Long-Term Effectiveness of Lithium in Bipolar Disorder:
A Multicenter Investigation of Patients With
Typical and Atypical Features

Anne Berghöfer, M.D.; Martin Alda, M.D.; Mazda Adli, M.D.;
Christopher Baethge, M.D.; Michael Bauer, M.D., Ph.D.; Tom Bschor, M.D.;
Tasha Glenn, Ph.D.; Paul Grof, M.D.; Bruno Müller-Oerlinghausen, M.D.;
Janusz Rybakowski, M.D.; Alexandra Suwalska, M.D.; and Andrea Pfennig, M.D., M.Sc.

Objective: Poor response to long-term lithium treatment has been reported to be associated with atypical features of bipolar disorder. The purpose of this study was to investigate the influence of atypical symptoms on the effectiveness and stability of long-term lithium treatment in a prospective, multicenter cohort of bipolar patients in a naturalistic setting.

Method: Patients were initially selected according to International Classification of Diseases, 8th Revision, criteria for bipolar disorder and required long-term treatment. Their diagnoses were reconfirmed according to DSM-IV upon its publication. They were prospectively followed for an approximately 20-year period ending in 2004 in 5 centers participating in the International Group for the Study of Lithium-Treated Patients. Examinations included a comprehensive psychiatric evaluation, an assessment of typical and atypical features on an 8-item scale, and an evaluation of clinical course using the morbidity index. Unbalanced repeated-measures regression models with structured covariance matrices were used to assess the extent to which the morbidity index was influenced by atypical symptoms, duration of treatment, and pretreatment features.

Results: A total of 242 patients were followed for a mean period of 10 years. In 142 patients, the number of typical features was greater than the number of atypical features, whereas in 100 patients the number of atypical features was greater than or equal to the number of typical features. The mean morbidity index remained stable over a period of 20 years in both groups of patients and was not significantly associated with the presence of atypical features, the duration of lithium treatment, the number or frequency of episodes, or latency from the onset of bipolar disorder to the start of lithium treatment.

Conclusion: Our study suggests that long-term response to lithium maintenance treatment is stable both in patients with typical and in patients with atypical features. The predominance of either typical or atypical features did not result in different responses to long-term lithium treatment in this sample of bipolar patients.

Lithium is recommended as a first-line prophylactic treatment for bipolar disorder in all major international guidelines.1–5 However, although the efficacy of lithium prophylaxis in bipolar disorder has been demonstrated in a number of controlled studies,6–7 the effectiveness of long-term lithium treatment in naturalistic settings has been reported to be much lower.8–10 Furthermore, some studies have suggested that the efficacy of lithium prophylaxis diminishes over time in a subsample of patients.11,12 The differences in effectiveness reported for lithium prophylaxis can be attributed to a variety of factors, including a broadening of the diagnostic criteria for
bipolar disorder, changes in the course of illness encountered by clinicians treating patients over the long term, and methodological differences in the studies being compared.16

Many investigators have also discovered an association between atypical features of bipolar disorder and a poor response to lithium.17–21 Patients with symptoms such as mixed states, incomplete remissions, cognitive impairment between episodes, psychiatric comorbidity, and poor social functioning are considered less likely to benefit from lithium treatment.22–26 Conversely, patients with several features of typical bipolar disorder, such as complete remission between episodes, family history of bipolar disorder, and a mania-depression-euthymia episode sequence, are likely to have an excellent response.27,28

The purpose of our study was to evaluate a prospective, multicenter cohort of patients with bipolar disorder to determine whether the long-term effectiveness of lithium prophylaxis remains stable over time and to investigate whether patients with atypical features show a poorer response to long-term lithium treatment compared to patients with typical features.

METHOD

Inclusion Criteria

Patients were selected according to classical criteria for bipolar disorder. When documentation of the first patients started in the 1980s, the International Classification of Diseases, 8th Revision (ICD-8)29 was in use. This diagnostic system was later replaced by the 9th Revision (ICD-9).30 After 1994, all patients were rediagnosed according to DSM-IV. All patients met the DSM-IV criteria31 for bipolar disorder and required long-term treatment, as defined by the presence of at least 1 manic episode or at least 2 overall episodes in the patient’s history. There was no age limitation for study entry.

