Eur Arch Psychiatr Neurol Sci (1987) 236:349–357

European Archives of Psychiatry and Neurological Sciences

Prediction of Therapeutic Response in Acute Treatment with Antidepressants

Results of an Empirical Study Involving 159 Endogenous Depressive Inpatients

H.J. Möller, G. Fischer, and D. v. Zerssen

The Max-Planck Institute for Psychiatry, Munich, in collaboration with the Psychiatric Clinic of the Technical University Munich, Federal Republic of Germany

Summary. A study was carried out involving 159 endogenous depressive inpatients treated with antidepressants. Using standardized rating instruments, a large set of potential predictor variables was tested. After cross-validation according to the split-half technique, only very few of these proved to be suitable as predictors for the main outcome criteria. These predictors included premorbid maladjustment, neurotic traits of premorbid personality, intensity of depressive-apathetic symptoms at admission, and the self-evaluated mood assessment score 3 weeks after commencement of treatment. This last criterion demonstrated the highest predictive value. Several predictors known from the literature could not be reproduced in this study.

Key words: Endogenous depression – Therapy with antidepressants – Prediction of response to antidepressants

Introduction

The great efficacy of antidepressants in the acute treatment of patients suffering from endogenous depression has been confirmed by many studies (e.g. Morris and Beck 1974; Möller 1985). However, one problem of treatment using antidepressants is that the condition of the individual patient does not always improve during the course of the antidepressant administration (Heimann 1974; Klerman and Cole 1965; Lehmann 1974). Specifically, patients who previously responded poorly to antidepressants administered in ambulatory treatment programs tend to accumulate in inpatient treatment areas. The results of the research into which patients respond well to customary antidepressant dosages and which profit little or hardly from this treatment have been unsatisfactory and often contradictory (Angst 1965; Ananth 1978; Bielski and Friedel 1976; Fähndrich 1983; Helmchen 1974; Levine and Raskin 1974; Philipp et al. 1985; Woggon 1983). The elucidation of this question is especially important with regard to the recognition of poor responders before beginning inpatient therapy, since such recognition could serve as the basis for more expedient therapy strategies, e.g., higher dosage, combination therapy, etc.

The results found in the literature concerning predictors of response to antidepressants seem to concur somewhat in the following points. Those characteristics which appear to affect the course of treatment unfavorably are: unmarried, neurotic personality traits, high number of phases, phase duration of more than 1 year at the start of therapy, poor response to antidepressants during the last phase, weak depressive symptoms, and presence of depressive illusional ideation (Bielski and Friedel 1976; Woggon 1983).

In recent years attempts have been made to investigate intervention-related variables with regard to their predictive significance for the postacute phase of therapy. In this context, early response to antidepressants has been revealed as a promising prognostic possibility (Woggon 1980). On the other hand, approaches based on biological or biochemical characteristics, such as the reaction to sleep deficit, rapid eye movement latency, urinary 3-methoxy-4-hydroxyphenyl-glycol excretion, the dexamethasone suppression test, and skin resistance, have not resulted in predictions of the therapy success with antidepressants which are unequivocally superior to clinical prognostic tools (Beckmann 1978; Fähndrich 1983; Gaertner et al. 1982; Greden et al. 1983; Giedke et al. 1986; Möller et al. 1986).

The variance proportions explained by the individual predictors are for the most part so small that they cannot be utilized for practical prognosis. With few exceptions (Levine and Raskin 1974; Woggon 1983), the possibility of combining predictors as a means of optimizing prognoses has not been investigated, and comprehensive sets of variables have only rarely been analysed with regard to their prognostic significance. Another disadvantage of many studies is the fact that only relatively small patient collectives were investigated; the degree of selection in such samples is reinforced by the inclusion of only patients who could be included in pharmacological studies. Results from studies based on large and less preselected collectives, i.e., results which better reflect the routine care situation and which can be more effectively generalized, are very rare (Downing and Rickels 1972; Report to the Medical Research Council 1965). Cross-validation for testing the stability of those predictors which were presented was almost never carried out (Woggon 1983).

In most of the work referred to above, only general prognostically relevant characteristics of the improvement in acute depressive symptoms under antidepressant therapy are described, without differentiation between predictors of a favorable spontaneous course and predictors of good response to antidepressants (Bielski and Friedel 1976). The majority of research groups employed imipramine or amitryptiline in their investigations; no consistent indications of medication-specific predictors could be formulated (Hollister et al. 1964; Hordern et al. 1963; Sandifer et al. 1965).

Offprint requests to: H.J. Möller

The following study was an attempt to improve the methods of prediction of the response to antidepressants in endogenous depressive patients. The study was based on the Routine Documentation System for the Psychiatric Department of the Max Planck Institute for Psychiatry (Barthelmes and von Zerssen 1978; Möller et al. 1983). To a great extent, this approach may avoid the methodological pitfalls in earlier prognosis studies. For, along with the large case collective, the extensive data banks, and the standardized processing of the majority of the variables under consideration, it is important that the patients involved were not part of a psychopharmacological research project, but rather were integrated into routine clinical procedures. The question to be answered here is: which patients suffering from an acute manifestation of an endogenous depression are not adequately treated by a routine therapy regimen in which antidepressants are administered in customary dosages?

