Psychotropic Drug Response:
Advances in Prediction

Edited by Philip R.A. May, M.D.
and J.R. Wittenborn, Ph.D.

Charles C. Thomas Publisher,
Springfield, IL. 1969

Chapter 8

SOME NONPHARMACOLOGIC MODIFIERS OF THE RESPONSE TO IMIPRAMINE IN DEPRESSED, PSYCHONEUROTIC OUTPATIENTS: A CONFIRMATORY STUDY

E. H. Uhlenhuth, David B. Duncan, and Lee C. Park

THE RESULTS to be reported here are based on data gathered in a study on the influence of imipramine in relieving depressive symptoms in psychoneurotic outpatients (6). The main problem concerned the possible influence of a nonpharmacologic variable, the individual psychiatrist, upon the effect of the drug (defined as the difference in response to drug and to placebo).

METHOD

The design, in brief, was a double-blind crossover with two medications, seven doctors, and forty-two psychoneurotic patients. The distribution of patients who completed the study was slightly unbalanced with regard to treatment conditions. The patients took each medication for four weeks.

Each patient's condition was assessed every two weeks during a brief interview with his treating psychiatrist, structured around the completion of a checklist of sixty-two symptoms adapted from that of Parloff *et al.* (5). Each symptom was scored 0, 1, 2, or 3, according to the degree of distress reported by the pa-

Note: This work was supported by grants K3-MH-18, 611, MH-06350 and MH-04732 from the NIMH. The computations were performed at the Computing Center of the Johns Hopkins Medical Institutions, which is supported by Research Grant No. FR-00004 from the NIH. This is paper number 380 from the Department of Biostatistics.

tient. A total weighted distress score at each interview was computed by summing the scores for the individual symptoms.

Eight senior psychiatrists independently classified the sixty-two symptoms into the following subgroups: depression (13 symptoms), anxiety (13 symptoms), secondary (10 symptoms such as phobias, obsessions, conversions), overlapping (12 symptoms which might belong in 2 or 3 of the first 3 subgroups), and miscellaneous (8 symptoms which did not belong in any of the first 3 subgroups). A symptom was included in the depression, anxiety, secondary, or miscellaneous subgroups only if at least five psychiatrists classified it in the same subgroup. Six symptoms were segregated into a "reject" subgroup since higher scores on these items were expected to accompany *less* distress. The weighted distress score for each symptom subgroup was computed by summing the scores for the individual symptoms included in the subgroup. Other types of assessments also were employed but are not pertinent to the present chapter.

EARLY RESULTS

The results using the total weighted distress score as the criterion of change may be summarized as follows: Patients experienced relief in proportion to their initial level of distress, irrespective of medication. This relationship illustrates the "law of initial value" (8) with respect to a psychologic variable. During the first two weeks of the study, patients taking imipramine improved more than patients taking placebo, even when the higher initial distress level among the patients taking imipramine was taken into account. After the first two weeks, both groups of patients experienced about the same amount of relief. Effects due to the seven individual doctors and effects due to the interaction between doctors and medications could not be demonstrated at any time, although the conditions of the study were unduly stringent, as pointed out in the original report.

NEW METHODS OF ANALYSIS

Information about some other variables that might modify the drug effect also was collected in the original study. The possible

effects of these variables were not examined originally because of practical limitations in the methods available for analyzing jointly the effects of multiple variables.

In the meanwhile, however, a general multiple covariance procedure for the electronic computer has been developed. This procedure can handle in the same analysis both quantitative data such as age and qualitative (classification) data such as sex. It also deals effectively with a disproportionate distribution of subjects among categories in classification data. The procedure offers the option of searching stepwise a pool of independent variables for the combination which best describes the change in the dependent variable. This option selects only one of several highly correlated variables. The method combines in effect the functions of analysis of variance, analysis of covariance, and multiple regression analysis in a form sufficiently flexible to cope with many problems in the statistical evaluation of quantitative observations in real-life situations.*

With this new method, two sets of preliminary, nonsearch analyses were performed. One set analyzed the patients' relief after two weeks of treatment, using the change in total weighted distress score from visit 1 to visit 2 as the dependent variable. The other set analyzed the patients' relief after four weeks of treatment, using the change in total weighted distress score from visit 1 to visit 3 as the dependent variable.

Each analysis of relief included as independent variables (a) medication, (b) one or more possible modifying variables selected from the original data on the basis of other reports in the literature, and (c) variables representing the statistical interaction between medication and each possible modifier.

Results of Non-search Analyses After Two Weeks

The results of the analyses after two weeks of treatment are easy to describe: none of the new variables produced any sub-

*For investigators who may wish to consider this approach in more detail, a discussion of the statistical considerations is included as Appendix 8 A and a practical guide to the computer program as Appendix 8 B. Materials for the use of the program are available from Dr. Uhlenhuth.

stantial reduction in the error mean square, as compared to a reference analysis including only medication and initial total distress score as independent variables. Thus the effectiveness of imipramine during the first two weeks of treatment appears to have been quite general within this sample of depressed, psychoneurotic outpatients. The remainder of this chapter deals only with the analyses of relief after four weeks of treatment.

Results of Non-search Analyses After Four Weeks

Effect of Patient's Initial Symptomatology

Kiloh and his associates (2, 3) have stressed the importance of the detailed symptomatology as a guide to the pharmacologic treatment of depressed outpatients. In the present study, the patients' initial scores on the six symptom subgroups offered a more specific characterization of their presenting clinical pictures, at least on the subjective side. Therefore an analysis of relief was performed with independent variables including medication and, instead of the initial total distress score, the initial scores for each of the six symptom subgroups and the scores for their interactions with medication. The interaction scores were obtained by multiplying the patient's medication score by his initial score for each symptom subgroup in turn.

This analysis provided a more precise (error mean square = 483.27) overall description of relief than the reference analysis (error mean square = 576.35). Several of the individual variables within the analysis showed reliable effects. Patients with a higher initial score on the depressive symptom subgroup experienced significantly* ($p < .05$) more relief, irrespective of medication (Fig. 8-1). This result probably reflects the operation of the "law of initial value."

This finding, however, disagrees with the report of Kiloh *et al.* (2), who found that more severely depressed patients did not respond well to imipramine. The respective ranges of severity

*All p values are for two-tailed tests. In each figure, change in total weighted SCL scores is adjusted for the effects of all independent variables that enter into the analysis, but do not appear on the figure.

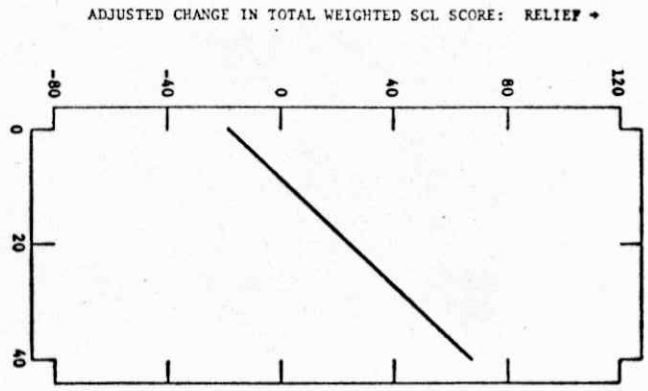

FIGURE 8-1. Initial weighted depression score.

represented in the two studies are difficult to compare. Kiloh's range may be the more severe. If so, then the divergent results of the two studies suggest that the effect of severity is nonlinear, with a moderate severity representing the best outlook for response to imipramine. Differences between the two studies, in the variables included and in the correlations among variables, may mean that the contrasting results are more apparent than real.

