and bipolar affective disorders, disorders resulting from use of alcohol, obsessive–compulsive disorders and disorders resulting from the use of psychoactive substances. The CIDI section rating nicotine dependence was not assessed. Lifetime diagnoses are not reported because we found them to be unreliable (Andrews *et al.* 1999).

The Global Assessment of Functioning Instrument (GAF; American Psychiatric Association, 1987) from the DSM-III-R over the 'past year' was rated for patients based on 'function' descriptors only (e.g. no more than slight impairment in social, occupational, or school functioning), ignoring 'symptom' descriptors. Each of two or three raters independently judged a GAF score range based on a case presentation and review of all data (no impairment as GAF of 81–90, mild 61–80, moderate 41–60, severe $\leqslant 40$), and a consensually determined rating was recorded ($\kappa = 0{\cdot}85$). Ratings were also made for deceased patients by asking informants to rate the patient's highest level of functioning for at least a few months during the year prior to death (avoiding the pre-terminal period).

The 'psychosocial functioning' subsection from the Longitudinal Interval Follow-up Evaluation assessment (LIFE; Keller *et al.* 1987) was employed to assess current social (past month) and occupational (past week) functioning and satisfaction (past week), each area being rated on a 6- to 8-point ordinal scales. Both patients and raters also made global ratings of social adjustment.

Dementia was diagnosed at a case conference, including two psychogeriatricians (H. B., C. P.) and two psychologists. Cases were determined based on best available information. The consensus diagnosis of dementia was considered definite if confirmed by either a research psychogeriatrician documenting that the patient met DSM-III-R criteria for dementia or documentation of a diagnosis of dementia from at least two independent doctors. Otherwise patients had to have evidence from at least four of the following: Mini-Mental State Examination (Folstein *et al.* 1975) score $\leqslant 24$, Clinical Dementia Rating Scale (Hughes *et al.* 1982) score $\geqslant 1$, Informant Questionnaire on Cognitive Decline in the Elderly (Jorm & Jacomb, 1989) score $> 3{\cdot}6$, Global Assessment of Functioning (American Psychiatric Association, 1987) score $\leqslant 60$, Activities of Daily Living (Katz & Apkom, 1976) score $\geqslant 2$ (impairment not due to physical causes), Instrumental Ac-

https://doi.org/10.1017/S0033291701004743 Published online by Cambridge University Press

tivities of Daily Living (Lawton & Brody, 1969) score $\geqslant 3$ (impairment not due to physical cause), death certificate diagnosis, retrospective case audit to meet DSM-III-R dementia criteria or a diagnosis of dementia recorded in nursing home notes. Dementia was considered probable for cases meeting two or three of these criteria.

Other variables assessed but not reported in the current paper included family psychiatric history, patient medical history, neurological and physical examinations and a detailed neuropsychological assessment. Socio-economic class at 25 years was determined using the Daniel prestige scale (Daniel, 1983, pp. 196–206).

The differing number of subjects for each variable pertains first to the valid number of subjects, and secondly to the reliability of the data obtained. For example, in the depressive group HRSD ratings could only be completed on the 41 living subjects and only 34 subjects, i.e. those not dead or demented, could complete the BDI. Data not considered reliable were excluded. Reliability was determined by the interviewer based on the consistency of responses to several lines of questioning, hospital records and death certificates.

Statistical methods

Basic analyses compared the amalgamated ED/ND subgroups (i.e. DEPs) with either the PSYCON or the surgical control (SURGCON) groups. Despite original pairwise matching, missing data meant that control comparisons were groupwise. Where appropriate each of these main analyses was repeated for the comparison of the ED and ND subgroups. For normally distributed continuous data, two-sample t tests were employed. In the case of skewed data (such as years of education and the measures of current depression) non-parametric Mann–Whitney U analyses were used (denoted by Uz).

Pearson's chi-square analyses were used to determine associations between categorical data, except in the case of two by two cells (i.e. when there is only one degree of freedom) where Yates' continuity correction was applied (denoted by CC).

The comparison of the annualized rates of hospitalizations between follow-up periods for the DEP group only was analysed using the non-parametric Wilcoxon signed rank test for paired observations (denoted by Wz). To investigate whether there was a difference over time (i.e. change between follow-ups) between the two DEP subgroups (i.e. ED and ND) in terms of change in annualized depressive episode rates, a change variable was created (second follow-up rate − first follow-up rate).

RESULTS

Patient characteristics

Was the present sample comparable with the baseline sample recruited 25 years ago during their index admission? There were no differences at index assessment between those probands followed for 25 years and those followed for shorter periods of time according to diagnosis ($\chi^2 = 2{\cdot}33$, df = 2, $P = 0{\cdot}312$), socio-economic status ($\chi^2 = 5{\cdot}01$, df = 3, $P = 0{\cdot}171$), gender (CC $\chi^2 = 0{\cdot}01$, df = 1, $P = 0{\cdot}903$), HAM-D ($t = -0{\cdot}80$, df = 210, $P = 0{\cdot}427$, 95% CI −2·52, 1·07), BDI ($t = -1{\cdot}03$, df = 207, $P = 0{\cdot}305$, 95% CI −4·49, 1·41) or IPAT (Institute for Personality and Anxiety Scale) total score ($t = -0{\cdot}67$, df = 208, $P = 0{\cdot}504$, 95% CI −2·46, 5·00). As might be expected, those probands who were followed longest were significantly younger at index ($t = -2{\cdot}81$, df = 165·32, $P = 0{\cdot}006$, 95% CI −10·61, −1·85) and older index patients were more likely to have died by the 25 year follow-up ($t = 2{\cdot}43$, df = 143, $P = 0{\cdot}016$, 95% CI −12·99, −1·34).