Materials and Methods

The clinical material consisted of 245 consecutively admitted patients having endogenous depression as the primary diagnosis (ICD criteria); these persons were treated as inpatients between 1972 and 1982 in the Max Planck Institute for Psychiatry (MPIP). Complete external evaluation data sets comprising the Routine Documentation existed for all cases. Following a review of the patients' records, 86 individuals were excluded, either because diagnostic uncertainties were still present ($n = 15$), because contrary to medical indication patients were treated with antidepressants for less than 2 weeks or not at all ($n = 24$), or because patients additionally received electroconvulsive therapy ($n = 11$). Ten patients were excluded for various other reasons.

A brief presentation of the distributions of several important characteristics serves to describe this collective more closely. The average age of the patients was 46.5 years, with the range extending from 18 to 75 years. The sex distribution showed the overrepresentation of women relative to men (70% to 30%) which is typical for this illness. According to the social class scoring of Moore and Kleining (1960), 7% of the patients were from the upper and upper middle classes, 56% from the middle class, 35% from the upper lower class and 3% from the lower class. With respect to the ICD diagnostic subgroups, the following distribution resulted: 29% involution depression, 42% monopolar depression, and 29% depressive phase of a bipolar affective psychosis. The average duration of hospitalization was 56 days.

The patients were treated by administering customary and individualized dosages of antidepressant. The medication of first choice was amitryptiline, administered orally in dosages per day of 150 mg. Where response was poor this dosage was increased according to the specific case. For 2 admitting years, the first choice medication was maprotiline or mianserine (Cording-Tömmel and von Zerssen 1982). Although the use of these two substances resulted in certain differences in treatment effectiveness, it appears justified to neglect these differences in the present study. Patients who did not respond adequately after 4 to 6 weeks with the first choice antidepressant received another antidepressant, usually from a different substance class or with a different mechanism of effect. This procedure was repeated in those cases where resistance to therapy persisted, whereby at this stage i.v. administration was also implemented. In addition to the main medication, tranquilizing antidepressants or neuroleptics were administered in low doses according to need (e.g., for sedation or for sleep induction).

A very comprehensive set of potentially prognostically relevant characteristics was included in this investigation. The Routine Documentation supplied the following predictor variables for all patients, so that these did not have to be determined retrospectively:

(a) sociodemographic data
(b) psychopathological symptoms at admission, as judged by the Inpatient Multidimensional Psychiatric Scale – IMPS (Lorr 1974)
(c) self-evaluation of psychopathological symptoms at admission using the Paranoid-Depression Scale (PD-S) and the Adjective Mood Scale (Befindlichkeitsskala, BfS) and their parallel forms (von Zerssen 1976, 1986).
(d) different dimensions of premorbid personality, including cyclothymic disposition, typus melancholicus, and schizoid, oral, anancastic, hysteric, or neurotoid personality structure. Information regarding premorbid personality was obtained for all the personality dimensions under discussion by administering self-evaluation questionnaires, for some dimensions by having relatives complete corresponding questionnaires. The self-evaluation questionnaires instructed the patient to describe him/herself as experienced before the onset of illness (von Zerssen 1980, 1982). The questionnaires were normally completed during the first week after admission.

Extroversion-Neuroticism-Rigidity Scale (Brengelmann and Brengelmann 1960), Anancasm-Hysteria-Orality Scale and its parallel form (AHOS and AHOSS, von Zerssen 1979): in addition to these three dimensions, these scales also measure neurotoid personality structure (neuroticism) as a character trait predisposing to various forms of neuroses. Schizoidism Scale, self-evaluation and external evaluation versions: not only schizoidism, but by means of a supplemental scale certain aspects of the "typus melancholicus" are also measured (see below). Questionnaires for the determination of "typus melancholicus" and "cyclothymic temperament" (von Zerssen 1974, 1982), for which two versions exist: the self-evaluation F-Form and the external evaluation K-Form.

In order to ensure the comprehensiveness and validity of the personality test battery employed, the number of items in the self-evaluation scales presented above was subsequently reduced according to retest reliability criteria (von Zerssen in preparation). The resulting item set was reformulated to produce the dimensions of the "Premorbid Personality Inventory" (PPI) which was then incorporated into the procedure of this study. The dimensions of the PPI are extroversion, frustration tolerance, self-insecurity, schizoidism, and orderliness.
(e) intelligence, tested for orientation purposes at admission by the subtest general knowledge from the Hamburg-Wechsler Intelligence Test (Wechsler 1964).

Retrospectively, the following potential predictor characteristics were taken from the patients' records: a list of historically relevant variables, including broken home, psychosis in the family, number of prior depression phases, total duration of inpatient psychiatric treatment in the 2 years prior to the index admission, total duration of occupational disintegration (disability, unemployment, premature retirement) in the 2 years prior to the index admission, restriction of occupational performance in the year prior to the index admission (Huber et

al. 1979), duration of the last phase before admission, duration of the symptomless interval between the last phase and the previous one, response to antidepressants in the previous phase, psychosocial stressors prior to the initial manifestation, psychosocial stressors prior to the index admission (after DSM III), degree of social adaptation during the year previous to the index admission (after DSM III), acuity of the initial manifestation, acuity of the index manifestation, duration of the index manifestation prior to the start of inpatient treatment, psychopathological characteristics of the index manifestation, e.g., depressive delusional ideation, suicidality, inhibition, agitation, insomnia, weight loss, vitally depressed mood; disturbances in premorbid social adaptation, as determined by the scale of Gittelman and Klein (1969), and disturbances in premorbid social adjustment, as measured by the abridged Phillips Scale (Harris 1975).

