Eur Arch Psychiatr Neurol Sci (1985) 235:65–70

European Archives of Psychiatry and Neurological Sciences

Self-rating Data as a Selecting Factor in Clinical Trials of Psychotropic Drugs

U. Baumann, F. Eckmann, and R. D. Stieglitz

Institut für Psychologie der Universität Salzburg, Akademiestrasse 22, A-5020 Salzburg, Austria

Summary. In psychiatry the relationship between rating by others and self-rating has been discussed again and again. The question has been raised as to the influence of complete data material (all patients with self-rating and rating by others) on results, in comparison to results where patients without self-rating are not taken into evaluation. This question was answered by conducting a double-blind study which included two pharmacologically active substances (amitriptylinoxide and trazodone) and two sub-samples of patients (patients with rating by others and self-rating, and patients with rating by others alone). The trial was performed in a total of 57 in-patients suffering from endogenous depression (ICD 296.1) who received drug treatment for 21 days. The effect was assessed multi-methodically. While no differences between the two drugs were seen in the sub-sample "self-rating and rating by others", amitriptylinoxide proved to be superior to trazodone in the sub-sample "rating by others". On the basis of these findings the use of rating by others and self-rating procedures will be considered.

Key words: Rating by others – Self-rating – Anti-depressant drugs (amitriptylinoxide, trazodone) – Methodology

1. Introduction

It is generally agreed that human behaviour and experience have to be recorded multi-methodically (Seidenstücker and Baumann 1978). Thus, the following aspects are distinguished:

— Data basis: basic unit of consideration (biochemical, physiological, psychological, sociological, ecological perspective).
— Source of data: data giver (patient, therapist, nursing staff, reference person, neutral observer, etc.).
— Functional range: partial aspect within a data basis (i.e. experience, behaviour, working capacity).

Due to these considerations, a multi-method approach for the evaluation of psychotropic drugs and psychotherapy research is generally required (Busch and Müller-Oerlinghausen 1978; Waskow and Parloff 1975; Kazdin 1980) to cope with the complexity of the phenomena studied.

The question to what degree the different data bases and data givers cohere, has found special interest. If different parameters per time are identical, this is called concordance, if they are not, this is called discordance. If the course of different parameters is parallel over the time, we have synchronicity, if not, we have desynchronicity (Baumann and Seidenstücker 1977).

In psychiatry, where most of the time there are different data givers for the psychological data basis (the patient on the one hand and the physician/psychiatrist on the other as the most important sources) the question of rating by others and self-rating has been raised again and again. It may be answered in many ways (Debus 1974; AMDP 1983):

— Assessment of the same facts by the same method by patient and rater (Shapiro and Post 1974).
— Assessment of the same facts by different methods (i.e. self-rating by items and resulting scale value, rating by others by global assessment of the scale value, Kass et al. (1983).
— Structural comparison of data obtained from rating by others and self-rating (i.e. by factor analysis, von Zerssen and Möller 1980), or more complex procedures by the multi-trait multi-method approach (Campbell and Fiske 1959; Wiggins 1973).
— Comparison of the statements obtained from rating by others and self-rating data (i.e. which statistical method can better identify the difference between the two substances and show sensitivity where implicitly a real difference is taken for granted; McNair 1974; Kieback 1982; Edwards et al. 1984; Simhandl et al. 1984).

Due to the basic considerations, different interpretation problems result from the multi-methodical approach (Peterson 1983), and elaborate methodical studies are necessary for an analysis of the exact inter-relations, simple correlation analyses alone are not sufficient. Each discordance and each desynchronicity needs a theoretical explanation (Fahrenberg 1984).

Regarding the question of the relationship between self-rating and rating by others, analyses are always based on complete data material, i.e. data obtained from rating by others and self-rating of all patients when available. In psychiatric research we often have only incomplete data, self-rating data are not always available due to the illness of the patient. Baumann et al. (1974) stated that in a mixed psychiatric sample 33% of 123 investigated patients did not complete the trial after a second self-rating had been demanded. In this case it must be considered that patients had already been preselected with regard to self-rating. There may be fewer dropouts if by comprehensive dialogue the therapist succeeded in improving the motivation for self-rating (von Zerssen 1976).

