SENSITIVITY OF SYMPTOM AND NONSYMPTOM-FOCUSED CRITERIA OF OUTPATIENT DRUG EFFICACY[1]

RONALD S. LIPMAN, PH.D.,[2] JONATHON O. COLE, M.D.,[3]
LEE C. PARK, M.D.,[4] AND KARL RICKELS, M.D.[5]

[Reprinted from THE AMERICAN JOURNAL OF PSYCHIATRY,
Vol. 122, No. 1, July, 1965]

SENSITIVITY OF SYMPTOM AND NONSYMPTOM-FOCUSED CRITERIA OF OUTPATIENT DRUG EFFICACY[1]

RONALD S. LIPMAN, PH.D.,[2] JONATHON O. COLE, M.D.,[3]
LEE C. PARK, M.D.,[4] AND KARL RICKELS, M.D.[5]

INTRODUCTION

The evaluation of clinical change in psychiatric outpatients has proved most difficult, and much time and effort have gone into the development of measurement techniques(1, 2, 3). In part, the observed variability of results in outpatient drug evaluations(4) may stem from the use of varied outcome criteria by different investigators.

While everyone would agree that an adequate drug trial should include criterion measures which reliably reflect relevant therapeutic change, and that, at a minimum, with the minor tranquilizers we typically look for an improvement in mood and a reduction in the psychic and somatic discomfort of the patient, questions remain with regard to the use of multiple raters and the actual choice of outcome criteria.

For these reasons we have employed the strategy of including a variety of different outcome measures using both the doctor and patient as raters. In proceeding from study to study, insensitive measures would be deleted and new measures added. Proceeding in this way it is feasible to develop a core battery of maximally sensitive criteria.

BACKGROUND

The findings reported here were gathered as part of the Collaborative Outpatient Program involving the Psychopharmacology Service Center of the National Institute of Mental Health, the Henry Phipps Psychiatric Clinic of The Johns Hopkins University, the Psychiatric Clinic of the University of Pennsylvania and the Psychiatric Clinic at Philadelphia General Hospital. These findings were derived from a meprobamate(5) and chlordiazepoxide trial in which we employed some traditional and some relatively unique criterion measures. Both studies were double-blind and placebo controlled, with the hospitals following identical protocols. Patients accepted for treatment were anxious, tense neurotics without sociopathy, organic impairment, alcoholism or marked depression.

Presenting complaints, in deceasing order of frequency as marked on checklists, were "nervousness or shakiness inside," "feeling easily annoyed or irritated," "headaches" and "feeling fearful." Whites and Negroes were about equally represented; females outnumbered males by about 2:1; age ranged from 18 to 65, with the mean age in the early 30's; roughly three-fourths of the sample had prior exposure to psychotropic drugs; and approximately 80 per cent of the patients were from the two lower social classes as defined by the Hollingshead classification(6). An *N* of 138 and 204 patients, respectively, completed each drug project.

CRITERION MEASURES

Our criterion measures can be roughly classified as symptom- or nonsymptom-focused. Symptom-focused measures included: (a) a 65-item Symptom Check List (SCL), adopted from The Johns Hopkins Distress Check List(7) and covering the more common psychoneurotic complaints; (b) target symptoms—defined at

---

[1] Read at the 121st annual meeting of The American Psychiatric Association, New York, N. Y., May 3-7, 1965.

The data were collected from studies supported by two NIMH-PSC research grants from the National Institute of Mental Health to the University of Pennsylvania (MH 04731) and to The Johns Hopkins University (MH 04732).

[2] Program Head, Outpatient Studies, Psychopharmacology Service Center, National Institute of Mental Health, Bethesda, Md.

[3] Chief, Psychopharmacology Service Center, NIMH.

[4] Assistant Professor of Psychiatry, The Johns Hopkins University School of Medicine, Baltimore, Md.