The patient's initial score on the secondary symptom subgroup was another important variable in the present analysis (Fig. 8-2). It showed a significant ($p < .01$) interaction* with medication: patients with a low initial level of secondary symptoms responded about the same to imipramine and to placebo, but patients with a high initial level showed less response to imipramine than to placebo. This result corresponds to Kiloh's report of an unfavorable outlook for treatment with imipramine in the presence of such complications as hypochondriasis and hysterical features.

The initial score on the overlapping symptom subgroup also showed a reliable ($p < .005$) interaction with medication, but in the opposite direction: patients with a high initial level

*The presence of significant interaction implies significant effects for each of the component variables. The significance of each component effect depends on the level of the other component variable.

showed a greater response to imipramine than to placebo (Fig. 8-3). The outstanding feature of the overlapping symptoms was the clinicians' inability to assign symptoms exclusively to the depressive, anxious, or secondary subgroups. A review of the symptoms in the overlapping subgroup, however, suggests that, with

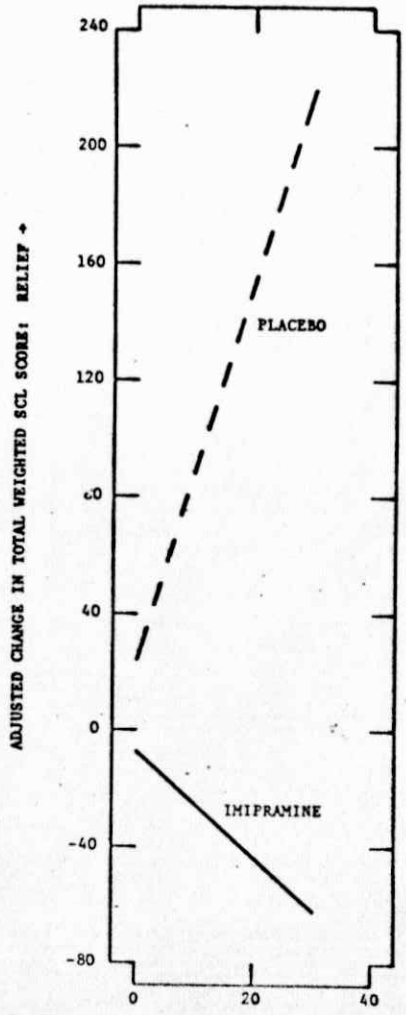

FIGURE 8-2. Initial weighted secondary S score.

depressed patients, many of these symptoms ma[illegible]ect a coloring of the depressive syndrome similar to Kiloh [illegible]litatively different from normal" characterization, which i[illegible]ated with the diagnosis of endogenous depression and a favo[illegible]response to imipramine (3).

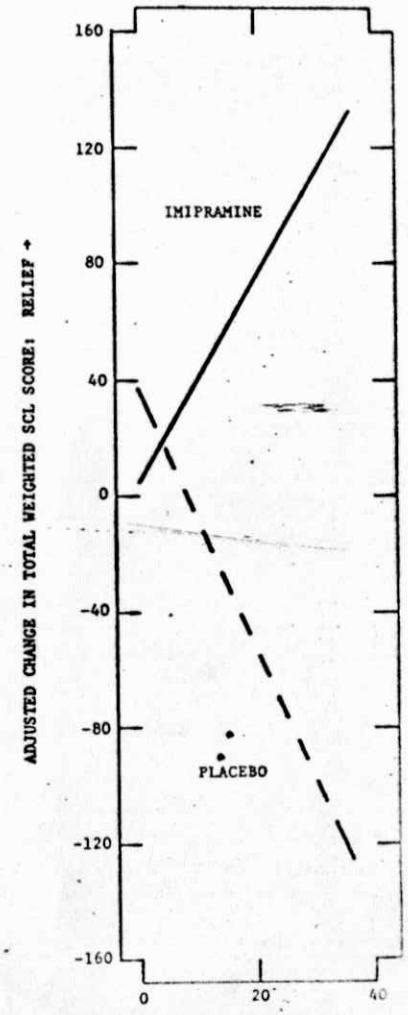

FIGURE 8-3. Initial weighted overlapping S sc[illegible]

Effect of Patient's Age

In addition to the symptom picture, Kiloh and Garside (2) reported age as an important variable modifying the response to imipramine, with older patients responding more favorably. In the present study, an analysis of relief with independent variables including medication, the initial level of total distress and age tended to confirm this finding (Fig. 8-4). Older patients responded more to imipramine than to placebo ($p < .10$, error mean square = 553.61). It is interesting that the crossover in the interaction, that is, the age at which patients responded equally to imipramine and placebo, occurred at about thirty-five years, not far from the age of 40 chosen by Kiloh *et al.* as the boundary between their two age groups.

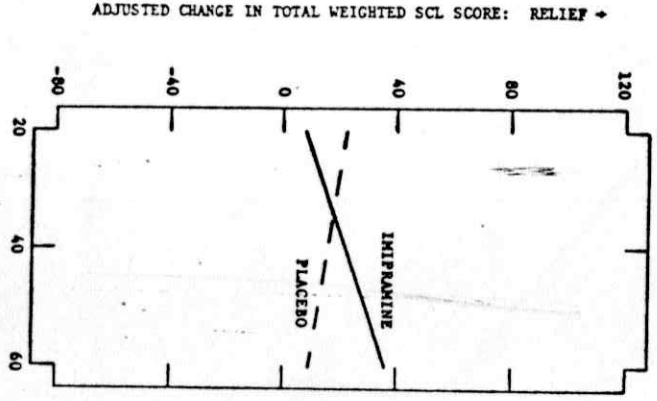

FIGURE 8-4. Patient's age.

Effect of Patient's Suicidal Trends as Rated by Doctor

Wittenborn, working with neurotic, though hospitalized, depressed patients, found that patients with a history of suicidal attempts tended to respond poorly to imipramine (9). In the present study, the treating psychiatrist initially rated each patient as suicidal or not suicidal. An analysis of relief with independent variables including medication, the initial level of total distress, and the doctor's initial suicidal rating showed a reliable ($p < .01$, error mean square = 465.23) interaction between the

effects of medication and the suicide rating as determined by the doctor (Fig. 5). Suicidal patients, in confirmation of Wittenborn's findings, showed less response to imipramine than to placebo.

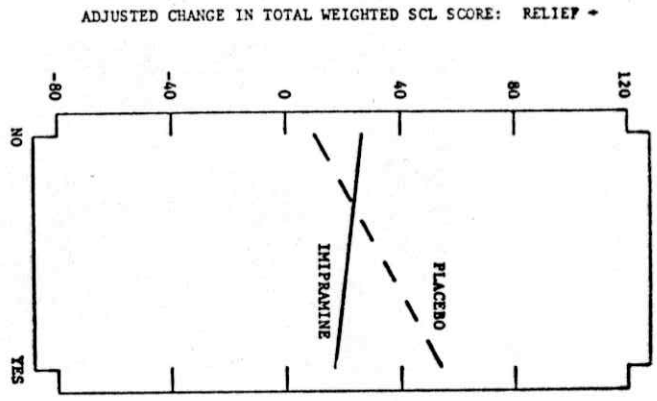

FIGURE 8-5. Patient suicidal?