Patients were followed in the outpatient clinics of 5 participating International Group for the Study of Lithium-Treated Patients (IGSLI) centers (Berlin, Germany; Halifax, Hamilton, and Ottawa, Canada; and Poznan, Poland) from presentation at the clinic until 2004. Founded in the 1980s, these centers adopted a standard research program for patients receiving long-term prophylactic treatment with lithium and other medications for the management of unipolar mood disorder, bipolar mood disorder, or schizoaffective disorder (www.igsli.org).32

Patient Assessment

Patients were evaluated by a psychiatrist during each visit. The evaluation involved (1) a psychiatric assessment, including the patient’s case history, and any clinical or pharmacologic interventions required; (2) an assessment using standard mood rating scales (Bech-Rafaelsen Melancholia and Mania Scales, Hamilton Rating Scale for Depression,35 Young Mania Rating Scale36); (3) a physical examination; and (4) documentation of any adverse events. Serum lithium levels were also obtained. Patients averaged 7 to 8 visits each year, depending on age, comorbidity, and severity of illness. The number of visits per year was higher than in normal outpatient settings, leading to optimal control of lithium long-term medication. During additional, unscheduled visits, patients were able to draw support from psychiatric nurses and social workers. All patients were treated continuously with lithium, and the above-mentioned data were documented at the respective outpatient clinics over a period of at least 1 year. Before enrollment in the prospective cohort, patients were thoroughly informed about the study procedures, the treatment, and any possible side effects. All participants provided written, informed consent.

Onset of bipolar disorder was defined as (1) the first documentation of diagnosis or, if such documentation was lacking, as (2) documentation of the first symptoms clearly related to bipolar disorder.

A recurrent episode was defined as the presence, in a previously remitted patient, of symptoms that required either psychotherapeutic or psychopharmacologic intervention. All recurrences were recorded and graded for severity, polarity, and duration. Remission was defined as the absence of affective symptoms as measured by standard mood rating scales. All data were ascertained prospectively.

The morbidity index was used as the outcome variable. The morbidity index was first introduced by Coppen and Abou-Saleh37 and includes severity and length of episodes. We rated severity in a semiquantitative manner using 3 different degrees: symptoms that do not require therapeutic action are rated as degree 1; symptoms that require psychotherapeutic or psychopharmacologic intervention but are manageable in an outpatient setting are rated as degree 2; and symptoms leading to inpatient treatment are rated as degree 3. We included symptoms of degree 2 and degree 3 in the analysis and calculated the morbidity index using the following formula:

\[
\text{Morbidity index total over a period of 1 year} = \frac{(\text{no. weeks with degree 2}) \times 2 + (\text{no. weeks with degree 3}) \times 3}{52 \text{ weeks}}
\]

For each year, the morbidity index was calculated for all affective episodes (total morbidity index [MItotal]) and separately for depressive episodes (depressive morbidity index [MIdep]) and manic episodes (manic morbidity index [MIman]). Patients had to be sufficiently compliant, which was defined as maintaining serum lithium levels of at least 0.5 mmol/L throughout the documentation period. Any prophylactic medication administered in addition to lithium after 3 months of remission was...
recorded, including data on the type of medication used and the duration of its administration. Additional prophylactic medication included antipsychotics, antidepressants, anticonvulsants, and high-dose thyroxine.

**Atypical Features**

The atypical or typical features present in each patient were scored retrospectively using the scale shown in Table 1. The scale was developed for this particular research question as part of a consensus procedure among IGSLI members and incorporated features of the patients’ psychopathological history and course of illness. Items known to be characteristic of typical or atypical illness were designated as core features. In contrast, optional features were defined as those features for which there was limited evidence of potential relevance for bipolar disorder. Data were obtained from patient records or from interviews with the patients and their relatives, if possible. Items were scored as “yes” or “no.” A total score was achieved by adding the number of affirmative responses. Patients who did not score on a given item (e.g., because they did not show any features) did not count as “gain” or “lose” a point with regard to typicality or atypicality. Similarly, patients who had 3 or 4 episodes per year prior to the index episode also did not “gain” or “lose” a point with regard to typicality or atypicality. The Minnesota Multiphasic Personality Inventory (MMPI) profile,38 which has been shown to be useful for measuring interepisodic remission, was assessed in only one fourth of the patients, and all of these patients were in a euthymic state at the time of assessment. A “normal MMPI profile with subthreshold precarious findings” includes residual abnormities within the normal psychological profile (i.e., all MMPI scales with t scores less than 70). A family history of nonepisodic disorders includes any psychiatric diseases that do not have a clear episodic course, such as anxiety disorders or personality disorders.

Because many patients experience both typical and atypical features, our patients were categorized into 2 groups based only on core features. Group 1 had more typical features than atypical features, and group 2 either had more atypical features than typical features or an equal number of both.