Offprint requests to: U. Baumann

Provided that in stationary psychiatry data received from rating by others can always be obtained, and self-rating data can only partly be obtained, then the relationship between rating by others and self-rating must additionally be clarified with regard to selection effects. The question arises as to the influence of the complete data material (all patients with self-rating and rating by others) on the results, since this is only possible by patient selection, i.e. if patients without self-rating data are not taken into the evaluation.

In the following, this question was examined by a clinical study of anti-depressant drugs. It will test how far the statements regarding the drugs are comparable in two sub-samples, one of them containing data obtained from self-rating and rating by others (SR), the other data obtained only from rating by others (R).

2. Methods

Selection of Patients

For this trial in-patients of the Landeskrankenhaus Schleswig (State Hospital Schleswig, West Germany) were chosen who were admitted to hospital because of a diagnosed acute depressive state and who seemed to need anti-depressant drug medication. Additional criteria for inclusion were: age (16–65 years), intelligence level (filling in questionnaires should be possible in principle). Criteria for exclusion were pregnancy and previous agranulocytosis.

Medication

The two anti-depressant drugs used were amitriptylinoxide and trazodone. Amitriptylinoxide belongs to the group of tricyclic anti-depressants which chemically results from a modification of amitriptyline by incorporating an oxygen atom in semipolar binding to the nitrogen atom of the side chain. In pharmacological studies amitriptylinoxide has shown an equal anti-depressant efficacy with a significantly less marked peripheral anti-cholinergic and cardio-depressive effect than amitriptyline (Wenzl 1978; Leuschner et al. 1978; Greeff and Köhler 1977). By means of receptor binding studies it has been proved that amitriptylinoxide reveals a significantly lower affinity for the various receptors which are responsible for the peripheral anti-cholinergic and cardio-toxic effects (Borbe and Müller 1984). Clinical use has shown that the anti-depressive effect of amitriptylinoxide equals that of amitriptyline but is faster acting and superior regarding cardiac and vegetative tolerance (Rapp 1978; Seefried 1982; Borromei 1982; Cassano et al. 1983; Platz and Bartsch 1984).

Chemically, trazodone is a completely new substance which does not belong to the tricyclic or to the tetracyclic anti-depressants. This is reflected in pharmacological studies, as trazodone is characterized by an absence of reserpine antagonism and lack anti-cholinergic effects (Benkert and Hippius 1980). Contrary to the tricyclic anti-depressants, trazodone shows no cardio-toxic effects (Cassano et al. 1982). In the treatment of depressive patients it has proved to be of equal therapeutic efficacy as amitriptyline and imipramine with a better vegetative tolerance (Cassano et al. 1982; Kerr et al. 1984; Costa and Racagni 1982).

Both drugs were given at daily doses of 150 mg. The trial was performed double-blind. Pre-treatment with other anti-depressants was stopped 3 days before starting the trial. No concomitant medication with other anti-depressants was allowed. Anti-depressant drugs such as neuroleptics and tranquilizers were only permitted in justified exceptions.

Table 1. Time screen of the trial

	Procedure	Source of data	Trial day				
			0	4	7	14	21
Psychological data	Hamilton Depression Scale	Physician	×	×	×	×	×
	AMDP 5		×	×	×	×	×
	NOSIE	Nursing staff	×	–	×	×	×
	Bf-S	Patients	×	×	×	×	×
	B-L, B-L'		×	–	×	×	×
	Addition of numbers		×	–	×	×	×
	D-S		×	–	–	–	×
	Biological parameters		×	–	–	–	×

Assessment

Based on a multi-methodic approach, the following parameters were chosen:

(1) biological data: usual laboratory tests in haematology and biochemistry, ECG.

(2) Psychological data: (which are ranked first in our trial)

— Rating by the physician: depression by Hamilton Depression Scale (if not mentioned otherwise: version CIPS 1981) (Scale value basing on 17 items), side-effects by AMDP system (Part 5: somatic signs: 40 symptoms).

— Rating by nursing staff: NOSIE (7 scale values).

— Self-rating by the patient: "Befindlichkeitsskala" by von Zerssen (Bf-S, 1 scale), "Beschwerdenliste" by von Zerssen (BL, halved scale value of B-L and B-L'), Depression Scale by von Zerssen (D-S, 1 scale).