Since there were only six suicidal patients (3 taking each medication) among the forty-two patients studied, it seemed advisable to examine the records of these six patients more closely for other characteristics which might distinguish them as a group from the other patients in the study. It was startling to find that four of the six patients rated as suicidal were under the care of the same psychiatrist, who rated only three of his seven patients as not suicidal. This situation suggested that (a) a large proportion of the suicidal patients in the study were assigned to this doctor by chance or (b) that this doctor had a systematic tendency to rate patients as suicidal by virtue of some unknown characteristic of his own.

Effect of Patient's Complaint of Suicidal Thoughts

Fortunately some data were available bearing upon both of these possibilities. One symptom on the checklist asked the patient to rate his suicidal thoughts on a four-point scale of intensity. These ratings, too, were subject to some influence by the psychiatrist, as they were made in his presence. However, the form of the complaint and the rating at least were standardized.

An analysis of relief, in which the patient's complaint of suicidal thoughts was substituted for the doctor's rating, showed a similar interaction ($p < .025$, error mean square = 436.19) between the effects of medication and suicidal preoccupations (Fig. 8-6). This finding, then, supported the notion that the patients' own suicidal inclinations affected their response to the drug and that the one psychiatrist in the study simply received a higher proportion of suicidal patients.

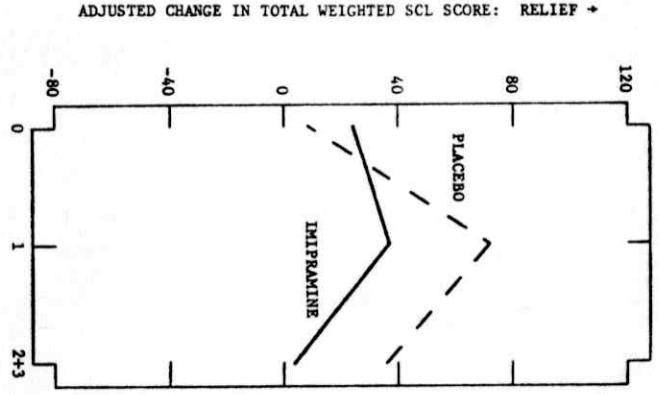

FIGURE 8-6. Patient's suicidal thoughts.

Effect of Doctor's Score on the A-B Scale

There was, however, some information indicating that this doctor also differed in a significant way from his colleagues. He had the lowest score on the A-B scale, which predicts psychotherapeutic success with schizophrenic patients [high score, "type A," see (7)] and with psychoneurotic patients [low score, "type B," see (4)]. Since the seven psychiatrists in the study were well distributed along the A-B scale, an analysis of relief was performed with independent variables including medication, the initial level of total distress, and the doctor's score on the A-B scale. A near-significant ($p < .10$, error mean square = 537.08) interaction emerged (Fig. 8-7). The patients of high-scoring (type A) doctors responded more to imipramine than to placebo, but the patients of low-scoring (type B) doctors responded about the same to the two medications. These findings suggest that the doctor's suicidal rating as well as the patient's subsequent response to medication may emerge from a complex process reflecting both the patient's suicidal preoccupations and his treating doctor's response to them.

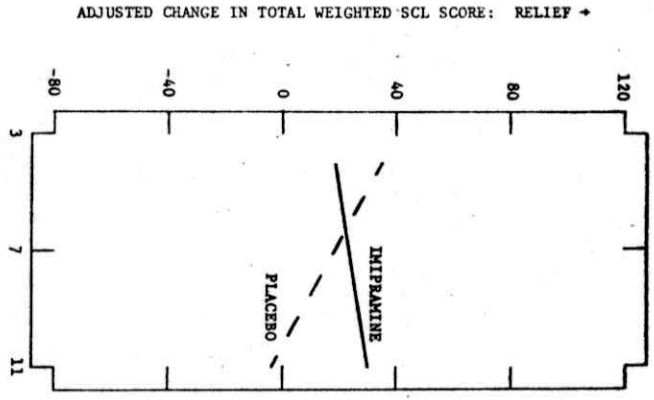

FIGURE 8-7. Doctor's A-B Scale Score.

Effect of Patient's Sex and Initial Anxiety Level

Hamilton offered perhaps the most complex suggestion about the effects of modifying variables upon the patient's response to imipramine (in comparison with phenelzine). In his study (1), the effect of imipramine was determined jointly by the patient's sex and the agitated or retarded quality of the patient's illness: agitated men responded better than retarded men, but retarded women responded better than agitated women. An analysis of relief in the present study employing this constellation of independent variables (with the initial score on the anxiety-symptom subgroup as an index of agitation) did not confirm Hamilton's hypothesis.

The group of potential modifying variables studied, as summarized in Table 8-I, includes several which replicate previous findings. This is striking, especially in view of the small number of patients and the differences in cultural settings, clinic settings, and specific procedures such as the criterion measures employed. Furthermore, the results of all the analyses undertaken are included in this chapter, excepting a few containing clerical errors.

TABLE 8-I

VARIABLES MODIFYING THE EFFECT OF IMIPRAMINE AFTER FOUR WEEKS OF TREATMENT

Variable (Modifier)		*Relief with Imipramine**
Increasing initial symptom subgroup scores:	Depressive	—
	Anxiety	—
	Secondary	less
	Overlapping	more
	Miscellaneous	—
	Reject	—
Increasing age		more
Suicidal trends present:	Doctor's rating	less
	Patient's complaint	less
Increasing score on A-B scale (type A doctor)		more
Sex and increasing initial anxiety symptom subgroup score		—

*Compared to placebo.

This is *not* a selected sample of analyses showing positive results.

Several of the modifiers in this group appeared related to one another. To whatever extent they were correlated, their respective effects upon the patient's response to medication were necessarily confounded. These considerations suggested that a joint analysis with certain independent variables selected from the entire group might yield a simpler and yet more precise description of relief in response to medication and its modifiers.

Results of Joint Analysis Using Search Option

A final analysis of relief was performed by submitting to the search option of the multiple covariance procedure as independent variables: medication, the entire group listed in Table 8-I (except for the doctor's rating of the patient's suicidal trend, which seemed too similar to the patient's complaint of suicidal thoughts), and the interactions between medication and each of the other variables. The critical F ratio was set at about the .05 confidence level (see Appendix 8-A). The analysis resulting

from this procedure met a tentative criterion of reliability, taking account of the added uncertainties in a search.* Medication and three modifying variables from the pool were selected for the final analysis (Table 8-II).

Figure 8-8 and Table 8-II (MED × SUICIDAL) show how the presence of suicidal thoughts modified the patient's response to imipramine. Nonsuicidal patients responded about the same to imipramine and to placebo. Patients who complained of suicidal thoughts, however, responded less to imipramine than to placebo.