There were 49 DEP patients (19 were originally diagnosed with endogenous depression, 30 with neurotic depression) and 22 with other neuroses (PSYCON) with outcome data at 25 years (Table 2). There were no significant differences between the DEP and PSYCON groups or the DEP and the surgical controls (SURGCON) on any demographic variable (see Table 2). However, the SURGCON group's average length of follow-up was approximately 2 years longer than the DEP group (see Table 2).

Diagnoses of definite dementia were made for eight patients (seven DEPs and one PSYCON). There were a further two cases of possible dementia (both DEPs). There were no cases of dementia in the surgical group. More detailed information regarding cases of dementia will be provided in a paper examining the neuropsychological data. Statistical comparison with the surgical controls was not appropriate because

https://doi.org/10.1017/S0033291701004743 Published online by Cambridge University Press

Table 2. *Demographic and clinical details of the 71 probands followed-up for 25 years and the 50 surgical controls*

Characteristic	ED ($N = 19$)	ND ($N = 30$)	DEP ($N = 49$)	PSYCON ($N = 22$)	SURGCON ($N = 50$)	DEP *v.* PSYCON			DEP *v.* SURGCON		
						Statistic	df	P†	Statistic	df	P
Length of follow-up (yrs)*	25·3 (1·5)	25·7 (1·4)	25·5 (1·4)	25·9 (1·4)	28·0 (1·6)	$t = -0{\cdot}89$	55	0·379	$t = -7{\cdot}23$	86	0·000
Age at interview or death (Range 40–92)*	73·4 (10·9)	57·8 (13·0)	63·8 (14·4)	60·1 (12·5)	61·0 (10·5)	$t = 1{\cdot}07$	69	0·289	$t = 1{\cdot}13$	87·9	0·262
Age at interview (Range 40–91)*	70·1 (10·6)	55·4 (10·9)	60·4 (12·8)	60·7 (13·0)	61·0 (10·5)	$t = -0{\cdot}06$	59	0·952	$t = -0{\cdot}22$	89	0·825
Gender: male, *N* (%)	11 (58)	6 (20)	17 (34·7)	9 (40·9)	19 (38·0)	CC $\chi^2 = 0{\cdot}06$	1	0·813	CC $\chi^2 = 0{\cdot}02$	1	0·894
Education, mean (S.D.)	9·7 (2·2)	10·3 (2·7)	10·1 (2·5)	10·6 (2·8)	11·5 (4·6)	$Uz = -1{\cdot}12$		0·262	$Uz = -1{\cdot}37$		0·170
Higher socio-economic class *N* (%)‡	5 (33·3)	3 (12·5)	8 (20·5)	4 (20·0)	17 (34·0)	$\chi^2 = 3{\cdot}93$	3	0·269	$\chi^2 = 8{\cdot}32$	4	0·081
Current marital status, *N* (%)											
Married/*de facto*	8 (57·1)	15 (60·0)	23 (59·0)	8 (42·1)	32 (64·0)						
Separated/divorced	1 (7·1)	7 (28·0)	8 (20·5)	6 (31·6)	8 (16·0)						
Widowed	4 (28·6)	3 (12·0)	7 (17·9)	1 (5·3)	5 (10·0)	$\chi^2 = 7{\cdot}88$	3	0·048	$\chi^2 = 3{\cdot}16$	3	0·367
Current depression§											
HAM-D, mean (S.D.)	4·3 (4·6)	3·8 (6·3)	4·0 (5·6)	3·8 (6·4)	1·9 (2·9)	$Uz = -0{\cdot}16$		0·869	$Uz = -1{\cdot}56$		0·120
BDI, mean (S.D.)	5·7 (1·9)	10·7 (7·4)	9·5 (6·8)	5·6 (8·3)	4·9 (4·7)	$Uz = -2{\cdot}68$		0·007	$Uz = -3{\cdot}51$		0·001
Current functional status ‖											
GAF impairment, No. (%)¶											
Nil	3 (15·8)	8 (26·7)	11 (22·4)	8 (38·1)	30 (60·0)	$\chi^2 = 2{\cdot}29$	2	0·319	$\chi^2 = 22{\cdot}59$	2	0·000
Mild	7 (36·8)	12 (40·0)	19 (38·8)	8 (38·1)	18 (36·0)						
Moderate–severe	9 (47·4)	10 (33·3)	19 (38·8)	5 (23·8)	2 (4·0)						
LIFE (global impairment), No. (%)											
Nil	8 (53·3)	17 (70·8)	25 (64·1)	13 (68·4)	39 (79·6)	$\chi^2 = 0{\cdot}19$	2	0·910	$\chi^2 = 5{\cdot}63$	2	0·060
Mild	3 (20·0)	5 (20·8)	8 (20·5)	3 (15·8)	9 (18·4)						
Moderate–marked	4 (26·7)	2 (8·3)	6 (15·4)	3 (15·8)	1 (2·0)				N/A		
Dementia											
Probable, *N* (%)	4 (21·0)	3 (10·0)	7 (14·3)	1 (4·5)		$\chi^2 = 2{\cdot}40$	1	0·16			
Possible, *N* (%)	2 (10·5)	0 (0·0)	2 (4·1)	0 (0·0)							
Total	6 (6·7)	3 (10·0)	9 (18·4)	1 (4·5)							