— Patient's ability: adding-up of numbers (maximally 90 additions of 3 one- to two-digit numbers for 5 min (Fahrenberg et al. 1977) with reference numbers of the fields dealt with and percentage of error (referred to field number).

Time Screen

The choice of investigation times was determined by considerations regarding object and organisation of the trial (see Table 1).

Statistical Analysis

The trial represents a two-factorial trial design with the factor medication (2 levels) and the factor time (2 respectively, 4 or 5 levels) with repeated measures. A closed evaluation for the whole trial was performed together with an open evaluation per time. For the closed evaluation the rank variance analysis with test repetition for factorial trial designs (Lienert 1978, p. 1036, p. 1051) employing the χ^2 test were chosen. In order to obtain the necessary equal sample sizes on the one hand and not to eliminate data on the other, two overlapping subsamples per evaluation step were built of each sample, i.e. amitriptylinoxide: 3 + 24 + 3 patients versus trazodone: 27 patients. The open evaluation was performed using distribution-free tests for ordinal data (rank variance analyses and *U*-tests).

Table 2. Details on sub-samples

		Amitriptylinoxide	Trazodone
All tests evaluated (SR)	M:F age $\bar{x}$, ± SD	10:10 ($n = 20$) 33.0, 8.5	6:8 ($n = 14$) 33.1, 7.1
Only data obtained from rating by others (R)	M:F age $\bar{x}$, ± SD	9:1 ($n = 10$) 41.7, 11.3	9:4 ($n = 13$) 40.3, 12.5
Total	M:F age $\bar{x}$, ± SD	19:11 ($n = 30$) 35.9, 10.2	15:12 ($n = 27$) 36.6, 10.5

$\bar{x}$ = arithmetic mean
± SD = single standard deviation of the mean

3. *Details on Samples*

Out of 60 patients who were scheduled for the trial, 3 left prematurely (discharge against medical advice, attempted suicides, all 3 patients belonged to the trazodone group). Only part of the remaining 57 patients were able to fill in the self-ratings due to the severity of their illness (particularly at the initial stage). Thus, the total sample divides into four sub-samples (see Table 2): amitriptylinoxide ($n = 30$): 20 with SR, 10 with R, and trazodone ($n = 27$): 14 with SR, 13 with R.

There was an overall drop-out rate of 40% which might have been lower if extremely intensive help had been given in filling in the questionnaires which was not possible due to lack of staff.

In spite of the numerical tendency to a greater number of drop-outs in the trazodone group (prematurely stopped, only ratings by others evaluable) the differences were not significant (χ^2 tests, $P > 0.10$). Regarding age and sex, the correspondingly adjoined sub-samples were comparable (amitriptylinoxide versus trazodone: total, SR, R). The two sub-samples SR and R only differ in age, as the average age of patients in R was higher than that in SR (t-test, $P < 0.10$).

All patients belonged to the group of endogenous depressions which to date have only shown a unipolar course (ICD 296.1; ICD 9th Rev.). Most of the patients showed psychomotor retardation (46 of 57), the remaining 11 a psychomotor agitation.

4. Results

4.1. *Sub-sample SR*

All self-rating data ("Befindlichkeitsskala" Bf-S, "Beschwerdenliste" BL, Depression Scale D-S) showed a significant time effect (rank variance analysis, $P < 0.05$) with constantly decreasing values, which means that they show an improvement. No significant results were found with either the factor medication or interaction, although numerically more patients on amitriptylinoxide revealed a marked improvement than on trazodone (see Table 3).

Similar results were obtained regarding data on patients' ability as the capacity increases in volume while the error percentage remains unchanged.

On the basis of the self-rating data, we can draw the following conclusions from medication: both drugs led to a significant improvement, and there was no difference between the two drugs (main effect, interaction), apart from the fact that amitriptylinoxide proved to be slightly superior numerically.