Figure 8-9 and Table 8-II (MED × OVERLAPPING) show how the patient's initial score on the overlapping symptom subgroup modified his response to imipramine. Patients with a low initial score responded less to imipramine than to placebo, whereas the reverse held true in patients with a high initial score.

Figure 8-10 and Table 8-II (MED × SEX × ANXY) show how the patient's sex and initial score on the anxiety symptom subgroup modified his response to imipramine. Men with high initial anxiety responded about the same to imipramine and to placebo, whereas men with low initial anxiety responded less to

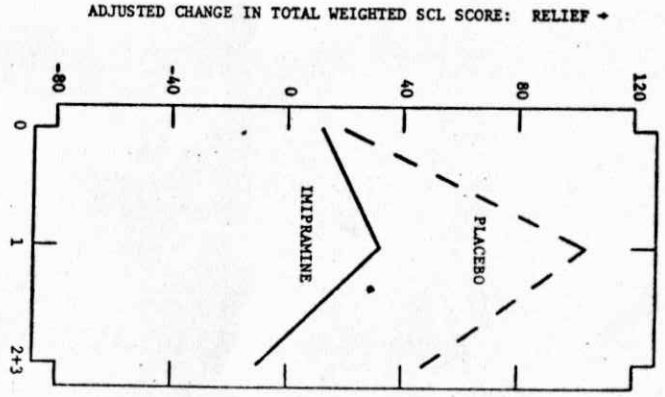

FIGURE 8-8. Initial score on suicidal thoughts.

*The "adjusted" F ratio for total regression was 3.16. (See Appendix 8 A for a more detailed discussion.)

TABLE 8-II

ANALYSIS OF VARIANCE AND REGRESSION OF CHANGE IN TOTAL DISTRESS AFTER FOUR WEEKS OF TREATMENT

Subset No.	*Name*	*DF*	*MS*	*F*	*Variable No.*	*B*	*SE*	*T*	*Mean*
	C TERM	1	18145.9285			26.1495			20.7857
2	MEDICATION	1	1366.1134		2	-58.3636	28.1004	2.0770	0.5238
4	SEX	1	1228.9406		4	47.8106	24.2702	1.9699	0.2619
6	SUICIDAL THOUGHTS	2	2948.4334		6	-28.7353	11.4463	2.5104	0.6905
					7	45.7587	24.2372	1.8879	0.1667
8	ANXIETY SX	1	625.0440		9	1.0722	0.7632	1.4049	16.5000
10	OVERLAPPING SX	1	165.3127		11	-0.7103	0.9831	0.7225	16.9524
14	MED × SEX	1	2810.7371		15	-100.8275	33.8441	2.9792	0.1667
16	MED × SUICIDAL	2	2043.6696	6.4533	17	50.9565	18.4817	2.7571	0.3095
					18	-3.9623	29.2469	0.1355	0.1429
18	MED × ANXIETY	1	198.2342		20	-1.0931	1.3816	0.7912	9.5714
20	MED × OVERLAPPING	1	1623.9128	5.1278	22	2.8951	1.2785	2.2645	9.5238
23	SEX × ANXIETY	1	801.1759		25	-2.1124	1.3281	1.5906	4.7143
24	MED × SEX × ANXY	1	1930.1310	6.0948	26	4.4250	1.7924	2.4688	3.0952
	REGRESSION	13	1682.2990	5.3122					
	ERROR	28	316.6852						
	TOTAL	41	749.6847						

FIGURE 8-9. Initial weighted overlapping S score.

FIGURE 8-10. Initial weighted anxiety score.

imipramine. Women responded less to imipramine at all levels of initial anxiety, but especially at high levels. This result, after all, favors Hamilton's hypothesis.

The discrepancies between the results of the individual analyses and the joint analysis reflect the complex pattern of correlations among the independent variables submitted for joint analysis. Such correlations in a particular sample, of course, may reflect meaningful relationships or chance occurrences. For example, the correlations between the patient's symptomatology and the

doctor's A-B type in the present study would be expected only by chance since patients were assigned to doctors at random. The failure of the search procedure to select the doctor's A-B type in the joint analysis therefore need not be interpreted as a clearly negative result.

On the other hand, "true" correlations among the patient's own symptoms and other chracteristics would be expected. The selection of the sex by anxiety interaction, for example, as an important modifier in this context suggests that it has a more striking *independent* effect upon drug response than do some of its correlates. This effect may have gone unnoticed in the individual analysis because this analysis also contained the initial total distress score, which, of course, includes the anxiety subgroup. Under these conditions, the positive results of the joint analysis would seem to merit the more weight.

Restructuring our understanding of psychologic processes in terms of the action and interaction of such multiple-variable fields inevitably raises knotty issues of interpretation. Their clarification depends upon (a) designing studies more suited to the independent assessment of factors contributing to outcome and (b) gaining further experience in applying such concepts and analyses to observed data.

DISCUSSION AND SUMMARY

PHILIP R. A. MAY: Dr. Uhlenhuth presents the results of reanalyzing in a different way data that were gathered several years ago in a double-blind, four-week study of the effects of imipramine and placebo on forty-two depressed psychoneurotic outpatients. The original analyses had shown that relief of overall distress was in proportion to the patients' initial distress level, and that during the first two weeks, patients taking imipramine improved more than those on placebo. After this, both groups of patients experienced about the same amount of relief.

Information about other variables which might act as modifiers of drug effect was also collected in the orginal study, but their possible value as predictors could not be examined because of practical limitations in the methods then available for multivariate analysis. Subsequently Uhlenhuth and Duncan developed a new method that, in essence, functions as a general multiple covariance program with or without a search feature. (Details are given in Appendices 8A and 8B.) This program can handle qualitative data (categorical classifications such as sex and religion) as well as ordered quantified data (such as age), performing on request a stepwise succession of search sequences to fit a multiple regression equation to a group it selects as the most important predictors. The main advantage of the new method lies in the fact that a large number of variables can be explored without arbitrarily rejecting potential predictors on a priori grounds. In this way, important effects may be detected that would otherwise be missed because they were never looked at. It also presents the results of the analysis in multiple covariance form, providing adjusted scores for the effect of each item taking into account the effects of all other items fitted.

When this method was used on the same data, the results early in treatment (after 2 weeks) were the same as before: they showed that imipramine had a general distress-reducing effect but no specific effects from any of the new variables. With longer treatment, however (4 weeks), several of the new variables showed significant effects indicating that the relief afforded by imipramine became restricted to a relatively narrow group of patients. The search procedure selected medication and four modifying variables as the most important predictors, indicating that the patient most suitable for treatment with imipramine is a man without suicidal tendencies and with high scores on two subscales—anxiety and a group of symptoms characterized as qualitatively different from normal and suggestive of endogenous depressions. The least-suitable candidate is a man with suicidal thoughts and with low scores on these same two subscales.

Dr. Uhlenhuth points out that there were some discrepancies between the results from individual analyses and from the search procedure, and I must agree heartily with his comment that application of a multivariate model must inevitably raise knotty problems of interpretation. However, the fact that a task is likely to be difficult is a poor excuse for not trying. Psychiatric

research has limped along on univariate analyses much too long for a field in which multidetermination appears so obvious to the clinician. The results from this new approach are encouraging, and it is to be hoped that other investigators will use it to enlarge our experience with multivariate analyses of psychiatric data.