ED, endogenous depression; ND, neurotic depression; DEP, combined ED + ND; PSYCON, other neurosis controls; SURGCON, surgical controls.

* Values are mean (S.D.).

† *P* value for gender analysis and dementia analysis modified by a Fisher's Exact Test.

‡ Daniel Prestige scale: analysed as high (1–3 on the 7-point scale, none classified 1), medium (4–5) and low (6–7).

§ Information was missing for 20 DEPs, 7 PSYCONs, 26 SURGCONs on the HAM-D and 22 DEPs, 6 PSYCONs, 6 SURGCONs on the BDI.

‖ Current refers to time of interview with patient or with proxy (LIFE, Longitudinal Follow-Up Evaluation, is the rater's assessment of social adjustment and as such contains proxy information for two deceased DEPs and one deceased PSYCON).

¶ Global Assessment of Functioning: nil, score of 81–90; mild, 61–80; moderate to severe, < 61.

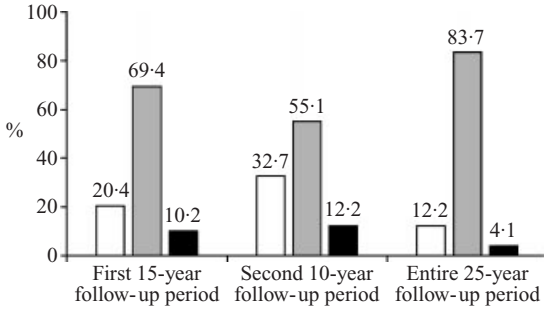

FIG. 1. Clinical outcome (using categories defined by Kiloh *et al.* 1988) of depression for the 49 patients diagnosed with depression at index and followed for 25 years. This does not include patients followed up for shorter periods of time (e.g. the nine DEPs and two PSYCONs who committed suicide in the first 15 years). (□, Recovered and continuously well; ▨, recovered with subsequent attacks and ■, always incapacitated or death by suicide.)

of the differences in recruitment methods for the two samples.

Depression outcome (DEPs only)

Clinical outcome

The clinical outcome categories are shown for the 0–15 year follow-up, the 16–25 year follow-up period and the two periods combined in Fig. 1. Forty-one (83·7%) of those 49 probands diagnosed with depression at index experienced at least one recurrence during the entire 25-year follow-up period. None of the probands developed bipolar affective disorder. There was no association between clinical outcome and length of follow-up for these 49 probands.

We were interested to examine change in outcome across the two major follow-up periods. Specifically, if someone had been rated as chronically depressed for 15 years, did they remain so for the next 10 years? Just over one-quarter (13/49 or 26·5%) had an improved clinical outcome during the second follow-up period, 55·1% remained in their original outcome category and 18·4% dropped to a poorer outcome category.

Of the five DEPs rated as chronically depressed for the 0–15 year period, only one an ED patient who had been psychotic, remained so for the 16–25 year follow-up. The patient was subsequently diagnosed as having possible dementia (see above for more detail). Of the remaining four DEPs who recovered, three had no evidence of gross cognitive impairment and one, who had a history of alcohol use, had an early dementia. The only two DEPs to receive ECT in the second follow-up period included the patient who remained chronic over the entire follow-up, and one of the DEPs who recovered.

Table 3. *Proportion of follow-up spent depressed by those with depression at baseline* (N = 41/49)

%	First 15-year follow-up period *N* (%)	Second 10-year follow-up period *N* (%)	Entire 25-year follow-up period *N* (%)
0	1 (2·4)	16 (39·0)	1 (2·4)
1–25	29 (70·7)	18 (43·9)	28 (68·3)
26–50	6 (14·6)	2 (4·9)	5 (12·2)
51–75	3 (7·3)	0 (0·0)	6 (14·6)
76–100	2 (4·9)	5 (12·2)	1 (2·4)

The index episode was the first episode of depression for 66% of the DEPs (61% of EDs and 69% of NDs; data were missing for two patients). The index episode was their first and only episode of depression for all of those who recovered and remained continuously well for the entire follow-up period. Conversely, both the DEP patient who was incapacitated throughout follow-up and the other who completed suicide (see below) had experienced previous episodes.