Table 3. Self-rating data, sub-sample SR (amitriptylinoxide $n = 20$, trazodone $n = 14$)

			d_0	d_4	d_7	d_{14}	d_{21}
Bf-S	Amitriptylinoxide	$\bar{x}$	52.2	47.8	19.7	8.6	6.1
		± SD	4.7	15.2	22.1	18.5	15.2
	Trazodone	$\bar{x}$	50.5	42.0	28.1	11.9	10.4
		± SD	4.3	20.5	21.6	19.5	19.8
$\frac{\text{B-L + B-L'}}{2}$	Amitriptylinoxide	$\bar{x}$	33.9		11.8	5.6	3.8
		± SD	5.1		11.7	9.7	6.9
	Trazodone	$\bar{x}$	33.6		16.6	9.8	8.6
		± SD	7.4		14.3	14.2	13.0
D-S	Amitriptylinoxide	$\bar{x}$	43.0				4.8
		± SD	4.5				10.7
	Trazodone	$\bar{x}$	43.4				10.2
		± SD	6.3				16.4

Table 4. Data obtained from rating by others: Hamilton Depression Scale

		d_0	d_4	d_7	d_{14}	d_{21}
Sub-sample SR						
Amitriptylinoxide	$\bar{x}$	33.0	29.3	16.0	7.2	4.5
	± SD	4.2	8.3	9.0	8.0	7.4
Trazodone	$\bar{x}$	34.1	26.0	16.8	8.6	6.6
	± SD	5.9	9.8	7.9	7.3	10.0
Sub-sample R						
Amitriptylinoxide	$\bar{x}$	42.8	38.1	29.3	17.1	10.6
	± SD	4.1	3.8	7.5	11.4	8.7
Trazodone	$\bar{x}$	39.4	36.7	34.5	28.8	25.9
	± SD	4.3	6.9	6.3	8.5	9.8

Data obtained from rating by others showed a similar picture. In the Hamilton Depression Scale (Table 4, Fig. 1) we found a constant decrease in score values which confirmed the improvement found in self-rating data. Besides the time effect no medication effect or interaction occurred. NOSIE data (see Figs. 2–5) confirmed the improvement in all 7 scales (time effect with $P < 0.01$ each): increase in social competence (1), social interest (2), personal neatness (3), decrease in irritability (4), manifest psychosis (5), retardation (6), depression (7). Although there were significant interactions in scales 2, 3, 4, 5 and 7, no significant differences between the two medication groups were observed in U-tests calculated per trial day (except in scale 2, where on day 21 the amitriptylinoxide group showed a clearer improvement than trazodone).

In summary, interpretations on the basis of rating by others are as follows: both drugs led to a significant improvement, and the two drugs did not differ in their main effect. Regarding interactions (NOSIE scales) we have numerically less marked but significant results. An exact analysis of the course values seems to given evidence of a superiority of amitriptylinoxide. Both self rating and rating by others led to identical statements.

4.2. *Sub-sample R*

On analysing patients' data obtained from rating by others only, results were significantly more marked in spite of smaller sub-samples, contrary to paragraph 4.1.

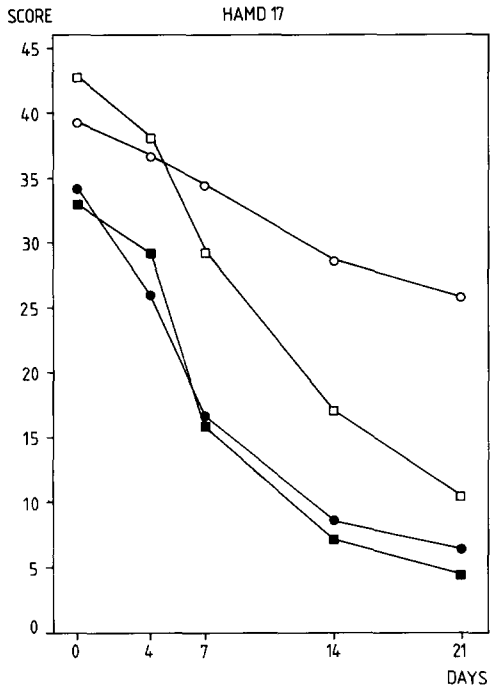

Fig. 1. Hamilton Depression Scale (HAMD): Comparison of sub-samples R and SR for both medication groups (arithmetic mean). Rating by others R: amitriptylinoxide □; trazodone ○. Self-rating and rating by others SR: amitriptylinoxide ■; trazodone ●