In this connection, I wonder whether the analyses might not have been improved (more efficient) if *final* scores had been used instead of *change* scores. Wittenborn (1966) has pointed out that change scores are inherently exposed to a double dose of measurement error, and there are additional methodological reasons for questioning the advisability of using a patient's initial score as a covariate when the criterion variable is a change score whose computation has involved the use of that same initial score.

E. H. UHLENHUTH: Dr. May's comments raise some very subtle and difficult questions about the adequacy of the usual statistical models to the analysis of fallibly measured variables. No entirely satisfactory answer for these questions is available at present.

Given the accepted multifactor models, however, the results are the same for a change score and a final score analysis incorporating initial criterion score, except for the regression on initial score itself, which is less by one for the final score analysis. In the present search analysis, which includes only that part of the initial criterion score attributable to two of the six symptom subgroups, a final score analysis is substantially similar, but not identical, to the change score analysis reported.

References

1. HAMILTON, M.: A Controlled Trial: Some Lessons. Given at a research conference on the treatment of depressive states sponsored by McGill University, Feb. 5-7, 1965.
2. KILOH, L. G.; BALL, J. R. B., AND GARSIDE, R. F.: Prognostic factors in treatment of depressive states with imipramine. *Brit Med J, 1*:1225-1227, 1962.
3. KILOH, L. G., AND GARSIDE, R. F.: The independence of neurotic depression and endogenous depression. *Brit J Psychiat, 109*:451-463, 1963.
4. MCNAIR, D. M.; CALLAHAN, D. M., AND LORR, M.: Therapist "type" and patient response to psychotherapy. *J Consult Psychol, 26*:425-429, 1962.
5. PARLOFF, M. B.; KELMAN, H. D., AND FRANK, J. D.: Comfort, effectiveness and self-awareness as criteria of improvement in psychotherapy. *Amer J Psychiat, 111*:343-351, 1954.
6. UHLENHUTH, E. H., AND PARK, L. C.: The influence of medication (imipramine) and doctor in relieving depressed psychoneurotic outpatients. *J Psychiat Res, 2*:101-122, 1964.
7. WHITEHORN, J. C., AND BETZ, B. J.: Further studies of the doctor as a crucial variable in the outcome of treatment with schizophrenic patients. *Amer J Psychiat, 117*:215-223, 1960.
8. WILDER, J.: Modern psychophysiology and the law of initial value. *Amer J Psychother, 12*:199-221, 1958.
9. WITTENBORN, J. R.: Factors which qualify the response to iproniazid and to imipramine. In WITTENBORN, J. R., AND MAY, P. R. A. (Eds.): *Prediction of Response to Pharmacotherapy*. Springfield, Thomas, 1966.

APPENDIX 8A

A General Multiple Covariance Search Procedure: Statistical Considerations

In many clinical studies in the fields of psychiatry and psychology, the data may be expected to conform to the model of a multiple covariance analysis with many classification criteria (race, sex) and many concomitant regressors (age, initial anxiety score). Because of the large number of factors potentially, but not assuredly, involved and because of the nonorthogonal nature of the data (with unequal numbers in the subclasses), however, a direct full-scale multiple covariance analysis does not offer a practical means of examining these data.

To deal with these problems, a "search" type of multiple covariance procedure for the IBM 7094 has been developed. This is similar in approach to the multiple regression procedure de-

veloped at UCLA* which "searches" a given set of regressors and fits a multiple regression equation to a group containing only the more important ones. However, by introducing subsets of "dummy regressors" to handle the fitting of class effects in a multiple regression form, the new program extends the search capability to a general multiple covariance analysis. The purpose of this text is to present the new procedure and to discuss its use.

Method of Analysis

OUTLINE. The main steps, apart from exceptions to be mentioned later, may be summarized as follows: Step one in a preliminary phase consists of introducing "dummy" regressors to accomplish the analysis of classification effects by using multiple regression. A subset of one, two, or more regressors is introduced for each classification containing two, three, or more classes. The number of dummy regressors required is one less than the number of classes. Step two consists of the generation of an additional regressor subset to accomplish the analysis of interaction for each pair of regressor subsets for which interaction is required. For any two subsets of p and q regressors, the interaction subset consists of pq regressors obtained by multiplying each regressor in one subset by each regressor in the other.

It should be emphasized that here, and from here onward, the term *regressor subset* often refers to a single regressor. This term may refer to one of the original concomitant regressors, a single dummy regressor for a two-class classification, or an interaction regressor for interaction between two single regressors. It may also refer to a subset of two or more dummy regressors representing a classification with three or more classes, or to a subset of two or more interaction regressors.

The main "search" phase of the analysis then commences. This consists of the repetition of one or more search sequences. At the beginning of any search sequence except the first, the procedure already has fitted a multiple regression equation to one

*See DIXON, W. J. (ED.): *BMD Biomedical Computer Programs*. Health Sciences Computing Facility, Dept. of Preventive Medicine and Public Health, School of Medicine. Los Angeles, U. of Calif., Jan. 1, 1964.

or more regressor subsets. The subsets in the regression equation may be termed the *in* subsets; the remainder, the *out* subsets.

The first step of any search sequence consists of working through the *out* subsets with a view to adding the most important one to the regression equation. The search is conducted by fitting, in effect, a multiple regression to the *in* subsets together with each of the *out* subsets taken in turn. The importance of each *out* subset is scored in terms of the ratio for the additional regression it gives, that is, the F ratio obtained from the additional regression sum of squares and the reduced error sum of squares brought about by fitting the additional subsets. This "add-search" step is concluded by adding the most important of the *out* subsets to the regression equation if its F ratio exceeds a prescribed critical ratio F_c. (F_c may be taken for, example, as the tabled 5% value of F.)

If the add-search step terminates successfully, that is, with the addition of a subset, it is followed by a "drop-search" step. This consists of working through the *in* subsets with a view to dropping the least important if it has ceased making an important contribution. This search is conducted by refitting, in effect, the multiple regression equation with each of the *in* subsets dropped in turn. The importance of each subset is again scored in terms of the appropriate F ratio. This is now obtained in terms of the reduction in sums of squares for regression brought about by dropping the subset and the error sum of squares before the dropping. The drop-search terminates by dropping the least important *in* subset if its F ratio is less than the critical value F_c.

A successful drop-search, that is, one followed by the dropping of a subset, is followed by further drop-searches in sequence until a multiple regression equation is reached in which no subset has an F ratio below F_c.