[5] Associate Professor of Psychiatry, School of Medicine, University of Pennsylvania, Philadelphia, Pa.

the first treatment visit as any complaint on the SCL checked as present by both patient and doctor; in effect, Target Symptoms represent the most salient complaints of the patient; *i.e.*, both those reported on a paper form and presented verbally to the treating doctor; (c) Anxiety, consisting of 15 adjectives culled from the Clyde Mood Scale and McNair-Lorr Tension-Anxiety Factor; (d) Depression, consisting of 15 adjectives culled from the Clyde Mood Scale and the VA Depression Factor; and (e) the six Clyde(8) Mood Scale Factors.

Nonsymptom-focused criteria included: (a) Global Improvement, measured by a seven-point scale ranging from seven, "very much worse," through four, "no change," to one, "very much better" (two reference points—"since the start of treatment" and "since the last clinic visit"—were indicated for this question); (b) significant life-situation events reported by the patient to the treating psychiatrist who, in turn, rated these events as positive or negative (collected only in the chlordiazepoxide trial and only at one clinic, $N = 62$); (c) Patient's Evaluation of the Doctor (P.E.D.), an instrument developed by Dr. Balter of the Psychopharmacology Service Center including such adjectives as "warm," "friendly," "interested in me as a person" which are rated by the patient to show the extent to which these describe his doctor; (d) the patient's desire, after completing the trial, to continue or discontinue taking the prescribed medication (only in the chlordiazepoxide trial); and (e) doctor medication guess—at each rating period the treating psychiatrist guessed whether the patient was receiving active or inactive medication.

PROCEDURE

In the meprobamate study patients were scheduled for four biweekly appointments and filled out criterion forms before seeing their assigned doctor. The SCL, Anxiety, Depression and the Clyde Mood Scale were completed at Weeks 0, II, IV and VI, the Global Measures at Weeks II, IV and VI and the P.E.D. at Week VI. Psychiatric residents completed these same criterion ratings (with, of course, the exception of P.E.D.) independently and immediately after each patient visit.

In filling out the distress and mood measures, patients rated themselves with respect to how they had felt during the prior week. Doctors based their ratings on information elicited from the patient during the 15- to 30-minute treatment visit. Frequent nondirective probes were used. A "not elicited" category was checked by the doctor whenever appropriate.

Similar procedures obtained in the chlordiazepoxide trial, which lasted only one week. At one clinic patients were also asked whether anything "significant" or "important" had happened to them during the medication week. With the very few patients who seemed confused by this request, an example of a marriage or death in the family was cited by the doctor.

Following the last trial visit all patients were told that the psychiatrist in charge of the clinic would see them in order to plan future treatment. During this disposition interview patients were asked whether they wanted to continue taking their medication.

RESULTS AND DISCUSSION

Rater type. In general, although a fair degree of concordance was found between patient and doctor ratings on the same measures, there were sufficient differences to consider both kinds of raters as providing useful information. For example, doctors tended to pick up drug effects early and patients late in the trial. More detailed results are presented elsewhere(9).

Measurement period. Our experience has strongly indicated the desirability of obtaining pretreatment criterion scores on patients. By employing covariance procedures to statistically adjust posttreatment scores for individual differences in initial distress levels, we have found, particularly with patient ratings such as Target Symptoms, that as much as 60 per cent of the postscore variance is removed, thus permitting a much more sensitive test of treatment outcome. More than twice the number of patients would have been needed to provide this same degree of sensitivity, assuming constant mean differences between medication groups, had only posttreatment scores been obtained.

Sensitivity. Of the sympton-focused criterion measures, Target Symptoms consis-

tently revealed the most reliable drug-placebo differences. The SCL and the Anxiety and Depression Scales fell in a middle range of sensitivity while the factors of the Clyde Mood Scale (Score Key #3) showed poor sensitivity; subsequent revisions of this scale may prove more sensitive.

Since the Clyde Mood Scale was developed with college students, it is not surprising that it has proved insensitive in low education clinic populations both in these studies and others(10, 11), while proving quite useful in acute drug studies in university settings(12).

Our anxiety depression data do not really comment on the sensitivity of the corresponding VA mood factors since only 5 of the 15 items in each scale were taken from the larger VA factors. The current status of the VA mood factors in neurotics has recently been reviewed by McNair and Lorr(13).