Since both of these latter scales refer to definite chronological development stages independently of the initial manifestation of the illness, a modified form of the Phillips Scale was employed to determine premorbid disturbances in the strict sense, i.e., disturbances which already existed prior to the symptoms observed by the patient or his/her relatives (Premorbid Scale). Determination of important variables relating to general prognosis by means of an adapted version of the Strauss-Carpenter Prognosis Scale (Strauss and Carpenter 1974), which was originally conceived for the prognosis of schizophrenic disorders.

Since the patients' records kept by the MPIP were relatively well-structured and for the most part very informative, the retrospective determination of the set of characteristics listed above could be carried out without great difficulty by the examining physician, who was, of course, "blind" to the discharge data (G. Fischer). Noteworthy "gaps" in the data were present only in the scales for determining disturbances in premorbid social adjustment. In this respect, the abridged Phillips Scale after Harris supplied the most data, a circumstance probably related to the fact that this scale was specifically developed for the retrospective evaluation of patients' records.

The data supplied by the IMPS, PD-S, and the premorbid personality scales were not compared by statistical analysis on the item level, since the evaluation of individual symptoms often causes considerable problems regarding reliability. In addition to this, the large number of the ensuing correlations (the IMPS alone has 90 items) would have possibly generated pseudo-significances and thus led to difficulties in interpretation. Therefore, the values of selected item groups were subsumed under syndrome scores in keeping with the evaluation directions stated in the respective manuals. With particular regard to the IMPS data, this meant that not only the original 12 syndrome scores described by Lorr (1974) were considered in the calculations, but also the 5 superfactors under which the original 12 syndromes were subsumed by von Zerssen and Cording (1978). Of these 5 superfactors, however, only the depressive-apathetic syndrome, which combines the depressive syndrome, the apathetic syndrome, and impaired functioning, was used in this study.

Beyond these data gained from the histories and the findings at admission, intervention-related variables, as experienced by the mood assessment scores after 3 weeks and the mood improvement quotient after 3 weeks, were also considered in the analysis of potential predictors.

The formulation of efficacy criteria proceeded according to the principle of multiple outcome measurement, i.e., the fact was taken into account that different, incompletely correlating criteria are necessary in the evaluation of therapy efficacy. Since in the various previous studies more or less different efficacy criteria were employed, for this investigation it seemed expedient to implement a number of efficacy criteria to compare all results adequately. As was the case with the set of predictors, among the efficacy criteria the distinction was made between those criteria derived retrospectively from the records (and which therefore could be expected to be of inferior quality) and those which could be established within the context of routine documentation at the time when the data arose. Criteria of the latter type are not only those derived from the data of the IMPS, the PD-S and the Bf-S – all concerning psychopathology at discharge – but also a global evaluation based on the scale of the Clinical Global Impressions (CGI, Guy 1976), which registers the extent of improvement at discharge; and the mood assessment data at fixed intervals 3 and 6 weeks after admission. Information relating to length of hospitalization was also used in formulating efficacy criteria. Finally, on the basis of data from patients' records, the extent of improvement after 3 weeks was evaluated using the CGI and the time needed to reach a definite improvement was determined.

Following the selection of the relevant dimensions of the IMPS and the self-evaluation scales, 16 efficiency criteria were formulated:

1. IMPS factor depressive syndrome at discharge
2. IMPS superfactor depressive-apathetic syndrome at discharge
3. Improvement quotient for the IMPS factor depressive syndrome[1]
4. Improvement quotient for the IMPS superfactor depressive-apathetic syndrome[1]
5. Depression factor (PD-S) at discharge
6. Improvement quotient for the depression factor (PD-S) at discharge[1]
7. Mood score after 3 weeks
8. Mood score after 6 weeks
9. Improvement quotient for mood score after 3 weeks[1]
10. Improvement quotient for mood score after 6 weeks[1]
11. Mood score at discharge
12. Improvement quotient for mood score at discharge[1]
13. Length of hospitalization
14. CGI of the psychopathological state after 3 weeks
15. Number of inpatient days elapsed until definite improvement
16. CGI of the psychopathological state at discharge.

Thus, 16 efficacy criteria were included in the statistical analysis. However, it was possible to reduce this rather broad-based measurement of efficacy to a few primary efficacy criteria; most appropriate in this regard were the IMPS superfactor depressive-apathetic syndrome and the corresponding improvement quotient, a presumption which was confirmed by the product-moment correlation of all outcome criteria. The efficacy criteria which reflect the condition of the patient at discharge showed the closest correlations with the predictor variables, less close were between the latter and the condition of the patient after 3 or 6 weeks. Although the potential pre-

[1] The improvement quotients were calculated as follows: score upon admission minus score at specific times (3 weeks, 6 weeks, discharge, etc.), divided by the score upon admission.

dictors did not correlate at all with the length of hospitalization, this "hard" criterion was nevertheless chosen as a primary criterion, since among the patients discharged in a sufficiently improved condition there were those who only improved after a relatively lengthy treatment. If merely the psychopathological condition at discharge were to be chosen as an efficacy criterion, then these "poor responders" would be left unconsidered.

In the sense of an orienting analysis, the product-moment correlations between the potential predictor variables and the above efficacy criteria were calculated. Only correlations significant at the level of $P < 0.01$ or better were judged as being relevant for practical prognosis and will be presented here. In order to test the stability of the results, a cross-validation employing the "split-half" technique was carried out. Those predictors were viewed as consistent whose correlation coefficients relative to the respective efficacy criteria in both subcollectives were at the $P < 0.05$ or better significance level.

