Time to recovery of an inception cohort with hitherto untreated unipolar major depressive episodes

TOSHIAKI A. FURUKAWA, TOSHINORI KITAMURA and KIYOHISA TAKAHASHI

Background Generalisability of existing studies on the naturalistic history of major depression is undermined by overrepresentation of in-patients and tertiary care academic centres, inclusion of patients already on treatment and/or incomplete follow-up.

Aims To report the time to recovery of an inception cohort of unipolar major depressive episodes.

Method A multi-centre prospective follow-up study of patients with a mood disorder, who had been selected to be representative of the untreated first-visit patients at 23 psychiatric settings from all over Japan.

Results The median time to recovery of the index episode after treatment commencement was 3 months (95% CI 2.5–3.6): 26% of the cohort reached asymptomatic or minimally symptomatic status by 1 month, 63% by 3 months, 85% by 12 months and 88% by 24 months.

Conclusions Our estimate of the episode length was 25–50% shorter than estimates reported in the literature.

Declaration of interest No conflict of interest. Funding from the Ministry of Health and Welfare, Japan.

How long does an episode of unipolar major depression last after we begin to treat it? Are there any prognostic factors that exert substantive influences on the estimated episode duration? Undoubtedly these are important questions for clinicians and patients. Studies available to date suggest that the median time to recovery after study entry may be 6–12 months (Klein *et al*, 1988; Goering *et al*, 1992; Keller *et al*, 1992; Wells *et al*, 1992). Few clinical features have been identified that predict chronicity (Angst & Preisig, 1995). However, these studies each suffer from methodological flaws (Keller *et al*, 1984) and may fail to provide sound evidence on which we can rely in our day-to-day practices. We would therefore like to report the course of an inception cohort of hitherto untreated patients with unipolar major depressive disorder presenting to various psychiatric facilities. We will present: a survival curve for unipolar major depressive episodes since the time they presented to psychiatrists and received treatment for the first time; a survival curve for unipolar major depressive episodes since the onset of the episodes, retrospectively ascertained; exploratory analyses of predictor variables for the course, measured at commencement of treatment.

METHOD

The Group for Longitudinal Affective Disorders Study (GLADS) has been conducting detailed prospective serial assessments of a cohort of patients with broadly defined affective disorders under naturalistic conditions. The 23 collaborating centres included psychiatric departments of 13 university hospitals and 6 general hospitals, 3 mental hospitals and 1 community mental health centre from all over Japan. In Japan we do not have the family doctor system and psychiatrists are often the first-line doctors that people consult when they realise that their problems are mental rather than physical.

Participating psychiatrists at each centre administered a semi-structured interview called the Psychiatric Initial Screening for Affective Disorders (PISA; Kitamura, 1992) to a representative subset of its first-visit patients in order to ascertain the patients' eligibility. The details of the predetermined rules on how to select a subset of first-visit patients were left to individual centres, depending on their human and logistic resources: some centres administered PISA to all their first-visit patients, others did so with those on a certain day of the week, and still others did so with those seen by one or two collaborating psychiatrists only. The eligibility criteria were:

(a) depressive state, defined as presenting with depressed mood or anhedonia lasting longer than 4 days, or manic state, defined as presenting with elated, expansive or irritable mood lasting longer than 4 days;

(b) having received no antidepressant or antipsychotic medication in the preceding 3 months;

(c) aged 18 years or older;

(d) absence of conditions such as mental retardation, dementia or hearing disability, which would render detailed psychopathological assessment difficult.

Out of all the eligible subjects, each participating centre was expected to enter the first such patient every 1 or 2 months. Written informed consent was obtained from all participants after full disclosure of the purposes and procedures of the study.