The whole search phase consists of a succession of search sequences which may be diagrammed as in Figure 8A-1. It starts with an add-search which differs from the others in that there are no *in* subsets at the beginning. The F ratios tested in this add-search are for each of the regressor subsets taken individually and alone. From then onward, each search sequence starts

FLOW CHART FOR SEARCH PHASE

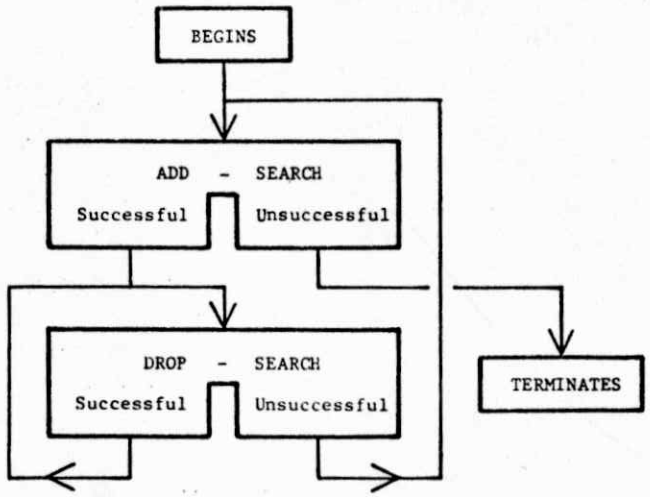

FIGURE 8A-1.

with an add-search step, and this is followed by one or more drop-search steps. If a search sequence terminates in an unsuccessful drop-search, the procedure goes on to the add-search step at the beginning of the next sequence. The entire search phase terminates at the first unsuccessful add-search.

The last phase of the analysis organizes the results of the final fitted equation. These are developed in the form of an analysis of variance showing the sums of squares, degrees of freedom, and so forth, for the respective regressor subsets and for error. The table also presents the estimates of the regression coefficients and their standard deviations. Finally, the coefficients for the dummy regressors are transformed and presented in the usual analysis of covariance form as adjusted class means.

For readers concerned with the mechanics of the procedure, the essential details of the more unusual features follow below.

THE DUMMY REGRESSORS. The method of generating dummy

regressors in the preliminary phase can be best indicated by an example:

	x_1	x_2	x_3	x_0
Class 1	1	0	0	1
Class 2	0	1	0	1
Class 3	0	0	1	1
Class 4	0	0	0	1

The center columns of the example show the values of the three dummy regressors x_1, x_2, and x_3 which are introduced for a classification with four classes. Any individual from class 1 is scored one for the regressor x_1, zero for the regressor x_2, and zero for the regressor x_3. Any individual from class 2 is scored zero for x_1, one for x_2, zero for x_3, and so on.

A multiple regression fitted on x_1, x_2, and x_3 together with other regressors including x_0, the dummy regressor for fitting the intercept β_0 shown later in Equation II provides all the information necessary for fitting the effects of the four classes. The sums of x_1, x_2, and x_3, for example, give the numbers in the first three classes

$$\sum_j x_{1j} = n_1, \sum_j x_{2j} = n_2, \sum_j x_{3j} = n_3$$

The sum of the dummy regressor x_0 likewise gives the total number of observations n. This, in turn, together with the numbers in the first three classes, provides the number in the fourth class:

$$n_4 = n - n_1 - n_2 - n_3$$

If we let $\hat{\mu}_1$, $\hat{\mu}_2$, $\hat{\mu}_3$, $\hat{\mu}_4$ denote the adjusted means for the four classes and put μ_1, μ_2, μ_3, μ_4 for the four corresponding expectations the homogeneity hypothesis for these classes may be written as H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$. The F ratio for testing this hypothesis is identical with the F ratio for testing the regressor subset null hypothesis H_0: $\beta_1 = \beta_2 = \beta_3 = 0$. It is the F ratio for the latter form of H_0 which is computed in the analysis.

In the final phase of the analysis, the adjusted means for the four classes are obtained from the coefficients b_1, b_2, b_3 fitted for

the dummy regressors and from $\bar{y}$, the overall mean of the dependent variable y. The equations used for this are

$$\begin{aligned}
\hat{\mu}_1 &= \bar{y} + (1-\bar{x}_1)b_1 - \bar{x}_2 b_2 - \bar{x}_3 b_3 \\
\hat{\mu}_2 &= \bar{y} - \bar{x}_1 b_1 + (1-\bar{x}_2)b_2 - \bar{x}_3 b_3 \\
\hat{\mu}_3 &= \bar{y} - \bar{x}_1 b_1 - \bar{x}_2 b_2 + (1-\bar{x}_3)b_3 \\
\hat{\mu}_4 &= \bar{y} - \bar{x}_1 b_1 - \bar{x}_2 b_2 - \bar{x}_3 b_3
\end{aligned} \qquad \text{[I]}$$

which we may write

$$\begin{bmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \\ \hat{\mu}_3 \\ \hat{\mu}_4 \end{bmatrix} = \begin{bmatrix} 1 & 1-\bar{x}_1 & -\bar{x}_2 & -\bar{x}_3 \\ 1 & -\bar{x}_1 & 1-\bar{x}_2 & -\bar{x}_3 \\ 1 & -\bar{x}_1 & -\bar{x}_2 & 1-\bar{x}_3 \\ 1 & -\bar{x}_1 & -\bar{x}_2 & -\bar{x}_3 \end{bmatrix} \begin{bmatrix} \bar{y} \\ b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

or, more briefly for convenience, as

$$\hat{\mu} = Cb$$

The mathematics for getting these can be seen by rewriting the regression model

$$y_j = \beta_0 x_{0j} + \beta_1 x_{1j} + \ldots + \beta_r x_{rj} + \epsilon_j \qquad \text{[II]}$$

for the jth observation, in the familiar deviations-from-means form

$$y_j = \mu + \beta_1 \tilde{x}_{1j} + \ldots + \beta_r \tilde{x}_{rj} + \epsilon_j$$

where $\tilde{x}_{ij} = x_{ij} - \bar{x}_i$, $i=1, \ldots, r$, $j=1, \ldots, n$, and $\bar{x}_i$ is the mean of the ith regressor. The "true" adjusted mean μ_1 in the first class, for example, is given by taking the expectation for any individual in the first class and setting all of the regressors beyond the three concerned at their mean values (i.e. by putting $\tilde{x}_{ij} = 0$, $i > 3$). This gives

$$\mu_1 = \mu + \beta_1 \tilde{x}_{1j} + \beta_2 \tilde{x}_{2j} + \beta_3 \tilde{x}_{3j}, \text{ for all } j \text{ in class 1}$$

or

$$\mu_1 = \mu + (1-\bar{x}_1)\beta_1 - \bar{x}_2 \beta_2 - \bar{x}_3 \beta_3$$

On replacing the parameters by their estimates this gives

$$\hat{\mu}_1 = \bar{y} + (1-\bar{x}_1)b_1 - \bar{x}_2 b_2 - \bar{x}_3 b_3$$

which is the first equation of the entire Equation I. The remainder come from working in the second, third, and fourth class in the same way.

The variance-covariance matrix $V(\hat{\mu})$ of the adjusted class means is then obtained using

$$V(\hat{\mu}) = C \; V(b) \; C'$$

where $V(b)$ is the variance-covariance matrix of $\bar{y}$ and b_1, b_2, b_3, which is available from the regression analysis.

INTERACTIONS. In testing for the importance of adding an interaction subset, a modification is made to the add-search step. If either or both of the main effects involved are not already in the regression equation, these are included among *in* subsets before adding the interaction subset. If the interaction subset qualifies for addition to the regression equation (based on the F ratio for the interaction subset alone but adjusted for all of the others), both of the main effects are added to the regression equation along with the interaction if they are not already in.