Results

In comparison to control values from a representative sample of the general population (Wittchen et al. in preparation), 24% of the patients attained pathological scores in the IMPS criterion depressive syndrome and could thus be viewed as poor responders. The following presentation of the predictors is restricted to significant correlations ($P < 0.01$) between predictor variables and the three primary efficacy criteria. The other results will be mentioned only if they proved to be of particular interest. In the tables, only those relationships between predictor variables and efficacy criteria are shown for which at least one significant correlation was present.

The sociodemographic characteristics did not demonstrate noteworthy prognostic relationships in any cases, in particular not to any of the primary efficacy criteria. Likewise, none of the items of the Gittelman-Klein Scale correlated with these criteria; however, almost all of the individual items correlated with an insufficient improvement in the patient's condition as evaluated after 3 weeks, and some items correlated significantly with the degree of mood disturbance after 3 weeks. The items of the Phillips Scale did not correlate with any of the primary efficacy criteria, but they did correlate significantly with the other criteria; in particular, the total score of this scale correlated with the degree of mood disturbance (Table 1). The Premorbid Scale revealed itself to be more important as regards prognostic significance: the total score and the item few social contacts, among others, correlated significantly with the depressive-apathetic syndrome at discharge and with the poor improvement in this syndrome.

Of the numerous premorbid personality dimensions tested in this study, only the characteristic orality (AHOS) correlated significantly with the depressive-apathetic syndrome at the time of discharge, but also with the depressivity factor of the PD-S (Table 1). Several of the other premorbid personality dimensions showed significant relationships to some secondary efficacy criteria, especially to self-evaluated depressivity (PD-S) at discharge, and to the degree of self-evaluated mood disturbance after 3 to 6 weeks, but not, however, to any of the primary outcome criteria. Schizoidism and cyclothymia did not correlate with any of the efficacy criteria. Interestingly enough, there were some significant correlations between the self-evaluated personality traits and the duration of hospitalization; the traits concerned were neurotoid personality structure and typus melancholicus (F-Form). The intelligence quotient, measured in this study by the subtest "general knowledge" of the HAWIE, showed no prognostic value.

The 12 IMPS factors, reflecting the psychopathological symptomatology upon patient admission, also demonstrated few significant correlations, i.e., only for three dimensions and with just one of the primary efficacy criteria, namely with the improvement quotient for the depressive-apathetic syndrome (Table 2). High scores for both the depressive and apathetic syndromes upon admission correlated with good improvement, but the tightest correlation in this regard was between the superfactor depressive-apathetic syndrome upon admission and the improvement quotient for this same syndrome ($r = 0.39$). It is interesting that none of these three characteristics correlated with the self-evaluated efficacy criteria. Conversely, although self-evaluated depressivity (PD-S) and mood disturbance (Bf-S) showed almost no correlation with externally evaluated depressivity (the notable exception being the correlation between mood at admission and the primary criterion improvement in the depressive-apathetic syndrome). There were a number of correlations among the self-evaluated depressivity and mood parameters, to the effect that a high admission score for a particular characteristic was related to a high score for the same characteristic after 3 and/or 6 weeks and at the time of discharge, as well as to a great improvement at the respective time. Self-evaluated depression at the time of admission also correlated significantly with the duration of hospitalization.

On the whole, very close relations existed between the mood score after 3 weeks and various efficacy criteria (Table 2). This Bf-S value correlated significantly with the depressive-apathetic syndrome at the time of discharge, with the depressivity score, with the mood assessment score after 6 weeks and at discharge, with the improvement quotient for mood at discharge, with the length of hospitalization, with the time period elapsing before improvement was registered, with the global state at discharge, and with an insufficiently improved global state at discharge. Thus, this intervention-related variable often achieved much closer correlations to several efficacy criteria than the other potential predictor characteristics examined here. The improvement quotient for mood after 3 weeks achieved a far lesser prognostic significance, especially with respect to the number of significant correlations. However, a close relationship between this quotient and the improvement quotient for the depressive-apathetic syndrome at discharge existed.

As a supplement to the correlations of the IMPS syndromes, several specialized psychopathologic aspects (evaluated retrospectively using data from the records), which are highly regarded in the prognosis literature relating to depressive patients, were investigated with respect to establishing their predictive value. The results showed that for the characteristics depressive ideational illusions, psychomotor inhibition, psychomotor agitation, insomnia, weight loss, depressed feeling, diurnal mood fluctuations, and suicidality no significant correlations existed with the primary efficacy criteria.

In contrast, the characteristics based on the history of illness yielded several correlations, three of which involved the primary efficacy criterion depressive-apathetic syndrome at the time of discharge: low degree of social adjustment in the year prior to the index manifestation, gradual beginning of the index manifestation, and duration of the depressive phase prior

Table 1. Dimensions of premorbid social adjustment and personality as predictors (correlation shown only if $P < 0.01$)

	Depressive-apathetic syndrome (IMPS)	Improvement in depressive-apathetic syndrome	Depressive syndrome (IMPS)	Improvement in depressive syndrome	Depressivity (PD-S)	Improvement in depressivity	Mood disturbances after 3 weeks	Improvement in mood disturbances after 3 weeks	Mood disturbances after 6 weeks	Improvement in mood disturbances after 6 weeks	Mood disturbances at discharge	Improvement in mood disturbances at discharge	Poor therapeutic response after 3 weeks	Days elapsing before initial improvement	Poor therapeutic response at discharge	Length of hospitalization
Few friends (premorbid scale)	0.28 (134)	−0.25 (134)	0.28 (134)													
Score sum (premorbid scale)	0.28* (132)	−0.23 (132)	0.26 (132)	−0.27 (132)									0.23 (132)			
Orality (AHOS)	0.26* (143)		0.22 (143)		0.27 (130)											
Neurotic structure (AHOS)							0.24 (126)									0.24* (126)
Typus melancholicus (F-Bogen)																0.22 (146)