The patients eligible for and consenting to the study were then interviewed within 1 week of entry by a psychiatrist using the entry version of the Comprehensive Assessment List for Affective Disorders (COALA; Furukawa, 1992). The COALA consists of a series of semi-structured interviews that enable serial assessment of the cohort; these include the entry version, monthly follow-up version and 6-monthly follow-up version. It provides depression severity scores according to the 17-item Hamilton Rating Scale for Depression (HRSD; Hamilton, 1986). The reliability of the PISA and COALA has been reported to be good to excellent (Furukawa *et al*, 1995). The cohort was followed up monthly until treatment termination and 6-monthly thereafter up to 2 years. The course chart

https://doi.org/10.1192/bjp.177.4.331 Published online by Cambridge University Press

for the first 24 months of the follow-up was constructed based on these data. In this naturalistic study, the cohort received, on average, 60 (s.d.=44) mg of imipramine or equivalent per day on entry and 85 (s.d.=73) mg at 1 month (Furukawa *et al*, 2000).

The present paper focuses on the course of the subset of the cohort who were diagnosed as suffering from unipolar major depressive disorder, not superimposed on dysthymic disorder, according to DSM–IV (American Psychiatric Association, 1994). We excluded major depressive disorders superimposed on dysthymic disorder (double depression) because there is empirical evidence that these have a distinctively poorer prognosis (Klein *et al*, 1988; Wells *et al*, 1992). We defined recovery from a major depressive episode in accordance with the US National Institute for Mental Health (NIMH) definition (Keller *et al*, 1992) as a consecutive 2 months with no more than one or two mild depressive symptoms. The duration of an episode was calculated excluding these last 2 months in remission.

We used the statistical package SPSS for Windows 8.0 (SPSS Inc., 1997) to perform Kaplan–Meier survival analyses to depict survival curves of major depressive episodes, and Cox regression analyses for exploratory analysis of their predictors.

Table 1 Clinical characteristics of the cohort (n=90)

Age (years), mean (s.d.)	43.7 (14.9)
Gender, *n* (%) female	52 (58%)
Education (years), mean (s.d.)	11.7 (2.9)
Marital status	
Single, *n* (%)	25 (28%)
Married, *n* (%)	65 (72%)
Treatment settings	
University hospital, *n* (%)	55 (61%)
General hospital, *n* (%)	23 (26%)
Mental hospital, *n* (%)	12 (13%)
Axis I comorbidity	
Panic disorder, *n*	3
Generalised anxiety disorder, *n*	2
Social phobia, *n*	1
Anorexia nervosa, *n*	1
Alcohol intoxication, *n*	1
Vascular dementia, *n*	1
In-patient status at entry, *n* (%)	14 (15%)
Length of episode before entry (months), median (range)	3.0 (0.47–48.0)
HRSD score during the worst week of index episode, mean (s.d.)	25.8 (7.3)
HRSD score during the week preceding the intake interview, mean (s.d.)	20.2 (8.6)

HRSD, Hamilton Rating Scale for Depression.

RESULTS

Of the 126 patients who entered the study, 95 met the DSM–IV criteria for major depressive disorder, either single episode (n=67) or recurrent (n=28). The major depressive disorder was superimposed on pre-existing dysthymia in five of these patients. In the following analyses, we will therefore concentrate on the 90 subjects who were diagnosed with unipolar major depressive disorder, not superimposed on dysthymic disorder, and who had received no antidepressant therapy for the index episode before study enrolment (Table 1).

We recorded recovery in 78 (87%) of our cohort. Seven never satisfied the recovery criteria for the 24 months of follow-up, and one committed suicide at 7 months without ever attaining recovery. The follow-up was therefore incomplete in only four (4%) of the total sample.

Figure 1 shows the cumulative probability of remaining in the index major depressive episode for the total cohort: 74% (95% CI 65–83%) were still in the index episode 1 month after entry into the study; 63% (52–73%) recovered within 3 months, 77% (68–86%) recovered within 6 months and 85% (77–93%) recovered within 12 months; 12% (5–18%) were still in the index episode at 24 months of follow-up. The median duration of an episode after entry into the study was 3.0 months (2.5–3.6 months). The mean duration with the upper limit of 24 months was 5.6 months (4.1–7.1 months).

Figure 2 shows the survival curve of the major depressive episode since its onset: 24% (15–32%) recovered within 3 months, 47% (37–57%) within 6 months, 65% (55–75%) within 1 year, 86% (79–94%) within 2 years and 92% (84–99%) within 4 years. The median duration of the total major depressive episode was 7.0 months (5.2–8.8 months).