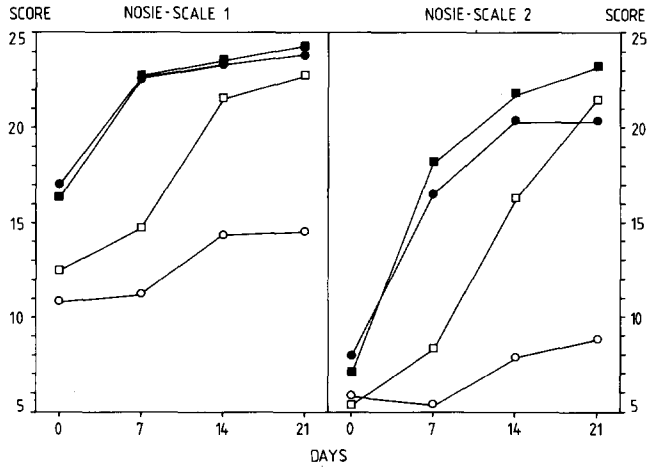

Fig. 2. NOSIE Scale 1 and 2 (social competence and social interest). Comparison of sub-samples R and SR for both medication groups (arithmetic mean). Symbols see Fig. 1

In the closed evaluation of Hamilton scores a time effect was found, but in evaluations per trial day it became evident (Table 4 and Fig. 1) that the amitriptylinoxide group exerted a clearer improvement than the trazodone group (note days 14 and 21, difference $P < 0.01$).

NOSIE scales showed similar results: a time effect was observed in all 7 scales ($P < 0.01$) and the interaction of scales 1, 2, 4, 5, 6, 7 ($P < 0.01$) became manifest in marked differences per trial day (Fig. 2–5).

The amelioration of social competence (1) and social interest (2) was much stronger in the amitriptylinoxide group than in the trazodone group. Particularly striking was the strong decrease in scale values irritability (4), manifest psy-

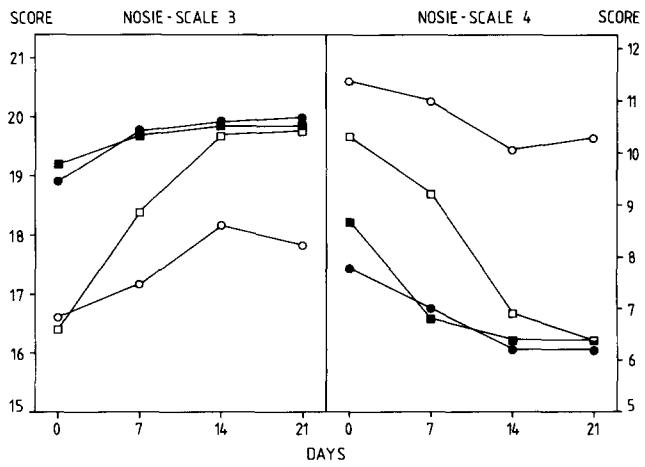

Fig. 3. NOSIE Scale 3 and 4 (personal neatness and irritability). Comparison of sub-samples R and SR for both medication groups (arithmetic mean). Symbols see Fig. 1

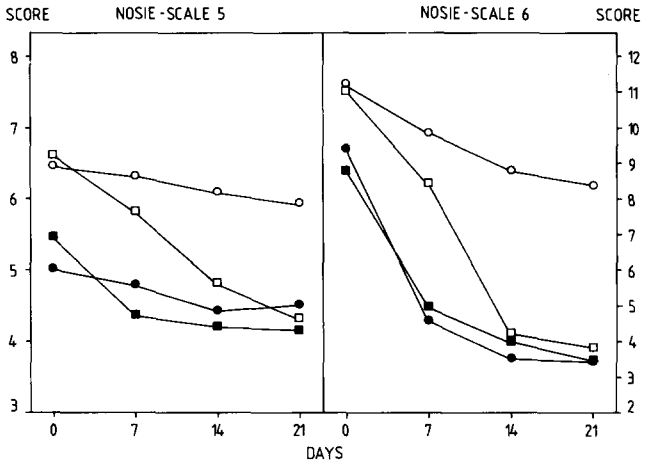

Fig. 4. NOSIE Scale 5 and 6 (manifest psychosis and retardation). Comparison of sub-samples R and SR for both medication groups (arithmetic mean). Symbols see Fig. 1