* Statistical significance confirmed by cross-validation

Table 2. Psychopathological symptoms at admission as predictors (correlation only shown if $P < 0.01$)

	Depressive-apathetic syndrome (IMPS)	Improvement in depressive-apathetic syndrome	Depressive syndrome (IMPS)	Improvement in depressive syndrome	Depressivity (PD-S)	Improvement in depressivity	Mood disturbances after 3 weeks	Improvement in mood disturbances after 3 weeks	Mood disturbances after 6 weeks	Improvement in mood disturbances after 6 weeks	Mood disturbances at discharge	Improvement in mood disturbances at discharge	Poor therapeutic response after 3 weeks	Days elapsing before initial improvement	Poor therapeutic response at discharge	Length of hospitalization
Depressive syndrome (IMPS)		0.34 (159)		0.49* (159)												
Apathetic syndrome (IMPS)		0.26* (159)														
Depressive-apathetic syndrome (IMPS)		0.39* (159)		0.41* (159)					0.26 (99)			0.22 (144)				
Depressivity (PD-S')			0.22 (140)		0.33 (126)		0.-30* (120)		0.31 (90)							0.27 (140)
Mood state (Bf-S) at admission		0.26 (144)		0.27 (144)			0.27 (126)	0.28 (105)		0.38 (82)	0.25 (130)		0.24 (144)	0.22 (141)		
Mood state after 3 weeks	0.25* (123)				0.30 (132)			−0.25 (111)	0.62* (96)		0.49* (120)	−0.34 (127)	0.59* (134)	0.50* (129)	0.23 (134)	0.50 (134)
Improvement of mood state after 3 weeks		0.43* (121)			−0.25 (111)					0.91* (85)		0.54* (106)				

* Statistical significance confirmed by cross-validation

Table 3. Characteristics of patient history as predictors (correlation only shown if $P < 0.01$)

	Depressive-apathetic syndrome (IMPS)	Improvement in depressive-apathetic syndrome	Depressive syndrome (IMPS)	Improvement depressive syndrome	Depressivity (PD-S)	Improvement in depressivity	Mood disturbances after 3 weeks	Improvement in mood disturbances after 3 weeks	Mood disturbances after 6 weeks	Improvement in mood disturbances after 6 weeks	Mood disturbances at discharge	Improvement in mood disturbances at discharge	Poor therapeutic response after 3 weeks	Days elapsing before initial improvement	Poor therapeutic response at discharge	Length of hospitalization
Number of hospitalizations							0.32* (133)									0.23 (158)
Low degree of social adjustment in the year prior manifestation	0.22 (143)		0.22 (143)		0.30 (125)	−0.25 (122)									0.31* (143)	
Gradual beginning of index manifestation	0.21 (155)					−0.29 (130)										
Duration of depression prior to start of index treatment	0.22 (156)				0.29* (135)	−0.25 (131)	0.30 (131)		0.27 (97)						0.33* (156)	

* Statistical significance confirmed by cross-validation

to starting the index therapy (Table 3). These characteristics also showed themselves to be prognostically relevant for self-evaluated depressivity at discharge and for the global assessment at discharge. The number of inpatient treatments in the history was merely related to a longer period of hospitalization and to the mood score after 3 weeks.

A poor response to antidepressants during the previous phase correlated merely with the mood assessment score after 3 weeks. A number of the tested characteristics were shown to be prognostically irrelevant; these included mental illness within the family, age at the time of the initial manifestation, duration of the illness, duration of the previous phase, diagnostic subclassification into monopolar or bipolar, psychosocial stressors prior to the index treatment, degree of social adjustment prior to the index treatment, etc.

Along with the total score, the following items of the Strauss-Carpenter Scale correlated significantly with one or more of the primary efficacy criteria (Table 4): occupational incompetence in the year prior to the index admission, few social contacts in the year prior to admission, duration of the longest period of severe psychiatric symptoms, lack of fullness of life in the year prior to admission, and the absence of subjective distress in the month prior to the index admission. All of these characteristics were related to a high score for the criterion depressive-apathetic syndrome at the time of discharge and/or to a slight improvement in this syndrome. A part of these characteristics was related to other efficacy criteria, this being especially the case for the characteristics occupational incompetence in the year prior to admission and lack of fullness of life in the year prior to admission, both of which correlated significantly with the majority of the efficacy criteria. The latter characteristic achieved nearly the same prognostic significance as the total score of the Strauss-Carpenter Scale, which also correlated closely with a majority of the efficacy criteria. The total duration of previous hospitalization correlated most markedly with the length of the index hospitalization and with the conditions at the time of discharge from hospital.

The orienting product-moment correlation analysis revealed but few prognostically relevant relationships between the various predictor variables and the primary efficacy criteria. In the face of the large number of correlations calculated, those few significant correlations which did result must be regarded critically with respect to their possibly being chance phenomena. The results of the cross-validation were even more disappointing, even though the minimum significance level was reduced to $P < 0.05$ (Table 5). If the restriction regarding only significant correlations with the primary efficacy criteria was removed and all prognostic correlations between potential predictor variables and efficacy criteria which have been confirmed by cross-validation were allowed, as expected, a larger number of prognostically relevant characteristics resulted.