Suicide

Of the 49 probands followed through 25 years, one additional person completed suicide during the final 10-year follow-up period. The documented cause of death for this proband was 'multiple injuries', which were sustained after falling from a window the day following discharge from a private psychiatric clinic (admission followed patient complaints of depression). A further three DEPs (out of a total of 38 with reliable information) attempted suicide during the 16–25 year follow-up, each of them making two attempts. It should be noted that in the first 15 years, nine DEP probands and two PSYCONs suicided.

Proportion of follow-up spent depressed and number of episodes of depression

Forty-one DEPs had reliable information on the percentage of time spent depressed during the 25-year follow-up period (Table 3). The depressed probands had 2·4 (S.D. 1·5) episodes over the first

https://doi.org/10.1017/S0033291701004743 Published online by Cambridge University Press

15 years (or 0·17/year) and 0·6 (S.D. 0·9) over the next 10 years (0·06/year); in summary 3·0 (S.D. 1·8) episodes over the 25 years or 0·11 per year. There was a significant decline in episode rate per year over time ($Wz = -4{\cdot}46$, $P = 0{\cdot}000$).

Treatment received for depression (whole sample)

Hospitalizations for affective disorder

Fifty-eight per cent ($N = 25/43$) of the DEPs were hospitalized for depression at least once more during the entire follow-up period, 51·2% during the first 15 years and 23·3% during the final 10-year period. The average annual rate of hospitalization for depression post-index for the first follow-up period was 0·09 (S.D. 0·13), and during the 16–25 year follow-up was 0·04 (S.D. 0·08) for the DEP group, a reduction which was just significant ($Wz = -2{\cdot}03$, $P = 0{\cdot}042$). For the entire 25-year period the average annual rate of hospitalization for depression for the DEP group was 0·06 (S.D. 0·09), and for the PSYCON group 0·09 (S.D. 0·19, 7/18 hospitalized), and the SURGCON group 0·001 (S.D. 0.007, 2/50 hospitalized; DEP *v.* PSYCON, $Uz = -0{\cdot}49$, $P = 0{\cdot}627$; DEP *v.* SURGCON, $Uz = -5{\cdot}83$, $P = 0{\cdot}000$).

Number of ECT treatments received

Over the 25 years 36·6% (15/41) of the DEPs, 23·5% (4/17) of the PSYCONs and none of the surgical controls received at least one course of ECT. All of these 15 DEPs had ECT during the first follow-up period, and four also received ECT during the second follow-up period. The group differences are significant when comparing DEPs and SURGCONs only (DEP *v.* SURGCON, CC $\chi^2 = 19{\cdot}33$, df = 1, $P = 0{\cdot}000$).

Time spent on antidepressants

Over the entire follow-up period, 29 of 39 DEPs (74·4%) spent some time on antidepressant medication compared to 11 of 18 (61·1%) PSYCONs and four of 49 SURGCONs (8·2%) DEP *v.* PSYCON, CC $\chi^2 = 0{\cdot}50$, df = 1, P = 0·481; DEP *v.* SURGCON, CC $\chi^2 = 37{\cdot}82$, df = 1, P = 0·000). During the first 15 years of follow-up 25 DEPs took antidepressants, and during the final 10 years 18 DEPs used antidepressants, four of whom had not taken antidepressants during the first follow-up period.

Time spent on antipsychotics

Four DEPs took antipsychotics at some time during the entire 25 years (10·3%, 4/39), an equivalent number during the first follow-up period, and only one of these four DEPs continued their use during the second follow-up period. No one commenced antipsychotic use during the second follow-up period. Two of 16 PSYCON patients (12·5%), and two of 49 SURGCON patients (4·1%) took antipsychotics at some stage during the entire 25 years, however numbers were too small for statistical comparisons.

Current ratings of depression (see Table 2)

As there were significant amounts of missing data, comparisons of the ratings of current depression should be interpreted cautiously. Clinician ratings of current depression severity were uniformly low and there were no significant differences between the groups (HAM-D, DEP *v.* PSYCON, $P = 0{\cdot}869$; DEP *v.* SURGCON, $P = 0{\cdot}120$). However, the DEPs' assessments of their own level of depression on the BDI were significantly higher than either of the control groups (DEP *v.* PSYCON, $P = 0{\cdot}007$; DEP *v.* SURGCON, $P = 0{\cdot}001$). Five of 27 DEP (18·5%), two of 16 PSYCON (12·5%) and one of 44 SURGCON patients (2·3%) had BDI scores $\geqslant 16$, which are indicative of significant current depression (expected frequencies too low for analysis). However, clinician ratings indicated significant depression (i.e. HAM-D 16) in only one of 29 DEP (3·5%) and one of 15 PSYCON (6·7%) patients (expected frequencies too low for analysis).