We next performed exploratory analyses of predictor variables, measured at treatment commencement, of the time to recovery of the major depressive episodes not superimposed on dysthymia. Variables entered in univariate Cox regression analyses included age, gender, education, marriage status, treatment setting, in-patient status, length of the index episode before treatment, single episode or recurrent, HRSD scores during the worst week of the index episode and during the week preceding the intake interview, endogenicity as defined by DSM–IV, psychotic features, panic attack, physical illness, Axis I comorbidity and family history of major depression. Only two variables emerged as predictors that were statistically significant at the conventional P level of 0.05: the HRSD score during the worst week of the index episode (β=−0.034, P=0.04) and the presence of psychotic features (β=0.89, P=0.03). Entering two variables at the same time into Cox regression did not alter the estimated β values materially, and both remained statistically significant. When we examined subgroups defined by these two variables, however, the 95% CIs of the

https://doi.org/10.1192/bjp.177.4.331 Published online by Cambridge University Press

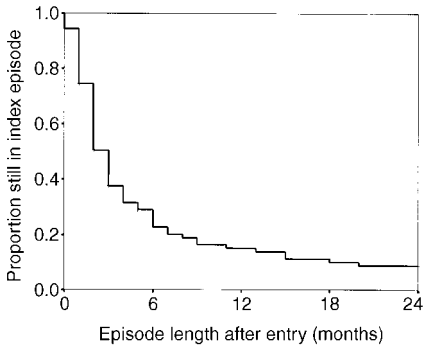

Fig. 1 Cumulative probability of remaining in the index episode after treatment commencement for the 90 probands with DSM–IV major depressive disorder not superimposed on dysthymia. Patients who recovered within a few days after treatment commencement were regarded as attaining recovery at 0 month.

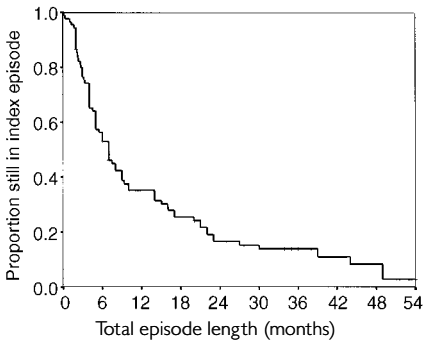

Fig. 2 Cumulative probability of remaining in the index episode since its onset for the 90 probands with DSM–IV major depressive disorder not superimposed on dysthymia.

estimated probabilities of recovery within 24 months for the subgroups overlapped with those for the whole group.

DISCUSSION

Prior studies and their limitations

A number of prospective studies are available on the naturalistic history of major depression but they each suffer from methodological limitations.

One of the first large-scale systematic studies on this subject is the Burghölzli study, in which Angst and his colleagues followed up 186 patients with unipolar depression hospitalised in the Zurich University Psychiatric Hospital (Burghölzli) at 5-year intervals for almost 30 years. The median duration of episodes for unipolar depression was 5.6 months. Chronicity, defined as an episode lasting at least 24 months without recovery, developed in 13% but it was impossible to predict chronicity from clinical variables (Angst & Preisig, 1995). The main weaknesses of this study are, first, that it was based on in-patients only, who are unlikely to be representative of patients with unipolar depression in general, because only a small minority of subjects with major depression are ever hospitalised. Second, the investigators were unable to use operationalised criteria to diagnose depression at the commencement of the study and did not employ any operationalised criteria to judge recovery and hence to determine the episode length.