Discussion

Despite the great number of correlation calculations – approximately 300 for those involving the three primary criteria alone – most of the significant relationships of the total collective discovered in this study certainly cannot be interpreted as pseudo-significances. At the 1% level of significance underlying this study, only three chance coincidences would

Table 4. Characteristics of the Strauss-Carpenter scale as predictors (correlation shown only if $P < 0.01$)

	Depressive-apathetic syndrome (IMPS)	Improvement in depressive-apathetic syndrome	Depressive syndrome (IMPS)	Improvement in depressive syndrome	Depressivity (PD-S)	Improvement in depressivity	Mood disturbances after 3 weeks	Improvement in mood disturbances after 3 weeks	Mood disturbances after 6 weeks	Improvement in mood disturbances after 6 weeks	Mood disturbances at discharge	Improvement in mood disturbances at discharge	Poor therapeutic response after 3 weeks	Days elapsing before initial improvement	Poor therapeutic response at discharge	Length of hospitalization
Occupational incompetence in the year prior to index admission	0.28 (155)		0.25 (155)		0.23 (133)		0.26 (130)		0.31* (97)				0.31* (155)		0.28* (155)	0.24 (155)
Few social contacts in the year prior to index admission	0.29 (101)	−0.27 (101)	0.26 (101)													
Total duration of previous hospitalization					0.27 (135)		0.31* (131)								0.21 (156)	0.27 (156)
Duration of longest period with severe psychiatric symptoms	0.21 (157)				0.34 (135)	−0.31* (131)					0.26 (136)				0.31 (157)	
Duration of longest period with severe and mild psychiatric symptoms					0.35* (135)	−0.29* (131)									0.29* (157)	
Absence of subjective distress in the month prior to index admission		−0.26 (159)		−0.41 (159)												
Relatively empty life in the year prior to index admission	0.26 (159)		0.29 (159)		0.31 (137)	−0.22 (133)	0.26 (134)				0.22 (138)				0.36* (159)	
Score sum of the Strauss-Carpenter scale	0.31 (87)		0.29 (87)		0.53 (77)	−0.32 (76)					0.30 (80)				0.43 (87)	

* Statistical significance confirmed by cross-validation

Table 5. Confirmed predictors of the primary efficacy criteria: confirmation by cross-validation

1. Disturbed premorbid social adjustment (score sum of the Premorbid Scale)
2. Orality (AHOS)
3. Neurotic structure (AHOS)
4. Apathetic syndrome (IMPS)
5. Superfactor depressive-apathetic syndrome (IMPS)
6. Depressive factor (PD-S)
7. Mood disturbances (Bf-S) after 3 weeks
8. Improvement of mood disturbances (Bf-S) after 3 weeks

be expected. In any case, the results were disappointing, and actually more so in the light of the cross-validation outcome. The majority of the predictors found in the recent literature could not be reproduced in correlations with the primary efficacy criteria; for the remainder, only disturbed premorbid adjustment, oral and neurotoid traits of premorbid personality, the intensity of depressive-apathetic symptoms at the time of admission, and the mood score along with its improvement after 3 weeks of therapy proved to be prognostically significant. The rather critical outcome of this study corresponds to the recently published results of Woggon (1983), who in a study involving 90 patients and employing standardized rating instruments found only four predictors, of which after cross-validation only the severity of depressive symptoms remained prognostically relevant.

Because of the tightness and number of its significant correlations, the mood score after 3 weeks was particularly prognostically relevant: a more serious mood disturbance beyond the 3rd week of therapy was an unfavorable sign for the future course. This intervention-related characteristic frequently achieved correlation coefficients approaching $r = 0.50$, e.g., in its relationship to the length of hospitalization, thus showing itself to be an especially important predictor. This corresponds well to another study by Woggon (1980), who described the degree of depression in the first 10 days of therapy (as defined by the rating scale of the Association for Methodology and Documentation in Psychiatry) as the most important predictor for the further course.

In comparing the present results with those from the literature, the marked dependence of predictor findings on the size and make-up of the patient collective becomes evident (Bielski and Friedel 1976). By carrying out internal cross-validation, a much more demanding standard has been established a priori in the present study and consequently substantially fewer numbers of predictors have been discovered here than in other studies, in which cross-validation was not carried out. If the other efficacy criteria examined in this study are also included, then after cross-validation, a greater number of prognostically relevant characteristics results. The content of a portion of these predictors is reflected in the descriptions of predictive characteristics in the literature, e.g., characteristics of the history of illness which indicate chronicity (Angst 1961, 1965; Deykin and DiMascio 1972; Kiloh et al. 1962; Lesse 1960), or the predictive importance of neurotic personality traits as confirmed here by cross-validation with respect to the primary efficacy criteria (Deykin and DiMascio 1972; Downing and Rickels 1973; Kiloh et al. 1962; Paykel et al. 1973; Raskin and Crook 1976); all of these traits can be regarded as unfavorable predictors. On the other hand, even after the limitations imposed by cross-validation and the restriction of the analysis to include only primary criteria were dropped, several of the predictors described in the literature as being particularly relevant could not be reproduced, which is not at all surprising, considering the instability of the findings presented in the literature.

The prognostic relationships found in this investigation were generally on a very low level of explained variance (usually under 10%), so that only group statistic prognostic differentiations and no individual prognoses may be formulated.