Psychological morbidity

Composite International Diagnostic Interview (see Table 4)

There were 105 CIDIs performed (37 DEP, 18 PSYCON, 50 SURGCON). We excluded one PSYCON interview because the informant was considered unreliable and two DEP patients because the probands were deemed to have significant memory problems and no available informant. Thirteen of the 35 DEP patients (37·1%) with reliable CIDI data met criteria for at least one current DSM-III-R diagnosis. Five of the 17 PSYCON probands (29·4%) and four of the 50 SURGCON patients (8·0%) met

https://doi.org/10.1017/S0033291701004743 Published online by Cambridge University Press

Table 4. *Current* psychological morbidity*

CIDI DSM-III-R diagnosis	DEP ($N = 35$) N (%)	PSYCON ($N = 17$) N (%)	SURGCON ($N = 50$) N (%)
Anxiety disorders	8 (22·9)	2 (11·8)	4 (8·0)
Substance use disorders	4 (11·4)	0 (0·0)	0 (0·0)
Eating disorders	1 (2·9)	1 (5·9)	0 (0·0)
Somatoform disorders	0 (0·0)	0 (0·0)	0 (0·0)
Depression/dysthymia†	4 (11·4)	2 (11·8)	1 (2·0)

DEP, combined endogenous depression + neurotic depression; PSYCON, other neurosis controls; SURGCON, surgical controls.

* Current, was defined as present within the month prior to interview.

† The SURGCON patient met criteria for a current CIDI rating of bipolar disorder, depressed, mild.

criteria for at least one current DSM-III-R diagnosis (DEP *v.* PSYCON, CC $\chi^2 = 0{\cdot}06$, df $= 1$, $P = 0{\cdot}811$; DEP *v.* SURGCON, CC $\chi^2 = 9{\cdot}18$, df $= 1$, $P = 0{\cdot}002$).

Psychosocial outcome

Global Assessment of Functioning (see Table 2)

Global Assessment of Functioning was rated as the highest level of functioning for the months prior to interview or death. Taking GAF scores of $\leqslant 80$ as indicative of some degree of impairment, 38 of 49 DEPs (77·6%), 13 of 21 PSYCONs (61·9%) and 20 of 50 SURGCONs (40·0%) were suffering some degree of impairment. The DEP group had a similar proportion of patients with functional impairment to the PSYCON group, but significantly more than the SURGCON group (DEP *v.* PSYCON, $\chi^2 = 2{\cdot}29$, df $= 2$, $P = 0{\cdot}319$; moderate and severe categories collapsed because of low expected frequencies; DEP-*v.*-SURGCON, $\chi^2 = 22{\cdot}59$, df $= 2$, $P = 0{\cdot}000$).

Social functioning

At the end of the psychosocial interview, raters made a global assessment of social adjustment and asked for the probands' assessment of their own social adjustment. Fourteen of 39 DEPs (35·9%), six of 19 PSYCONs (31·6%) and 10 of 49 SURGCONs (20·4%) received interviewer ratings of mild to marked impairment (DEP *v.* PSYCON: $\chi^2 = 0{\cdot}19$, df $= 2$, $P = 0{\cdot}910$; DEP *v.* SURGCON, $\chi^2 = 5{\cdot}63$, df $= 2$, $P = 0{\cdot}060$). In comparison, six of 32 DEPs (18·8%), three of 17 PSYCONs (17·7%) and six of 49 SURGCONs (12·3%) rated themselves as experiencing some degree of impairment in social adjustment, though none more than 'moderate' impairment (DEP *v.* PSYCON, $\chi^2 = 6{\cdot}90$, df $= 3$, $P = 0{\cdot}075$; DEP *v.* SURGCON, $\chi^2 = 5{\cdot}31$, df $= 3$, $P = 0{\cdot}151$).

Comparisons of the two depressive subtypes

We examined the association between depressive subtype (endogenous *versus* neurotic depression) and outcome of depression, proportion of follow-up spent depressed, hospitalizations for depression, current ratings of depression and psychosocial outcome (GAF score and social functioning). The only significant differences were demographic i.e. the ED group were older and more often male and in the change in episode rates between the two major follow-ups, the ED group experiencing a greater reduction in episode rate over time (0–15 year follow-up rate, 0·12/year to 16–25 year follow-up rate, 0·06/year) compared to the ND group (0·07/year to 0·06/year) (age – at interview $t = 4{\cdot}14$, df $= 39$, $P = 0{\cdot}000$, at interview or death $t = 4{\cdot}36$, df $= 47$, $P = 0{\cdot}000$; gender – CC $\chi^2 = 5{\cdot}80$, df $= 1$, $P = 0{\cdot}016$ (see Table 2); episode rate – $Uz = -2{\cdot}45$, $P = 0{\cdot}014$). The age differences were apparent at baseline and persisted through to the 25-year follow-up. The gender distribution by diagnostic subtype at baseline was equivalent at index; gender differences at the 25-year follow-up resulted from differential death rates.

DISCUSSION

The outcome of this sample at 15 years following index discharge showed about one-fifth of patients recovered and remained continuously well, and one-fifth remained incapacitated or suicided (Kiloh *et al.* 1988). Similar patterns were found at the 18 year follow-up of the Kendell cohort (Lee & Murray, 1988) and at the 10 year follow-up of the Basel site of the WHO Collaborative study (Thornicroft & Sartorius, 1993). Another site using these criteria, Nagasaki, found about one-third in the best outcome category, one-fifth in the worst (Thornicroft & Sartorius, 1993). The current study revealed that only one-eighth of the sample who were included in the 25-year follow-up remained continuously well following index,

and 4% remained incapacitated or suicided (more specifically, one of 49 patients with depression at index remained chronically depressed over the 25 years and one patient committed suicide during the second follow-up period).