Because estimates of episode length may show up to seven-fold variation, depending on the definition of recovery (Philipp & Fickinger, 1993), we will concentrate on studies that defined recovery in accordance with the NIMH criteria as the beginning of a period of at least 8 consecutive weeks with no more than one or two mild depressive symptoms. The NIMH Collaborative Depression Study (CDS) is a long-term, naturalistic cohort study of patients with mood disorder who sought treatment at five leading academic medical centres across the USA. A total of 431 patients entered the study in an episode of major depression, with no history of mania, hypomania, schizoaffective disorder or dysthymia. The Kaplan–Meier method, which takes into account the 26 patients (6%) with whom contact was lost before the first 5-year semi-annual follow-up, showed a cumulative probability of recovery of 54% at 6 months, 70% at 1 year and 81% at 2 years (Keller *et al*, 1992). In other words, the median duration of a major depressive episode after entry into the study was slightly less than 6 months. Because the median duration of the index episode before entry was about 6 months, the total episode length was estimated to be around 12 months (Keller *et al*, 1982*b*). Again, the large majority (77%) of the probands were in-patients at intake. Over 80% of these patients were receiving some treatment before being enrolled in the CDS (Keller *et al*, 1982*a*). These results are therefore subject to two kinds of bias: 'referral filter bias' because the patients were recruited in nationally renowned tertiary care centres; and 'lead-time bias' because they were not recruited at a similar point in time in the course of the disorder. The CDS investigators themselves acknowledge these weaknesses (Keller *et al*, 1984).

Other studies have similar methodological difficulties. For example, Goering *et al* (1992) dealt with in-patients only. The study by Klein *et al* (1988) and the Medical Outcomes Study (Wells *et al*, 1992) recruited out-patients, but in the former no mention was made of previous treatment and in the latter 10–30% had been on antidepressant medication at or just prior to baseline. Furthermore, the follow-up rates were less than satisfactory. The 6-month follow-up rate was 70% for the former study. In the latter, there was 30% loss for the first-stage screening, a further 40% loss for the second-stage interview and a further 26% loss for the one- and two-year follow-up telephone interviews.

The present study and its strengths

We planned the present study in order to surmount some of the difficulties noted above.

The strengths of the present study are as follows. First, our cohort was an inception cohort of patients with major depressive disorder who received antidepressant therapy for the first time for their index episode. The possibility of lead-time bias is minimised. Second, our cohort was representative of various psychiatric settings. Although the 23 participating centres of the GLADS project were not a random selection from all the psychiatric institutions in Japan, they consist of various types of institutions from all over Japan, and within each facility the selected sample was representative of the eligible first-visit patients during the study period. The cohort was not restricted to in-patients. The study is therefore less subject to referral filter bias than, for example, studies conducted on in-patients at one or a few academic institutions. Third, we performed prospective, serial assessments with reliable semi-structured interviews and applied predefined operational criteria to determine recovery. Our study was therefore able to minimise detection biases due to inaccurate recall or inconsistent application of decision criteria. Finally, the follow-up rate was satisfactory. The

https://doi.org/10.1192/bjp.177.4.331 Published online by Cambridge University Press

data concerning recovery or non-recovery by 24 months were available in 96% of the original cohort.

Possible weaknesses of the present study

The sample size may appear modest in comparison with some of the foregoing studies. However, a sample need only be large enough to allow precise estimates of the outcome (prevention of random error) and be representative enough to allow unbiased estimates thereof (prevention of systematic error). Ours was large enough to allow estimates with fairly narrow confidence intervals and reflected a wide range of psychiatric clinical settings within Japan.

With regard to certain infrequently observed prognostic variables, however, our study apparently suffers from low statistical power. Thus, for example, the statistically non-significant effect of Axis I comorbidity cannot exclude possible influences of some of the comorbid conditions. Nor would the lack of centre effect on the illness course rule out possible differences across treatment settings. We need a much larger cohort to elucidate these possibilities.

This study was designed from the perspective of what clinicians need to know about typical cases of depression in practice. We claim no generalisability beyond clinical samples, because only a limited proportion of people who suffer from major depression seek medical help. A community study is necessary to address the issue of the course of all persons with depression, both treated and untreated.

By the same token, the survival curve of the total episode of our cohort does not represent the total course of major depressive episodes in the general population. The course of a single episode since onset is influenced by the patient's illness behaviour and by filters along pathways to care in addition to treatment response, and our data do not allow examination of the relative contributions of these variables. Our estimate of the total episode length would, however, allow appreciation of the total duration of suffering among those patients whom clinicians start to treat.

CLINICAL IMPLICATIONS

- The median time to recovery of a hitherto untreated major depressive episode was 3.0 months: 26% of the cohort reached asymptomatic or minimally symptomatic status by 1 month, 63% by 3 months, 77% by 6 months, 85% by 12 months and 88% by 24 months.
- Psychiatrists starting to treat a unipolar major depressive episode and their patients are entitled to brighter prospects with regard to the index episode than heretofore suggested by the literature.
- At the same time, psychiatrists should not forget that 12% of patients remain in an episode for 24 months after starting treatment.