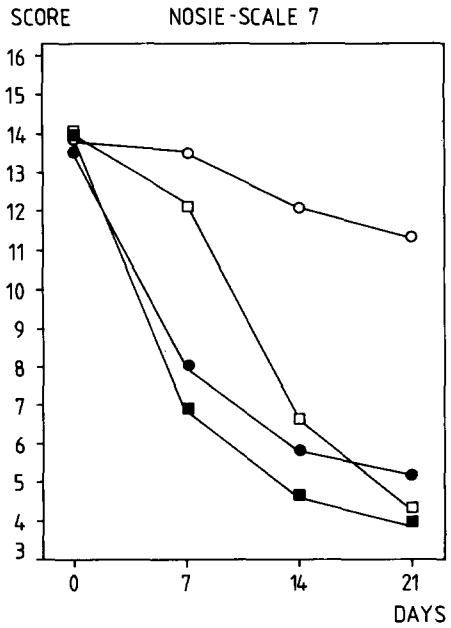

Fig. 5. NOSIE Scale 7 (depression). Comparison of sub-samples R and SR for both medication groups (arithmetic mean). Symbols see Fig. 1

Table 5. Side-effects (total sample) AMDP 5 (40 somatic symptoms)

Comparison	Amitriptylinoxide			Trazodone		
	+	−	0	+	−	0
0/4	16%	4%	80%	14%	8%	78%
0/7	29%	2%	79%	23%	4%	73%
0/14	34%	1%	65%	27%	4%	69%
0/21	37%	1%	62%	27%	3%	70%
Total	29%	2%	69%	23%	5%	72%

+ amelioration of symptoms compared to day 0
− deterioration of symptoms compared to day 0
0 symptoms unchanged (never present or equally strong)

choses (5), retardation (6), depression (7) in the amitriptylinoxide group compared to the trazodone group.

Summing up, interpretations on the basis of data obtained from rating by others are as follows: both drugs led to an improvement, and the improvement was more marked in the amitriptylinoxide group than in the trazodone group.

4.3. Total sample (only data obtained from rating by others)

Without considering the moderator variable self-rating, an evaluation for the total sample would have been performed as usual. The sample size for rating by others would then have been $n = 30$ (amitriptylinoxide) and $n = 27$ (trazodone) and for self-ratings ($n = 20$ amitriptylinoxide, $n = 10$ trazodone). In ratings by others the results would have been similar to those described in paragraph 4.2, i.e. the differences would not have been blurred by combining the two sub-samples. Results of sub-sample R were highly marked (mostly $P < 0.01$ compared per trial day) but less marked in the total sample (in spite of larger n, mostly $P < 0.05$).

Regarding side-effects (somatic symptoms according to AMDP 5) the difference can also be seen, if on the single trail days, the patients with deterioration or improvement of symptoms are counted (with reference to initial state; Table 5). In the trazodone group there were relatively more deteriorations and less improvements compared to the initial state. Relatively fewer deteriorations with amitriptylinoxide compared to trazodone were found regarding the symptoms: disturbances of falling asleep, decreased appetite, dry mouth, nausea, gastric discomfort, dizziness, palpitations, increased perspiration, seborrhoea, menstrual disturbances, backache, heaviness in the legs, hot flushes, conversion symptoms.

5. Discussion

Statements regarding drugs and the relation "rating by others: self-rating" differ according to origin.

A. Starting basis total sample (without considering selection criteria in self-rating data)
(1) the two drugs are different (better effect of amitriptylinoxide than trazodone), where most statements are made on the 5% level.
(2) While ratings by others show differences, self-ratings do not. Data obtained from ratings by others are more sensitive.

B. Starting basis sub-sample SR
(1) The two drugs appear to be similar apart from the slight numerical superiority of amitriptylinoxide.
(2) Data obtained from self-rating and rating by others show similar results, they are comparable in their conciseness of statements.

C. Starting basis sub-sample R
(1) The two drugs are significantly different (better effect of amitriptylinoxide than of trazodone, $P < 0.01$). Considering smaller sub-sample sizes compared to the total sample, the clinical relevance of the difference is greater.
(2) Rating by others—procedures are more important for testing the efficacy than self-rating procedures since the first can always be applied.

Each of the three statements is one-sided in itself and not correct, and has to be specified under consideration of selection flows. Finally, where the differences between the two sub-samples (SR/R) are to be found needs clarification. Regarding the initial diagnosis, the following can be seen (Table 4 and Figs. 1–5): in the Hamilton scale higher values of group R compared to group SR (R: 40.9, SR: 33.5), in the NOSIE scales higher values in scales 1, 2, 3 ($P < 0.01$ each), lower values in scales 4, 5 (at least $P < 0.10$), with no statistical difference in scales 6, 7.