The high rate of dementia, two possible and seven probable cases among the 49 DEPs compared to none of the SURGCONs cannot be merely attributed to different times of initial recruitment as there was only one of 22 PSYCONs who developed probable dementia. Also from age-adjusted published prevalence rates, we would expect only 2·041 cases within the DEPs (Jorm & Henderson, 1987, p. 8). These findings accord with previous reports and warrant further study (Jorm *et al.* 1991; Devanand *et al.* 1996). Despite this grim picture, we also found just over one-quarter of the depressed patients (26·5%) improved in clinical outcome category over the two major follow-up periods.

The vast majority (84%) of those diagnosed with a depressive disorder at index experienced a further episode of depression over the 25-year follow-up. This figure is slightly higher than the 0·74 probability risk at 25 years follow-up based on extrapolation by Surtees & Barkley (1994) of 12 year follow-up data. (These authors also found no significant differences between ED and ND groups in terms of recurrence.).

Currently, most patients were not clinically depressed – only one of 29 patients (3·5%) on the HRSD and four of 35 (11·4%) on the CIDI were rated as having significant depression at the 25-year interview. Equivalent data for the PSYCON patients are one of 15 (6·7%) (according to HRSD) and two of 17 (11·7%) according to CIDI) and for the SURGCONs none of 24 and one of 50 (2·0%) patients respectively. However, DEP patients were experiencing depressive symptoms subjectively and were functioning below expectations. Each of the three groups had large proportions of patients experiencing inferior psychosocial functioning – approximately three-quarters of the DEPs, about two-thirds of the PSYCON, and two-fifths of the SURGCON had some impairment when rated using the GAF. These findings did not appear to be specific to depressive diagnoses but may be applicable to other psychiatric diagnoses requiring hospitalization. These high rates of impaired social function are in accord with previous reports (Tsuang *et al.* 1979; Thornicroft & Sartorius, 1993; Surtees & Barkley, 1994). There were no significant differences in terms of treatment received between the DEP group and the PSYCON group, but this could be result of low overall rates of treatment. Also each of the index psychiatric groups experienced additional psychiatric morbidity.

Essentially, there was no difference in outcome between depressed patients rated as endogenous or neurotic at index admission. Longitudinally, the ED patients experienced a reduction in annual episode rate between the first (0–15 years) and second (16–25 years) follow-ups, but over the 16–25 year period both groups had a common episode rate of 0·06/year. Unlike the outcome of this cohort at 15 years (Kiloh *et al.* 1988), and O'Leary & Lee's 7-year results (Nottingham ECT study; O'Leary & Lee, 1996), we did not find a significantly greater occurrence of readmission for the ED group over the 25 years (65% of ED readmitted *versus* 54% of the ND group).

The study was both prospective and retrospective, used multiple retrospective and concurrent ratings, included assessment of depression, other psychiatric morbidity and psychosocial functioning, and relied on well proven instruments. However, we concede several limitations. First, although the initial number of patients was 212, after 25 years our final sample was small, increasing the risk of Type II error and the number of comparisons was large, increasing the risk of Type I error. Secondly, selective attrition may have biased our findings, but we note the lack of difference (apart from age) of those completing and those not completing this latest follow-up. Thirdly, the use of an in-patient sample may overestimate chronicity. For example, over 10 years recurrence occurred in only 40% of general practice 'depressive' patients (van Weel-Baumgarten *et al.* 1998). Fourthly, the use of surgical patients to control for hospital admission *per se* may have inadvertently included psychosomatic cases (Arolt *et al.* 1997) even though we specified prior psychiatric history as an exclusion criterion. Fifthly, the recruitment methods for the psychiatric and surgical control groups were different. The psychiatric samples were selected at admission and followed through 25 years

while attempts were made to contact surgical controls, who were identified from consecutive case records, 25 years after admission. Biases would have been thus introduced, for example the exclusion of surgical patients who had developed dementia or died. Sixthly, there are difficulties inherent in the use of retrospective data (Andrews *et al*. 1999) despite our use of structured instruments as well as contemporaneous assessments.

We conclude that patients requiring hospital admission for depressive disorder (or other non-psychotic psychiatric diagnoses) have a chronic disorder with only one in eight remaining continuously well. The usual pattern is for recurrences and persistent symptoms, psychiatric co-morbidity and impaired level of psychosocial functioning. On the positive side, this study demonstrates the potential for improvement after many years of chronic depression. After 15 years of chronicity, four out of five patients improved at least one category of outcome in the following decade. Our data do not support the contention that with age episodes become more severe, longer in duration and more frequent (Post, 1968). Neither do we find any differences in long-term outcome between rigorously diagnosed neurotic and endogenous depressives.