Nonsymptom-focused measures generally proved more sensitive to drug effects than symptom-focused measures, with the exception of Target Symptoms. Global Improvement, which allows the rater to make an assessment in which all facets of change may be computed and integrated psychologically, proved quite sensitive.

Doctor medication guesses at all time periods in both studies showed that doctors were reliably "breaking the double-blind." That is, they were correctly guessing the medication the patient was receiving—drug or placebo—reliably better than chance. When doctor medication guesses were related to whether or not the patient had reported side effects, no relationships were evident. However, highly significant relationships were found when doctor medication guesses were related to their global impressions of patient improvement (X^2's ranged from 11.4, $p < .001$ to 5.48, $p < .02$). Thus, their medication guesses in these data can be considered as a further index of differential clinical improvement.

A reliable drug effect ($p < .005$) was also indexed by the higher proportion of chlordiazepoxide-treated patients as compared with control patients who wanted to continue on medication. This measure has a particularly strong face validity since this information was elicited at a disposition interview when future treatment plans and options were being discussed with the patient.

The patients' reports of "significant" life situation events(14), mainly relating to their interaction with other people, and their evaluation of the treating doctor (P.E.D.) both revealed the presence of a reliable drug effect ($X^2 = 6.94$, $p < .05$; $F = 6.54, p < .025$). Chlordiazepoxide-treated patients perceived "significant" others more positively than did control patients.

Our findings with these last two measures seem particularly interesting and meaningful clinically, since they raise the intriguing hypothesis that the minor tranquilizers may influence the perceptual-processing component of neurosis. This finding may, however, be specific to chlordiazepoxide, since our meprobamate P.E.D. data did not reveal a reliable drug effect. Further, additional life-situation events data collected by Rickels in other meprobamate studies also do not show a greater incidence of positive events reported by meprobamate than by placebo-treated patients.

GENERAL REMARKS AND RECOMMENDATIONS

In addition to the measures we have discussed, the investigator should also be aware of differential drop-out rates(5, 10), dosage deviations(15, 16, 17), and in flexible dosage studies, differential prescribing rates of drug and placebo since all these factors may reveal useful criterion information.

We would most definitely recommend the inclusion of at least one criterion measure that is independently rated by both the patient and the doctor, particularly if placebo controls are not employed. A combination of symptom and nonsymptom-focused measures—such as Target Symptoms and Global Improvement—would seem desirable. In single medication studies, it is useful to determine, at the end of the trial period, whether the patient wants to continue medication. In cross-over studies where the patient is exposed to a sequence of medications, Rickels *et al.*(18) and Wheatley(19) have tapped patient medication preference at the end of the trial. This procedure has provided sensitive criterion information.

Although our findings are based on anxious neurotic patients seen in an outpatient clinic setting, we suspect that our findings have varying degrees of relevance for other patient populations and other treatment settings. Global ratings of improvement are also quite sensitive measures of drug-placebo difference in hospitalized acute schizophrenics(20) and in depressive states(21).

Finally, we should like to indicate that patients, seen in the clinic setting or in the office of the general practitioner or private psychiatrist, do not object to filling out rating scales when they are not overly complicated or lengthy and are presented as part of the medical routine useful to the treating doctor.

A judicious selection of criterion measures, appropriate to the medical setting and the patient sample, can go a long way toward increasing the sensitivity of drug evaluation and, consequently, of drug treatment.

BIBLIOGRAPHY

1. Frank, J., Gliedman, L., Imber, S., Stone, A., and Nash, E.: Patients' Expectancies and Relearning as Factors Determining Improvement in Psychotherapy, Amer. J. Psychiat. 115:961-968, 1959.

2. Lorr, M.: "Rating Scales, Behavior Inventories, and Drugs." In Uhr, L., and Miller, J., eds.: Drugs and Behavior. New York: John Wiley and Sons, Inc., 1960.