While the physician judged the grade of depression higher in group R than in group SR, the nursing staff did not (NOSIE scales 6, 7: no difference). In total, patients of group R were more severely ill than those of group SR. However, the grade of illness in the latter was severe enough to test the efficacy of anti-depressant drugs (Baumann 1976; Netter and Beckmann 1985).

In considering our results the following conclusions can be drawn with regard to the self-rating: rating by others ratio. Studies in psychiatric in-patients with complete data available from rating by others and self-rating have to be checked critically with regard to selection influences. The more severe the illness the more it is to be suspected that completeness of data material could only be obtained by means of methodically critical assessable influence factors (i.e. selection of patients without self-rating, reading the self-rating questionnaires to patients, which would no longer be a self-rating procedure (Lehrl et al. 1977). As far as self-rating and rating by others similar procedures could likewise be used, comparable or information completeness can be obtained. This depends on the parameters to be studied and cannot be answered definitely with regard to supplementation or redundancy. In single sub-samples, self-rating data can only be obtained by patient selection. In this case it is less problematical to do without self-rating data than without sub-samples which usually comprise particularly impaired patients.

References

AMDP (1983) (Hrsg) Testmanual zum AMDP-System. Verfaßt von U Baumann, RD Stieglitz. Springer, Berlin 1983

Baumann U, Meyer B, Suter U (1974) Formale und inhaltliche Überprüfung der PND-Skalen. Diagnostica 20:3–21

Baumann U (1976) Methodische Untersuchungen zur Hamilton-Depression-Skala. Arch Psychiatr Nervenkrank 222:359–375

Baumann U, Seidenstücker G (1977) Untersuchungsinstrumente. Zur Taxonomie und Bewertung psychologischer Untersuchungsverfahren bei Psychopharmakaprüfung. Pharmakopsychiatrie Neuro-Psychopharmakologie 3:165–175

Borbe HO, Müller WE (1984) Amitriptylinoxid – Ein Fortschritt in der Therapie mit trizyklischen Antidepressiva? Medizinische Welt 35:1438–1442

Borromei A (1982) A new antidepressant agent: Amitriptyline-N-oxide. Adv Biochem Psychopharm 32:43–47

Benkert O, Hippius H (1980) Psychiatrische Pharmakotherapie, 3. Auflage. Springer, Berlin Heidelberg New York

Busch H, Müller-Oerlinghausen B (1978) Psychopharmaka. In: Kuemmerle HP (Hrsg) Methoden der Klinischen Pharmakologie. München, Urban und Schwarzenberg

Campbell DT, Fiske DW (1959) Convergent and discriminant validation by the multitrait multimethod matrix. Psychol Bull 56:81–105

Cassano GB, Conti L, Dell'Osso L, Massimetti G (1982) Four new atypical antidepressants (Trazodone, Mianserin, Nomifensine, Caroxazone): Results of controlled trials. Adv Biochem Psychopharm 32:125–139

Cassano GB, Conti L, Massimetti G, Fornaro P, Re F (1983) A controlled clinical trial between amitriptyline and its N-oxide metabolite. Psychopharmacology Bulletin 19, 1:98–103

CIPS (1981) Internationale Skalen für die Psychiatrie. Beltz, Weinheim

Costa E, Racagni G (eds) (1982) Typical and atypical antidepressants. In: Advances in biochemical psychopharmacology, vol 32. Raven Press, New York

Debus G (1974) Selbst- und Fremdeinstufung des aktuellen Befindens in psychopharmakologischen Längsschnittuntersuchungen. Arzneimittel-Forschung 24:996–1001

Edwards BC, Lambert MJ, Moran PW, McCully T, Smith KC, Ellington AG (1984) A meta-analytic comparison of the Beck Depression Inventory and the Hamilton Rating Scale for Depression as measured of treatment outcome. Br J Clin Psychol 23:93–99

Fahrenberg J, Kuhn M, Kulick B, Myrtek M (1977) Methodenentwicklung für psychologische Zeitreihenstudien. Diagnostika 23: 15–36