An inherent problem of this study was that the primary efficacy criteria could not be established for a fixed time interval relative to the start of therapy; rather, the time interval varied according to the time of individual patient discharge, which besides being variable in itself is also influenced by many other factors, not the least of which is the improvement in or the reduction of the number of symptoms. However, as was shown by the congruency of several of the predictors from the present study with those from other investigations as well as by the comparison of efficacy criteria within this study, it would appear that this problem does not seriously distort the results regarding the predictors. In addition, the procedure in this form does full justice to the demand for predictors of inadequate therapy results under routine treatment conditions, thus opening up the possibility of adjusting the duration of treatment to the individual circumstances.

It can be regarded as a methodological failing of this study that a portion of the predictor variables and efficacy criteria were established retrospectively from patients' records. However, the definite advantages offered by the present method in comparison to other investigations dealing with predictors should not be overlooked: the relatively high case number in the collective examined, the multitude of predictor characteristics analyzed, the largely standardized evaluation of characteristics, the multiple evaluation of efficacy by means of various criteria, and the cross-validation of the univariate analysis results.

References

Ananth J (1978) Clinical prediction of antidepressant response. Int Pharmacopsychiatr 13:69–93

Angst J (1961) A clinical analysis of the effects of Tofranil in depression. Longitudinal and follow-up studies. Treatment of blood relations. Psychopharmakologia 2:381–407

Angst J (1965) Zur Prognose antidepressiver Behandlungen. Anglo-Germ Med Rev 2:733–751

Barthelmes H, Zerssen D v (1978) Das Münchner psychiatrische Informationssystem (PSYCHIS München). In: Reichertz P, Schwarz E (Hrsg) Informationssysteme in der medizinischen Versorgung, Ökologie der Systeme. Schattauer, Stuttgart, pp 138–145

Beckmann H (1978) Biochemische Grundlagen der engogenen Depression. Nervenarzt 49:557–568

Beckmann H (1981) Die medikamentöse Therapie der Depressionen. Nervenarzt 52:135–146

Bielski RJ, Friedel O (1976) Prediction of tricyclic antidepressant response. A critical review. Arch Gen Psychiatry 33:1479–1489

Brengelmann JC, Brengelmann L (1960) Deutsche Validierung von Fragebögen der Extraversion, neurotischen Tendenz und Rigidität. Z Exp Angew Psychol 7:291–331

Cording-Tömmel C, Zerssen D v (1982) Mianserin and Maprotilin as compared to amitriptyline in severe endogenous depression. A new methodological approach to the clinic evaluation of the efficacy of antidepressants. Pharmacopsychiatry 15:197–204

Deykin EY, DiMascio A (1972) Relationship of patient background characteristics to efficacy of pharmacotherapy in depression. J Nerv Ment Dis 155:209–215

Downing RW, Rickels K (1972) Predictors of amitriptyline response in outpatient depressives. J Nerv Ment Dis 154:248–263

Downing RW, Rickels K (1973) Predictors of response to amitriptyline and placebo in three outpatient treatment settings. J Nerv Ment Dis 156:109–129

Fähndrich E (1983) Clinical and biological parameters as predictors for antidepressant drug responses in depressed patients. Pharmacopsychiatry 16:179–185

Gaertner HJ, Kreuter F, Scharek G, Golfinopoulos G, Breyer-Pfaff U (1982) Zur klinischen Bedeutung einer biochemischen Differenzierung bei depressiven Erkrankungen. In: Beckmann H (Hrsg) Biologische Psychiatrie. Fortschritte psychiatrischer Forschung. Thieme, Stuttgart New York, pp 259–270

Giedke H, Axmann D, Gaertner HJ, Rein W, Rötzer-Zimmer FT (1986) Psychopathological predictors of antidepressant therapy response. Pharmacopsychatry 19:259–262

Gittelmann-Klein R, Klein D (1969) Premorbid asocial adjustment and prognosis in schizophrenia. J Psychiatr Res 7:35–53

Greden Jf, Gardner R, King D, Grunhaus L, Carroll BJ, Kronfol Z (1983) Dexamethasone suppression test in antidepressant treatment of melancholia. The process of normalization and test-retest reproducibility. Arch Gen Psychiatry 40:493

Guy W (1976) Clinical Global Impressions (CGI). ECDEU Assessment, manual for psychopharmacology, rev edn. Rockville, Maryland, pp 217–222

Harris JG (1975) An abbreviated form of the Phillips Rating Scale of premorbid adjustment in schizophrenia. J Abnorm Psychol 84: 129–137

Heimann H (1974) Therapy-resistant depressions: Symptoms and syndromes. Contributions to symptomatology and syndromes. Pharmakopsychiatr 7:139–144

Helmchen H (1974) Symptomatology of therapy-resistant depressions. Pharmakopsychiatr 7:145–155

Hollister LE, Overall JE, Johnson M, Penninton V, Katz G, Shelton J (1964) Controlled comparison of amitriptyline, imipramine and placebo in hospitalized depressed patients. J Nerv Ment Dis 139: 370–375

Hordern A, Holt NF, Burt CG, Gordon WF (1963) Amitriptyline in depressive states. Phenomenology and prognostic considerations. Br J Psychiatry 109:815–825

Huber G, Gross G, Schüttler R (1979) Schizophrenie. Eine verlaufs- und sozialpsychiatrische Langzeitstudie. Springer, Berlin Heidelberg New York