Operating on the same principle, a drop-search step does not "look" at any main effect subset with a view to dropping it so long as it is involved with any interaction still being retained in the equation. Once an interaction is dropped, however, any main effects depending on it alone for inclusion in the regression equation then become subject to dropping, depending on their own F ratios.

In the last phase of preparing the final search results for presentation, special steps are again taken for interactions. For an interaction between two concomitant regressors x_1 and x_2 two conditional regression lines are determined in the form

$$\hat{y} = a + bx_2$$

These show the regression of y on x_2 at a low value x_{1L} and at a high value x_{1H} of x_1. For the rationale of doing this, see, for example, *Experimental Designs.*•

The intercept a and slope b of these lines are determined by substituting x_{1L} and x_{1H} respectively for x_1 in the equation

$$\hat{y} = \bar{y} + b_1(x_1 - \bar{x}_1) + b_2(x_2 - \bar{x}_2) + b_3(x_3 - \bar{x}_3) \qquad \text{[IIA]}$$

•COCHRAN, W. G., AND COX, G. M.: *Experimental Designs*, 2nd ed. New York, Wiley, 1957, sect. 5.28.

where $x_3 = x_1x_2$. Thus at $x_1 = x_{1L}$, for example, the values of a and b are given by

$$a = \bar{y} + b_1(x_{1L} - \bar{x}_1) - b_2\bar{x}_2 - b_3\bar{x}_3$$

and

$$b = b_2 + b_3x_{1L}$$

For an interaction between two classifications with p and q classes respectively, the procedure estimates adjusted class means $\hat{\mu}_{ij}$ for each of the subclasses. The equations for doing this for a 2×3 classification, for example, are

[III]

$$\begin{aligned}
\hat{\mu}_{11} &= \bar{y} + (1-\bar{x}_1)b_1 + (1-\bar{x}_2)b_2 - \bar{x}_3b_3 + (1-\bar{x}_4)b_4 - \bar{x}_5b_5 \\
\hat{\mu}_{12} &= \bar{y} + (1-\bar{x}_1)b_1 - \bar{x}_2b_2 + (1-\bar{x}_3)b_3 - \bar{x}_4b_4 + (1-\bar{x}_5)b_5 \\
\hat{\mu}_{13} &= \bar{y} + (1-\bar{x}_1)b_1 - \bar{x}_2b_2 - \bar{x}_3b_3 - \bar{x}_4b_4 - \bar{x}_5b_5 \\
\hat{\mu}_{21} &= \bar{y} - \bar{x}_1b_1 + (1-\bar{x}_2)b_2 - \bar{x}_3b_3 - \bar{x}_4b_4 - \bar{x}_5b_5 \\
\hat{\mu}_{22} &= \bar{y} - \bar{x}_1b_1 - \bar{x}_2b_2 + (1-\bar{x}_3)b_3 - \bar{x}_4b_4 - \bar{x}_5b_5 \\
\hat{\mu}_{23} &= \bar{y} - \bar{x}_1b_1 - \bar{x}_2b_2 - \bar{x}_3b_3 - \bar{x}_4b_4 - \bar{x}_5b_5
\end{aligned}$$

where $\{x_1\}$, $\{x_2, x_3\}$, and $\{x_4, x_5\}$ are the subsets of dummy regressors for the two main effect subsets and the interaction subset respectively. The justification of these equations follows along the same lines as that of the equation $\hat{\mu} = Cb$ above, and the variance-covariance matrix is obtained in the same way.

In the case of an interaction between a classification with p classes and a concomitant regressor x, a conditional straight line regression is computed within each class. The method by which this is done can be explained as follows: By putting $x = 0$ or 1 we could, if we wished, generate $2 \times p$ adjusted values $\hat{\mu}_{ij}$. If we take a case with $p = 3$ classes, these would be given by the equations in Equation III. On consideration, it is clear that a regression line for each class could then be obtained as

$$\hat{y}_j = a_j + b_jx \qquad j = 1, 2, 3$$

where $a_j = \hat{\mu}_{2j}$ and $b_j = \hat{\mu}_{1j} - \hat{\mu}_{2j}$ since $\hat{\mu}_{2j}$ is the value a_j of $\hat{y}_j$ at $x = 0$ and $\hat{\mu}_{1j}$ is $a_j + b_j$ at $x = 1$.

From this it is seen that the required regression line estimates can be obtained directly in the form $b^\bullet = (a_1, b_1, a_2, b_2, a_3, b_3)'$

by changing the coefficient matrix C in Equation III to

$$C^\bullet = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{bmatrix} C$$

and using the direct equation

$$b^\bullet = C^\bullet b$$

The variance-covariance matrix is then obtained from

$$V(C^\bullet) = C^\bullet \; V(b) \; C^{\bullet\prime}$$

where b is the vector of regression coefficients in the right-hand side of Equation III and $V(b)$ is their variance-covariance matrix, both of which come from the multiple regression analysis.

SPECIAL CARE NEEDED IN "FIXING" MATHEMATICALLY DEPENDENT REGRESSORS. In obtaining the adjusted means $\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3, \hat{\mu}_4$ for the four category classes in Equation I, for example, the other regressors $x_4, x_5, \ldots, x_r$ have been implicitly fixed at (set equal to) their mean values $\bar{x}_4, \bar{x}_5, \ldots, \bar{x}_6$. This is appropriate provided each of these regressors is mathematically independent of the others. If any regressor is mathematically determined by the others this has to be modified. Thus if x_6, say, is a regressor for interaction between x_4 and x_5, that is, $x_6 = x_4.x_5$, once x_4 and x_5 are set equal to their means, x_6 is automatically fixed at $x_6 = x_4.x_5$. Each of the equations in Equation I would then have to include an extra term $(\bar{x}_4\bar{x}_5 - \bar{x}_6)b_6$ on the right hand side. The matrix C would have an extra column with $\bar{x}_4\bar{x}_5 - \bar{x}_6$ in every row and the column vector b would include b_6 as an extra element.

Similar additions would need to be made in Equation IIA and in Equation III. Corresponding adjustments would then appear in the intercept values a and a_j.

THE ADD-STEP. Let $Ab = g$ denote the initial normal equations which have been solved in fitting the current regression equation

and $A_n b_n = g_n$ the new ones to be solved in fitting an additional *out* subset. The new equations can be written in more detail as

$$\begin{bmatrix} A & A_a \\ A'_a & A_{aa} \end{bmatrix} \begin{bmatrix} b^\circ \\ b_a \end{bmatrix} = \begin{bmatrix} g \\ g_a \end{bmatrix}$$

where A_a, A_{aa}, g_a denote the additional sums of products (and squares); b_a denotes the added regression coefficients; and b° now denotes the modified values of the previous coefficients b. The F ratio for scoring the importance of the additional subset is obtained from

$$F = (N/r_a)/\left\{(E-N)/(f-r_a)\right\}$$

where $N = g'_a C_{aa} g_a$ E and f are the error sum of squares and its degrees of freedom from the initial analysis, and r_a is the number of regressors in the subset being tested for addition.