Part of this study was presented at the RANZCP conference in Sydney, in May 1997.

The project was supported by NHMRC grant no. 950425.

The authors wish to thank Annette Koschera, Kylie Fell, Jenny Grice, Leanne Prenter and Karen Berman for their help in the collection and preparation of these data. Dr Peisah was partly supported by an Eli Lily Fellowship. The late Professor Leslie Kiloh kindly gave permission for the continuation of the follow-up.

REFERENCES

American Psychiatric Association (1987). *Diagnostic and Statistical Manual of Mental Disorders, 3rd edn. Revised (DSM-III-R)*. APA; Washington, DC.

Andrews, G. (2001). Should depression be managed as a chronic disease? *British Medical Journal* **322**, 419–421.

Andrews, G., Kiloh, L. G. & Neilson, M. (1973). Patterns of depressive illness: the compatibility of disparate points of view. *Archives of General Psychiatry* **29**, 670–673.

Andrews, G., Neilson, M., Hunt, C., Stewart, G. & Kiloh, L. G. (1990). Diagnosis, personality and the long-term outcome of depression. *British Journal of Psychiatry* **157**, 13–18.

Andrews, G., Anstey, K., Brodaty, H., Issakidis, C. & Luscombe, G. (1999). Recall of depressive episode 25 years previously. *Psychological Medicine* **29**, 787–791.

Angst, J. (1986). The course of affective disorders. *Psychopathology* **19** (suppl. 2), 47–52.

Angst, J. (1997). A regular review of the long term follow up of depression. *British Medical Journal* **315**, 1143–1146.

Arolt, V., Driessen, M. & Dilling, H. (1997). The Lubeck General Hospital Study. 1: prevalence of psychiatric disorders in medical and surgical inpatients. *International Journal of Psychiatry in Clinical Practice* **1**, 207–216.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J. & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry* **4**, 561–571.

Brodaty, H., MacCuspie-Moore, C. M., Tickle, L. & Luscombe, G. (1997). Depression, diagnostic sub-type and death: a 25 year follow-up study. *Journal of Affective Disorders* **46**, 233–242.

Daniel, A. E. (1983). *Power, Privilege and Prestige: Occupations in Australia*, pp. 196–206. Longman Cheshire: Melbourne, Australia.

Devanand, D. P., Sano, M., Tang, M. X., Taylor, S., Gurland, B. J., Wilder, D., Stern, Y. & Mayeux, R. (1996). Depressed mood and the incidence of Alzheimer's disease in the elderly living in the community. *Archives of General Psychiatry* **53**, 175–182.

Folstein, M. F., Folstein, S. E. & McHugh, P. R. (1975). 'Mini-Mental State': a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* **12**, 189–198.

Hall, R. C. W. & Wise, M. G. (1995). The clinical and financial burden of mood disorders: cost and outcome. *Psychosomatics* **36**, S11–S18.

Hamilton, M. (1960). A rating scale of depression. *Journal of Neurology, Neurosurgery & Psychiatry* **23**, 56–61.

Hirschfeld, R. M. A., Keller, M., Bourgeois, M., Baldwin, D. S., Healy, D., Humble, M., Kasper, S. & Montgomery, S. A. (1998). Focus on social functioning in depression. *International Journal of Psychiatry in Clinical Practice* **2**, 241–243.

Hughes, C. P., Berg, L., Danziger, W. L., Coben, L. A. & Martin, R. L. (1982). A new clinical scale for the staging of dementia. *British Journal of Psychiatry* **140**, 566–572.

Jorm, A. F. & Henderson, A. S. (1987). *The Problem of Dementia in Australia*. National Health and Medical Research Council, Social Psychiatry Research Unit, Australian National University. Australian Government Publishing Service: Canberra, Australia.

Jorm, A. F. & Jacomb, P. A. (1989). The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): sociodemographic correlates, reliability, validity and some norms. *Psychological Medicine* **19**, 1015–1022.

Jorm, A. F., Van Duijn, C. M., Chandra, V., Fratiglioni, L., Graves, A. B., Heyman, A., Kokmen, E., Kondo, K., Mortmer, J. A., Rocca, W. A., Shalat, S. L., Soininen, H. & Hofman, A. (1991). Psychiatric history and related exposures as risk factors for Alzheimer's disease: a collaborative re-analysis of case-control studies. *International Journal of Epidemiology* **20** (suppl. 2), 43–47.

Judd, L. J. (1997). The clinical course of unipolar major depressive disorders. *Archives of General Psychiatry* **54**, 989–991.

Judd, L. J., Paulus, M. P., Wells, K. B. & Rapaport, M. H. (1996). Socioeconomic burden of subsyndromal depressive symptoms and major depression in a sample of the general population. *American Journal of Psychiatry* **153**, 1411–1417.

Judd, L. J., Akiskal, H. S., Maser, J. D., Zeller, P. J., Endicott, J., Coryell, W., Paulus, M. P., Kunovac, J. L., Leon, A. C., Mueller, T. I., Rice, J. A. & Keller, M. B. (1998). A prospective 12-year study of subsyndromal and syndromal depressive symptoms in unipolar major depressive disorders. *Archives of General Psychiatry* **55**, 694–700.