LIMITATIONS

- The sample size may appear modest in comparison with some prior studies, but was large enough to allow narrow estimates of the recovery rates and episode length.
- We found no substantively important prognostic factors that would enable us to make prognostic estimates for individual patients. Some of these non-significant findings with regard to predictor variables, however, may be due to the low statistical power of the present study.
- We claim no generalisability beyond clinical samples, because it is known that only a limited proportion of people who suffer from major depression present themselves for medical attention.

TOSHIAKI A. FURUKAWA, MD, Department of Psychiatry, Nagoya City University Medical School; TOSHINORI KITAMURA, FRCPsych, KIYOHISA TAKAHASHI, MD, National Center of Neurology and Psychiatry, Japan

Correspondence: Professor Toshi Furukawa, Department of Psychiatry, Nagoya City University Medical School, Mizuho-cho, Mizuho-ku, Nagoya 467-8601, Japan. Tel.: +81 52 853 8271; fax: +81 52 852 0837; e-mail: furukawa@med.nagoya-cu.ac.jp

(First received 27 September 1999, final revision 10 May 2000, accepted 15 May 2000)

The prognosis of a major depressive episode is more benign than previously reported

Within these constraints, we found that the median time to recovery of a major depressive episode, not superimposed on dysthymia, after it came under medical management was 3.0 months (95% CI 2.5–3.6 months). A newly treated episode reached asymptomatic or minimally symptomatic status in 26% of the cohort by 1 month, in 63% by 3 months, in 77% by 6 months and in 85% by 12 months. However, 12% did not recover even after 24 months. Adding the time spent in major depressive episode before treatment, the median duration of a single episode was 7.0 months (5.2–8.8 months).

Two prognostic factors emerged in *post hoc* explanatory analyses unadjusted for multiple comparisons: the sicker the patient was during the episode, the longer it took for him/her to recover; and patients with psychosis appeared to recover earlier than those without. The first predictor is intuitively plausible and is in line with other studies (Sargeant *et al*, 1990). The second is in contradiction with many earlier studies on the prognostic significance of psychotic symptoms in unipolar depression (Coryell *et al*, 1996). However, we examined nearly

https://doi.org/10.1192/bjp.177.4.331 Published online by Cambridge University Press

20 predictor variables and ours may very well be chance findings. The fact that the resultant prognostic factors may not be consistent with previous studies lessens their credibility. Stratifying the sample by these two factors did not reveal any subgroup with a distinctively and conclusively different prognosis. Being aware of the pitfalls of *post hoc* subgroup analyses, clinicians and patients are able to rely more safely on the overall estimates (Laupacis *et al*, 1994).

Our estimate of the length of episode after study entry (median=3 months; 95% CI 2.5–3.6 months) is 25–50% shorter than estimates from the foregoing studies: 6 months in the NIMH CDS (Keller *et al*, 1992) and the studies by Klein *et al* (1988) and Goering *et al* (1992), or 12 months in the Medical Outcomes Study (Wells *et al*, 1992). The total episode length (median=7 months; 95% CI 5.2–8.8 months) is again shorter than the 12 months taken from the NIMH CDS (Keller *et al*, 1992). It may be in line with the Burghölzli study (Angst & Preisig, 1995) but this comparison is hard to interpret because Angst *et al* did not clearly define what constituted recovery from an episode. Our initial suspicion that the foregoing studies were biased towards sicker patients due to referral filter bias and/or lead-time bias was largely borne out.

Psychiatrists starting to treat unipolar major depressive episodes and their patients are, on the average, entitled to brighter prospects with regard to their index episode than heretofore suggested by the literature. At the same time, we should not lose sight of the fact that 12% (95 CI 5–18%) of patients suffer continuously from a depressive episode for 24 months after coming under medical management.