**Statistical Analysis**

Data were analyzed using BMDP Statistical Software, Inc., Release 8.0 (BMDP Statistical Software, Inc.). Unbalanced repeated-measures regression models with structured covariance matrices were used (module 5V in BMDP) to assess the influence of atypical features and duration of treatment on the morbidity index. Separate calculations were made for MItotal, MIdep, and MIman. Maximum likelihood was used to estimate parameters where the expected response values are described as a linear function of the parameters. The main advantage of this approach was that all subjects could be included regardless of their length of treatment. Model selection was based on optimization of the Akaike information criterion.45 Significance of the independent variables was estimated using the Wald test. A 5% level of significance was established with 2-tailed tests. Using the same method, we also examined the impact of the number and frequency of episodes before the start of lithium treatment, as well as of treatment delay, on the morbidity index. In this analysis, the independent variables were modeled as covariates with length of treatment as a repeated measure.

**RESULTS**

A total of 242 patients receiving lithium treatment were followed for a mean (SD) period of 10 (6.4) years (range, 1–20) (Figure 1). Of these patients, 142 showed more typical features than atypical features and 100 showed more atypical features than typical features or an equal number of both. Baseline characteristics, stratified for both groups of patients, are shown in Table 2. The mean (SD) age of our patients at onset of bipolar disorder was 28.6 (10.9) years (range, 11–66), and lithium treatment was initiated with a mean (SD) latency of 9.4 (9.1) years (range, 0–44).

The mean (SD) MItotal for all patients decreased slightly from 0.150 (0.330) (N = 242; range, 0–1.52) in the twentieth year of the observation period, which suggests that the response was stable over the long term. The mean (SD) MItotal decreased in the typical group (Figure 2) from 0.129 (0.316) (N = 142;
The results of the repeated-measures regression found no significant associations between the number of typical and atypical symptoms on the one hand and the MItotal (p = .167), MIdep (p = .198), or MIman (p = .472, Table 3) on the other. Likewise, there was no significant time effect or time-by-group interaction that would indicate a change in the effectiveness of lithium over time for either of the 2 patient groups.

A subgroup analysis looking at the most extreme typical (4 typical symptoms, N = 50) and the most extreme atypical (3 or more atypical symptoms, N = 18) patients showed a similar MItotal (p = .215) and a similar MIdep (p = .909) in both groups. However, the MIman was significantly higher in the atypical group (p = .003).

The proportions of typical and atypical patients remained stable throughout the study, indicating that the attrition rate in both groups remained similar. Finally, we compared patients’ morbidity indices in their last year of treatment.
treatment with their mean morbidity index over the entire treatment period to test whether patients had discontinued their treatment due to a worsening of their symptoms. There was no evidence of this kind of bias in either group for any of the 3 indices.

In the sample of 242 subjects, the 3 morbidity indices were not significantly influenced by the number of episodes before the start of lithium treatment, the frequency of episodes before the start of lithium treatment, or latency between illness onset and the start of lithium treatment.

Atypical patients used significantly more additional prophylactic medication throughout the entire observation period (Mann-Whitney, p = .03). However, using the Wald test, we were unable to identify additional prophylactic medication as an independent variable with significant influence.
DISCUSSION

The affective morbidity indices remained stable over the entire observation period, confirming that the response to long-term lithium treatment was also stable. The number of typical and atypical symptoms did not have any effect on the morbidity index in patients receiving lithium long-term treatment. In addition, there was no association between the morbidity index and the number or frequency of episodes before starting lithium or the latency between onset of illness and starting lithium. When selecting only the extreme ends of the distribution, we obtained a significant difference for MIman, but not for MItotal or MIdep. It should be noted, however, that this was a very small subgroup, with only 12 extreme typical patients and 1 extreme atypical patient at year 20 of the observation period.

There are several reasons why our results differ from those seen in earlier studies that have shown a clear association between atypical features and a poor response to lithium treatment. First, to maintain consistency across different centers, we categorized patients according to their total number of typical or atypical features. However, some investigations of long-term lithium response have demonstrated that the contribution of individual features to explainable variance can differ dramatically, indicating that it might be useful to evaluate features individually or to give different weights to individual features. The atypical features assessed on our scale were not independent of each other and their impact may have varied considerably. In addition, our scale is comparatively crude, as it uses dichotomous rather than continuous variables for scoring.

Furthermore, many of the newer studies perform a survival analysis that uses time to new episode or rehospitalization as the main outcome measure for long-term prophylactic effectiveness. Although this type of survival analysis is well suited to relatively short efficacy trials that measure time until the first event, it is not optimal for long-term maintenance studies because it does not discriminate between different types of response. For instance, it does not afford proper assessment of the course of illness in patients who show substantial clinical improvement but still experience episodes. Given that bipolar disorder is characterized by wide variations in the length and severity of episodes, the morbidity index allows different forms of response and clinical course to be distinguished from one another in a precise manner.

In a retrospective analysis of lithium maintenance treatment in a small subsample of the present study over a maximum of 15 years, Berghöfer et al. found that, whereas the MItotal remained stable over the entire treatment period, the analysis of the absolute number of recurrences produced no conclusive results because of the general shift over the study period from outpatient to inpatient treatment.