Fahrenberg J (1984) Methodische Überlegungen zur Mehrebenen-Prozeßforschung. In: Baumann U (Hrsg) Psychotherapie: Makro-/Mikroperspektive. Hogrefe, Göttingen

Greef K, Köhler E (1977) Vergleich der Wirkung von Amitriptylin und Amitriptylin-N-oxid auf Herz und Kreislauf. Herz/Kreislauf 9:17–19

Kass F, Charles E, Klein DF, Cohen P (1983) Discordance between the SCL-90 and therapist's psychopathology ratings. Implications for clinical assessment. Arch Gen Psychiat 40:389–393

Kazdin AE (1980) Research in clinical psychology. Harper and Row Publishers, New York

Kerr TA, McClelland HA, Stephens DA, Ankier SI (1984) Trazodone—A comparative clinical and predictive study. Acta Psychiatr Scand 70:573–577

Kieback D (1982) Vergleich der Hamilton Rating Scale for Depression (HRS) mit der v. Zerssen-Befindlichkeitsskala (BS) am Beispiel einer klinischen Prüfung von Trazodon versus Amitriptylin. Pharmacopsychiatry 15:97–101

Lehrl S, Gallwitz A, Nusko G, Burkard G (1977) Warum geben Selbstbeurteilungsverfahren schwere Depressionen falsch wieder? Empirische Untersuchungen zur Einschränkung verschiedener Erklärungsmöglichkeiten. Therapeutische Umschau 10:687–693

Leuschner F, Neumann W, Doppelfeld IS, Erdös A, Fontaine R, Gabbar HA, Graf E (1978) General pharmacology of amitriptylinoxide. Arzneimittel-Forschung/Drug Res 28:1883–1893

Lienert GA (1978) Verteilungsfreie Methoden in der Biostatistik, Bd II. A Hain, Meisenheim an der Glan

McNair DM (1974) Original investigations self-evaluations of anti-depressants. Psychopharmacologia 37:281–302

Netter P, Beckmann H (1985) Procedings of the Congress of the AGNP Nürnberg 1983. Pharmacopsychiatry 18:1–178 (S 80ff)

Peterson L (1983) A brief methodological comment on possible inaccuracies induced by multimodal measurement analysis and reporting. J Behav Med 7:307–313

Platz W, Bartsch H (1984) Therapie depressiver Patienten – Doppelblindstudie Amitriptylin versus Amitriptylinoxid (Equilibrin® 30/60). Der Allgemeinarzt 6, 6:548–552

Rapp W (1978) Comparative trial of amitriptyline-N-oxide and amitriptyline in the treatment of out-patients with depressive syndromes. Acta Psychiatr Scand 58:245–255

Seefried S (1982) Equilibrin® bei älteren depressiven Patienten. Ergebnisse einer Therapiestudie. ZFA, 58:35/36

Seidenstücker G, Baumann U (1978) Multimethodale Diagnostik. In: Baumann U, Berbalk H, Seidenstücker G (Hrsg) Klinische Psychologie – Trends in Forschung und Praxis, Bd 1. Huber, Bern

Shapiro MB, Post F (1974) Comparison of self-ratings of psychiatric patients with ratings made by a psychiatrist. Br J Psychiatr 125:36–41

Simhandl Ch, Rogan M, Lesch ON, Musalek M, Strobl R (1984) Wertigkeit von Fremd- und Selbstbeurteilungsskalen bei chronisch Schizophrenen. Nervenarzt 55:371–377

Waskoff JE, Parloff MB (1975) Psychotherapy change measures. Rockvielle, MD, Nat Inst Ment Health

Wenzl H (1978) Special experiments on central nervous effects of amitriptyline oxide considering pharmacokinetic aspects. Arzneimittelfoschung/ Drug Res 28:1879–1883

Wiggins JS (1973) Personality and prediction. Principles of personality assessment. Addison-Wesley, Reading, MA

Zerssen D v (1976) Klinische Selbstbeurteilungsskalen Kst.-S. Allgemeiner Teil, Manual. Beltz, Weinheim

Zerssen D v, Möller HJ (1980) Psychopathometrische Verfahren in der psychiatrischen Therapieforschung. In: Biefang S (Hrsg) Evaluationsforschung in der Psychiatrie: Fragestellungen und Methoden. Enke, Stuttgart

Received May 13, 1985