3. Parloff, M., Kelman, H., and Frank, J.: Comfort, Effectiveness, and Self-awareness as Criteria of Improvement in Psychotherapy, Amer. J. Psychiat. 111:343-352, 1954.

4. Laties, V., and Weiss, B.: A Critical Review of the Efficacy of Meprobamate (Miltown, Equanil), in the Treatment of Anxiety, J. Chronic Dis. 7:500-519, 1958.

5. Fisher, S., Cole, J. O., Rickels, K., and Uhlenhuth, E. H.: "Drug-Set Interaction: The Effect of Expectations on Drug Response in Outpatients." In Bradley, P. B., *et al.*, eds.: Neuropsychopharmacology. Amsterdam: Elsevier Publishing Co. 1964, 149-156.

6. Hollingshead, A.: Two-Factor Index of Social Position. New Haven: A. B. Hollingshead, 1957.

7. Frank, J. D., Gliedman, L. H., Imber, S. D., Nash, E. H., and Stone, A. R.: Why Patients Leave Psychotherapy, A.M.A. Arch. Neur. Psychiat. 77:283-299, 1957.

8. Clyde, D. S.: "Self-ratings." In Urh, L., and Miller, J. G., eds.: Drugs and Behavior. New York: John Wiley and Sons, Inc., 1960.

9. Park, L. C., Uhlenhuth, E. H., Lipman, R. S., Rickels, K., and Fisher, S.: A Comparison of Doctor and Patient Improvement Ratings in a Double-blind Drug Trial, Brit. J. Psychiat. 3:535-540, 1965.

10. Rickels, K., Boren, R., and Stuart, H. M.: Controlled Psychopharmacological Research in General Practice, J. New Drugs. 4:138-147, 1964.

11. Wittenborn, J. R., Plante, M., Burgess, F., and Maurer, H.: A Comparison of Imipramine, Electroconvulsive Therapy and Placebo in the Treatment of Depressions, J. Nerv. Ment. Dis. 135:131-137, 1962.

12. DiMascio, A., Havens, L. L., and Klerman, G. L.: The Psychopharmacology of Phenothiazine Compounds: A Comparative Study of the Effects of Chlorpromazine, Promethazine, Trifluoperazine, and Perphenazine in Normal Males. II. Results and Discussion, J. Nerv. Ment. Dis. 136:168-186, 1963.

13. McNair, D. M., and Lorr, M.: An Analysis of Mood in Neurotics, J. Abnorm. Soc. Psychol. 69:620-627, 1964.

14. Lipman, R. S., Hammer, H. M., Bernardes, J. F., Park, L. C., and Cole, J. O.: Patient Report of Significant Life Situation Events: Methodological Implications for Outpatient Drug Evaluations, Dis. Nerv. Syst. In press.

15. Joyce, C. R. B.: Patient Cooperation and the Sensitivity of Clinical Trials, J. Chronic Dis. 15:1025-1036, 1962.

16. Park, L., and Lipman, R.: A Comparison of Patient Dosage Deviation Reports with Pill Counts, Psychopharmacologia 6:299, 1964.

17. Uhlenhuth, E., Park, L., Lipman, R., Rickels, K., Fisher, S., and Mock, J.: The Magnitude of Drug-Effect as Related to Dosage Deviation, J. Nerv. Ment. Dis. In press.

18. Rickels, K., Clark, T. W., Ewing, J. H., Klingensmith, W. C., Morris, H. M., and Smock, C. D.: Evaluation of Tranquilizing Drugs in Medical Outpatients: Meprobamate, Prochlorperazine, Amobarbital, Sodium, and Placebo, J.A.M.A. 171:1649-1656, 1959.

19. Wheatley, D.: The General Practitioner Research Group, Clin. Pharmacol. Ther. 4:542-547, 1963.

20. NIMH-PSC Collaborative Study Group: Phenothiazine Treatment in Acute Schizophrenics, Arch. Gen. Psychiat. 10:246, 1964.

21. Klerman, G., and Cole, J. O.: Clinical Pharmacology of Imipramine and Related Compounds, Pharmacol. Rev., 1965. In press.