Kiloh LG, Ball JR, Garside RF (1962) Prognostic factors in treatment of depressive states with imipramine. Br Med J 2:1225–1227

Klerman GL, Cole JO (1965) Clinical pharmacology of imipramine and related antidepressant compounds. Pharmacol Rev 17:101–141

Kupfer DJ, Spiker DG (1981) Refractory depression: Prediction of non-response by clinical indicators. J Clin Psychiatry 42:307–311

Lehmann HE (1974) Therapy-resistant depressions – a clinical classification. Pharmakopsychiatry 7:156–163

Lesse S (1960) The evaluation of imipramine hydrochloride in the ambulatory treatment of depressed patients. J Neuropsychiatry 1:246–252

Levine J, Raskin A (1974) Predicting treatment responsiveness-resistivness in a population of depressed patients. Pharmakopsychiatry 7:217–222

Lorr M (1974) Assessing psychotic behavior by the IMPS. In: Pichot P, Olivier R (eds) Psychological measurement in psychopharmacology. Modern problems in pharmacopsychiatry. Karger, Basel

Möller HJ (1985) Kontrollierte Untersuchungen zum Wirkungsnachweis von Amitriptylin unter besonderer Berücksichtigung des Stellenwertes von Amitriptylin gegenüber den neuen Antidepressiva. Literaturübersicht und Analyse methodischer Probleme. In: Beckmann H, Sieberns (Hrsg) Das ärztliche Gespräch. 38. pmi Verlag, Frankfurt, pp 135–147

Möller HJ, Barthelmes H, Zerssen D von (1983) Forschungsmöglichkeiten auf der Grundlage einer routinemäßig durchgeführten Basis- und Befunddokumentation. Psychiatria Clin 16:1–16

Möller HJ, Kissling W, Bottermann P (1986) the dexamethasone suppression test in depressive and schizophrenic patients under controlled treatment conditions. Eur Arch Psychiatr Neurol Sci 235:263–268

Moore H, Kleining G (1960) Das soziale Selbstbild der Gesellschaftsschichten in Deutschland. Kölner Z Soziol Sozialpsychol 12:86

Morris JB, Beck AT (1974) The efficacy of antidepressant drugs. A review of research (1958–1972). Arch Gen Psychiatry 30:667–674

Paykel ES, Prusoff BA, Klerman GI, Haskell D, DiMascio A (1973) Clinical responses to amitriptyline among depressed women. J Nerv Ment Dis 156:149–165

Philipp M, Beck V, Glocke M, Metz K, Scherhag R, Schmidt R (1985) Vorhersagbarkeit des Therapieansprechens depressiver Patienten auf Doxepin. In: Philipp M (Hrsg) Grundlagen und Erfolgsvorhersage der ambulanten Therapie mit Antidepressiva. Springer, Berlin Heidelberg New York Tokyo, pp 29–45

Raskin A, Crook TH (1976) The endogenous-neurotic distinction as a predictor of response to antidepressant drugs. Psychol Med 6:59

Report to Medical Research Council by its Clinical Psychiatry Committee (1965) Clinical trial of the treatment of depressive illness. Br Med J 1:881

Sandifer MG, Wilson IC, Gambill JM (1965) The influence of case selection and dosage in an antidepressant drug trial. Br J Psychiatry 111:142–148

Strauss JS, Carpenter WT (1974) Prediction of outcome in schizophrenia. II: Relationship between predictor and outcome variables. Arch Gen Psychiatry 31:37–42

Wechsler D (1964) Die Messung der Intelligenz Erwachsener. Huber, Bern Stuttgart Wien

Woggon B (1980) Veränderungen der psychopathologischen Symptomatik während 20tägiger antidepressiver oder neuroleptischer Behandlung. Psychiatria Clin 13:150–164

Woggon B (1983) Prognose der Pharmakotherapie. Klinische Untersuchung zur Voraussagbarkeit des Kurzzeittherapierfolges von Neuroleptika und Antidepressiva. Forum der Psychiatrie 16, Enke, Stuttgart

Zerssen D von (1979) Klinisch-psychiatrische Selbstbeurteilungs-Fragebögen. In: Baumann U, Berbalk H, Seidenstücker G (Hrsg) Klinische Psychologie. Trends in Forschung und Praxis, Bd II. Huber, Bern

Zerssen D von (1980) Psychopathometrische Verfahren und ihre Anwendung in der Psychiatrie. In: Peters (Hrsg) Psychologie des 20. Jahrhunderts, Bd X. Kindler, Zürich

Zerssen D von (1982) Personality and affective disorders. In: Paykel ES (ed) Handbook of affective disorders. Churchill Livingston, Edinburgh London Melbourne New York, pp 212–228

Zerssen D von (1986) Clinical self-rating scales (CSRS) of the Munich psychiatric information system (PSYCHIS München). In: Sartorius N, Ban TA (eds) Assessment of depression. Springer, Berlin Heidelberg New York, pp 270–303

Zerssen D von, Cording C (1978) The measurement of change in endogenous affective disorders. Arch Psychiatr Nervenkr 226:95–112

Zerssen D von, Koeller D-M (1976) Klinische Selbstbeurteilungs-Skalen (KSb-S) aus dem Münchener Psychiatrischen Informationssystem (PSYCHIS München). Manuale. a) Allgemeiner Teil; b) Beschwerden-Liste; c) Paranoid-Depressivitäts-Skala; d) Befindlichkeits-Skala. Beltz-Test, Weinheim

Received April 18, 1987