The duration of many earlier prospective studies of lithium treatment has been relatively short (i.e., less than 2 years). Only a few studies have had observation periods extending up to 5 years or 7 years. In our study, data were collected for a much longer period, covering a span of up to 20 years. Another reason why our results may differ from those observed in other studies is that the indications for starting lithium prophylaxis have expanded.

Our study only included patients who met both the classical Kraepelinian and the DSM-IV criteria for bipolar disorder.

The findings of several studies are consistent with our results. Using the morbidity index, Berghöfer and
Berghöfer et al.

colleagues reported on long-term response in a subgroup of bipolar patients over maximum time periods of 15 years and 20 years. In both studies, which included subjects from the present investigation, the severity and duration of recurrences remained stable, and even decreased, over the observation periods, albeit in small samples sizes. Two recent reviews also support our finding that the effectiveness of lithium prophylaxis does not diminish over time.

There has been some controversy as to whether the time span between illness onset and the start of prophylactic treatment may influence the response to long-term treatment. Recent studies do not show any association between negative outcome and latency. For this reason, we included latency of prophylactic treatment in our analysis.

Our study has several methodological shortcomings. Severity ratings for episodes or atypical features may have differed at the various centers. With multiple countries and cultures involved, treatment selection may have varied depending on the health care system, the regional facilities available, and individual patient preferences. At the same time, we evaluated the stability of response individually for each patient. As in any long-term study, treatment for up to 20 years inevitably means that patients were seen by a large number of therapists with varying degrees of training. Because many of the subjects experienced several years of illness before the index episode, we were not always able to reconstruct essential atypical features accurately. As a result, atypical features may be underreported in many patients.

The centers that participated in the study were specialized academic outpatient clinics that, for the most part, were treating patients with a poor course of illness or who required an above-average amount of care. Because of this, a selection bias has to be assumed, both for typical and atypical patients. Another reason that patients visit specialized outpatient clinics is simply because lithium treatment is often difficult to manage in general practice. However, the use of additional medication in our sample was quite low (see Table 2), which indicates that patients with a severe course of illness are unlikely to have been overrepresented in our study. Although we controlled for the influence of additional medication in our analysis, the threshold for initiating additional prophylactic treatment may have differed among the centers.

Atypical patients used significantly more additional prophylactic medication. As an independent variable, this result did not significantly influence the morbidity index. Nevertheless, it is difficult to interpret this finding, as it is impossible to tell from these data whether additional medication was effective and, if so, to what extent. The period during which patients used additional medication was rated with a morbidity index of at least 2. Thus, one could argue that the increased use of co-medication in atypical patients is already reflected in the morbidity index. It should be noted that the indications for any co-medication taken by patients in our study were not documented. As a result, it is not possible to say which symptoms or syndromes led to a particular prescription. The use of co-medication was documented only as a way to control for the influence of this co-medication on the long-term effect of lithium.

It should also be pointed out that the morbidity index does not fully reflect the effects and benefits of lithium in individual patients. For example, it is conceivable that one patient may have had a higher morbidity index than another patient during lithium treatment but may nevertheless have experienced a substantially higher reduction in his or her affective morbidity after treatment initiation. To show individual benefits, data on changes in morbidity index would have been helpful. However, it is important to note here that assessing the initial effects of lithium treatment was not the primary focus of our analysis.

The attrition rate in our study increased along with the duration of treatment. We have no data on subjects who dropped out of long-term observation. Our analysis shows that morbidity indices in the year before dropout were not higher than mean morbidity indices in the preceding years. However, it may still be the case that some patients stopped lithium treatment if they felt that their illness was gradually becoming worse, and a transient worsening would not necessarily be reflected by a significant change in morbidity index.

The relevance of the atypical features included in our scale has been the subject of controversy. For example, the lack of full interepisodic remission and lifetime anxiety comorbidity are not unanimously accepted as atypical features that are relevant to prognosis. The same applies to cognitive impairment between episodes, poor social functioning, and a mania-depression-euthymia sequence. Today, the term atypical is frequently used to describe mixed states of bipolar disorder. However, this is not the definition upon which our analysis is based. Instead, we use the term to describe deviations from the classical Kraepelinaian definition of bipolar disorder.

In conclusion, patients who met both the classical ICD-8 and ICD-9 criteria, as well as the DSM-IV criteria, for bipolar disorder benefited from long-term lithium treatment. Affective morbidity was not influenced by the presence of atypical features.

**Drug name:** lithium (Eskalith, Lithobid, and others).

**REFERENCES**

Long-Term Effectiveness of Lithium in Bipolar Disorder