ACKNOWLEDGEMENTS

This paper was prepared on behalf of the Group for Longitudinal Affective Disorders Study (GLADS). This study was supported by Research Grants 3A-6, 6A-4, 8B-2 and 10A-5 for Nervous and Mental Disorders from the Ministry of Health and Welfare, Japan.

REFERENCES

American Psychiatric Association (1994) *Diagnostic and Statistical Manual of Mental Disorders* (4th edn) (DSM–IV). Washington, DC: APA.

Angst, J. & Preisig, M. (1995) Course of clinical cohort of unipolar, bipolar and schizoaffective patients: results of a prospective study from 1959 to 1985. *Schweizer Archiv für Neurologie und Psychiatrie*, **146**, 5–16.

Coryell, W., Leon, A, Winokur, G., *et al* (1996) Importance of psychotic features to long-term course in major depressive disorder. *American Journal of Psychiatry*, **153**, 483–489.

Furukawa, T. (1992) *Comprehensive Assessment List for Affective Disorders (COALA)* (in Japanese). Ichikawa: National Institute of Mental Health, National Center for Neurology and Psychiatry.

—, Takahashi, K., Kitamura, T., *et al* (1995) The Comprehensive Assessment List for Affective Disorders (COALA): a polydiagnostic, comprehensive, and serial semistructured interview system for affective and related disorders. *Acta Psychiatrica Scandinavica Supplementum*, **387**, 1–36.

—, Kitamura, T. & Takahashi, K. (2000) Treatment received by depressed patients in Japan and its determinants: naturalistic observation from a multi-center collaborative follow-up study. *Journal of Affective Disorders* (in press).

Goering, P. N., Lancee, W. J. & Freeman, S. J. J. (1992) Marital support and recovery from depression. *British Journal of Psychiatry*, **160**, 76–82.

Hamilton, M. (1986) The Hamilton rating scale for depression. In *Assessment of Depression* (eds N. Sartorius & T. A. Ban), pp. 143–152. Berlin: Springer.

Keller, M. G., Klerman, G. L., Lavori, P. W., *et al* (1982*a*) Treatment received by depressed patients. *Journal of the American Medical Association*, **248**, 1848–1855.

—, Shapiro, R. W., Lavori, P. W., *et al* (1982*b*) Recovery in major depressive disorder: analysis with the life table and regression models. *Archives of General Psychiatry*, **39**, 905–910.

—, Klerman, G. L., Lavori, P. W., *et al* (1984) Long-term outcome of episodes of major depression: clinical and public health significance. *Journal of the American Medical Association*, **252**, 788–792.

—, Lavori, P. W., Mueller, T. I., *et al* (1992) Time to recovery, chronicity, and levels of psychopathology in major depression: a 5-year prospective follow-up of 431 subjects. *Archives of General Psychiatry*, **49**, 809–816.

Kitamura, T. (1992) *Psychiatric Initial Screening for Affective Disorders (PISA)* (in Japanese). Ichikawa: National Institute of Mental Health, National Center for Neurology and Psychiatry.

Klein, D. N., Taylor, E. B., Harding, K., *et al* (1988) Double depression and episodic major depression: demographic, clinical, familial, personality, and socioenvironmental characteristics and short-term outcome. *American Journal of Psychiatry*, **145**, 1226–1231.

Laupacis, A., Wells, G., Richardson, S., *et al* (1994) Users' guides to the medical literature: V. How to use an article about prognosis. *Journal of the American Medical Association*, **272**, 234–237.

Philipp, M & Fickinger, M. P. (1993) The definition of remission and its impact on the length of a depressive episode. *Archives of General Psychiatry*, **50**, 407–408.

Sargeant, J. K., Bruce, M. L., Florio, L. P., *et al* (1990) Factors associated with 1-year outcome of major depression in the community. *Archives of General Psychiatry*, **47**, 519–526.

SPSS Inc. (1997) *SPSS for Windows Version 8.0*. Chicago, IL: SPSS Inc.

Wells, K. B., Burnam, M. A., Rogers, W., *et al* (1992) The course of depression in adult outpatients: results from the Medical Outcomes Study. *Archives of General Psychiatry*, **49**, 788–794.

https://doi.org/10.1192/bjp.177.4.331 Published online by Cambridge University Press