Katz, S. & Apkom, C. A. (1976). A measure of primary sociobiological functions. *International Journal of Health Services* **6**, 493–508.

Keller, M. B., Lavori, P. W., Friedman, B., Nielsen, E., Endicott, J., McDonald-Scott, P. & Andreasen, N. C. (1987). The Longitudinal Follow-up Evaluation. A comprehensive method for assessing outcome in prospective longitudinal studies. *Archives of General Psychiatry* **44**, 540–548.

Kiloh, L. G. & Garside, R. (1963). The independence of neurotic depression and endogenous depression. *British Journal of Psychiatry* **109**, 451–463.

Kiloh, L. G., Andrews, G., Neilson, M. & Bianchi, G. N. (1972*a*). The relationship syndromes called endogenous and neurotic depression. *British Journal of Psychiatry* **121**, 183–196.

Kiloh, L. G., Andrews, G., Bianchi, G. N. & Neilson, M. (1972*b*). On studying depression. *Australian and New Zealand Journal of Psychiatry* **6**, 85–98.

Kiloh, L. G., Andrews, G. & Neilson, M. (1988). The long-term outcome of depressive illness. *British Journal of Psychiatry* **153**, 752–757.

Lawton, M. P. & Brody, E. M. (1969). Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist* **9**, 179–186.

Lee, A. S. & Murray, R. M. (1988). The long-term outcome of Maudsley depressives. *British Journal of Psychiatry* **153**, 741–751.

MIMS Australia (1995). *MIMS Annual 1995*. Tien Wah Press: Singapore.

Mueller, T. I., Leon, A. C., Keller, M. B., Solomon, D. A., Endicott, J., Coryell, W., Warshaw, M. & Maser, J. D. (1999). Recurrence after recovery from major depressive disorder during 15 years of observational follow-up. *American Journal of Psychiatry* **156**, 1000–1006.

Murphy, J. M., Olivier, D. C., Monson, R. R., Sobol, A. M., Federman, E. B. & Leighton, A. H. (1991). Depression and anxiety in relation to social status. A prospective epidemiologic study. *Archives of General Psychiatry* **48**, 223–229.

Murray, C. J. L. & Lopez, A. D. (eds.) (1996). *The Global Burden of Disease*. The Harvard School of Public Health on behalf of The World Health Organization and the World Bank: Geneva, Switzerland.

O'Leary, D. A. & Lee, A. S. (1996). Seven year prognosis in depression. Mortality and readmission risk in the Nottingham ECT cohort. *British Journal of Psychiatry* **169**, 423–429.

Opjordsmoen, S. (1989). Long-term course and outcome in unipolar affective and schizoaffective psychoses. *Acta Psychiatrica Scandinavica* **79**, 317–326.

Piccinelli, M. & Wilkinson, G. (1994). Outcome of depression in psychiatric settings. *British Journal of Psychiatry* **164**, 297–304.

Post, F. (1968). The factor of ageing in affective illness. In *Recent Developments in Affective Disorders: A symposium* (ed. A. Coppen and A. Walk), pp. 105–116. Headley Brothers Ltd.: Ashford, Kent.

Smith, E. M. & North, C. S. (1988). Familial subtypes of depression: a longitudinal perspective. *Journal of Affective Disorders* **14**, 145–154.

Stephens, H. & McHugh, P. R. (1991). Characteristics and long-term follow-up of patients hospitalized for mood disorders in the Phipps clinic, 1913–1940. *Journal of Nervous and Mental Disease* **179**, 64–73.

Surtees, P. G. & Barkley, C. (1994). Future imperfect: the long-term outcome of depression. *British Journal of Psychiatry* **164**, 327–341.

Thornicroft, G. & Sartorius, N. (1993). The course and outcome of depression in different cultures: 10-year follow-up of the WHO Collaborative Study on the Assessment of Depressive Disorders. *Psychological Medicine* **23**, 1023–1032.

Tsuang, M. T., Woolson, R. F. & Fleming, J. A. (1979). Long-term outcome of major psychoses. I. Schizophrenia and affective disorders compared with psychiatrically symptom-free surgical conditions. *Archives of General Psychiatry* **36**, 1295–1301.

van Weel-Baumgarten, E., van den Bosch, W., van den Hoogen, H. & Zitman, F. G. (1998). Ten year follow-up of depression after diagnosis in general practice. *British Journal of General Practice* **48**, 1643–1646.

Winokur, G. & Morrison, J. (1973). The Iowa 500: follow-up of 225 depressives. *British Journal of Psychiatry* **123**, 543–548.

Winokur, G. & Tsuang, M. T. (1996). *The Natural History of Mania, Depression and Schizophrenia*. American Psychiatric Press: Washington, DC.

World Health Organization (1967). *Manual of the International Classification of Diseases, Injuries and Causes of Death, revision 8* (*ICD-8*). WHO: Geneva.

World Health Organization (1993). *The Composite International Diagnostic Interview:* (*CIDI, Version 1.1*). WHO: Geneva.