The numerator term N is the additional regression sum of squares, and the matrix C_{aa} is the relevant part of the new inverse matrix as shown in

$$C_n = \begin{bmatrix} C^\circ & C_a \\ C'_a & C_{aa} \end{bmatrix} = A_n^{-1} = \begin{bmatrix} A & A_a \\ A'_a & A_{aa} \end{bmatrix}^{-1} \qquad \text{[IV]}$$

The major part of the add-search is in computing the C_{aa} matrices for the addition of each *out* subset using

$$C_{aa} = \left\{A - A'_a A^{-1} A_a\right\}^{-1}$$

When a subset is added at the end of an add-step, the new equation is obtained by computing

$$C_n = A_n^{-1} \quad \text{and} \quad b_n = C_n g_n$$

that is, by solving the regression problem afresh.

THE DROP-STEP. Let $Ab = g$ denote the initial normal equation as before, and $A_n b_n = g_n$ the new ones to be solved after dropping

an *in* subset. The *initial* equations now can be written in more detail as

$$\begin{bmatrix} A_n & A_d \\ A'_d & A_{dd} \end{bmatrix} \begin{bmatrix} b^\circ \\ b_d \end{bmatrix} = \begin{bmatrix} g_n \\ g_d \end{bmatrix}$$

where A_d, A_{dd}, g_d, and b_d are for a particular *in* regressor subset to be drop tested. The F ratio for scoring the importance of the *in* subset to be tested is obtained from

$$F = (N/r_a)/(E/f)$$

with terms given as above except that now $N = g'_d C_{dd} g_d$ and C_{dd} is defined as in

$$C = \begin{bmatrix} C^\circ & C_d \\ C'_d & C_{dd} \end{bmatrix} = \begin{bmatrix} A_n & A_d \\ A'_d & A_{dd} \end{bmatrix}^{-1}$$

The drop-search is accomplished quickly since the submatrix C_{dd} for each drop-test is already available from the previous computation of C.

When a subset is dropped at the end of a drop-step, the new equation is obtained by computing

$$C_n = A_n^{-1} \text{ and } b_n = C_n g_n$$

It should be remarked here, perhaps, that the inversion $C_n = A_n^{-1}$ at the end of the add-step and drop-step could alternatively be accomplished by "correction" steps similar in style to Equation IV applied to the previous inverse $C = A^{-1}$, and these would possibly be preferable on a desk calculator. A direct inversion of an $n \times n$ matrix, however, can be carried out on an electronic computer with no more effort than that of multiplying two $n \times n$ matrices and thus is quick and more accurate with this type of equipment.

TRANSFORMATIONS. In addition to generating product regressors for interaction, other transformations of the initial set of regressors may, of course, also be introduced. These may be usefully handled either as "junior" regressors which, like the product

interaction regressors, are tested and fitted only in the company of their "parent" regressors, or as "senior" regressors without these kinds of restrictions. A transform defined as the square x^2 of an initial regressor x would usually be treated as a junior regressor, for example, and not be tested or fitted without x also being present. On the other hand, a new regressor obtained as the logarithm of an initial regressor would be treated as a senior regressor and would be tested and fitted with or without the parent, depending on the parent's importance. With two such regressors, the fitting of one usually would tend to eliminate the other.

Discussion

ADDED UNCERTAINTIES OF A SEARCH PROCEDURE. A classical worry about this type of a search analysis is the possibility of fitting an untoward number of spurious regressions. If in fact the overall null hypothesis (all regressions zero) were true, this indeed would be a serious concern. The fitting of a spurious regression at each step would not only be a mistake in itself, but, by removing it from error, the search feature would bias the error mean square downward. This would produce added error probabilities (type 1) relative to a corresponding nonsearch procedure.

The answer, however, is not necessarily to be always more cautious. If in fact all of the true regressions were well away from zero, the search feature would produce an additional loss of power. Failures to discover real regression effects would not only be mistakes in themselves, but, by failing to remove them from error, the search feature would leave the error biased upward. This would produce added error probabilities (type 2) or decreases in power relative to a corresponding nonsearch procedure. A real situation like this would call for smaller critical F ratios.

In considering questions of this kind, it makes rough sense at least to think in terms of unimodal prior distributions for the standardized true regression coefficients $\beta'_1 = \beta_1/\sigma_{b_1}$, . . ., $\beta'_b = \beta_b/\sigma_{b_b}$ giving each a prior mean zero and an *average* prior variance of σ_{β}'. The prior standard deviation σ_{β}' may be thought

of as a rough measure of the overall true regression involved. In the extreme null and alternative cases mentioned above, the true regression according to this measure is zero ($\sigma_{\beta}'=0$) and very large ($\sigma_{\beta}'>>0$) respectively.

In practice, the amount of true regression will seldom approach close to zero as with the null hypothesis, nor will it approach especially high extremes. In such cases the added propensity for error in the search procedure will be far less serious than it might at first appear. If the investigator chooses his critical F ratios with the nature of his regressors in mind, he will naturally tend to choose small ones (say at the 5% or 10% level) if the true regression σ_{β}' is anticipated to be high and larger ones (say at the 1% or .1% level) if the true regression is anticipated to be low. Thus if small F ratios have been chosen for the search, the consequent downward bias in the error mean square will tend to be offset by the chances of the error containing more true regression since σ_{β}' is up. If large F ratios have been chosen for the search, the consequent upward bias in the error mean square by failing to remove enough regression will tend to be offset in the opposite way. With σ_{β}' small, there probably will be very little true regression to be removed from the error in the first place.

Notwithstanding these considerations, it would seem good practice, of course, to run the analysis again with larger critical F ratios if there is any independent evidence that the error mean square is on the low side. By the same token it should be run again with smaller critical F ratios if there is evidence that the error mean square is on the high side.

Also, by rough analogy with a similar situation in multiple comparisons,* there would be grounds for severely discounting the significance of any *individual* effects if evidence of *overall* heterogeneity gets too low. A rough assessment of this is given by computing a crude adjusted F ratio for total regression. An adjusted mean square for regression is first obtained from

$$MSR_a = (r_1 MSR + r_2 MSE)/(r_1+r_2)$$

*See DUNCAN, D. B.: A Bayesian approach to multiple comparisons. *Technometrics*, 7:171-222, 1965, Fig. 1.

where MSR and MSE are the mean squares for total regression and error, and r_1 and r_2 are the number of regressors fitted and not fitted, respectively. The crude F ratio is then computed as

$$F_a = MSR_a/MSE$$

Any tendency for this ratio to drop below 2.5 would call for caution.

This check is like the preliminary F ratio test in the Fisher protected least-significant-difference rule often used in a multiple comparisons problem. In both situations, the observed ratio will tend rarely to go below the critical value of 2.5. Nevertheless, in case it does, this rough check is an important one.

Much more research is needed on the problems just raised, and work in this direction is going forward. In the meantime there are so many advantages to the search approach that, even with the present roughness of the inference aspects, there is a considerable gain.

Advantages of the Overall Procedure. The main advantage of the method lies in the large number of regressors it can explore. In dealing with problems by comparable nonsearch techniques, large numbers of potential regressors often have to be arbitrarily rejected from the study at the outset. For many of these, there is no assurance on prior grounds that the regression is negligible. Thus, for study after study, the same possibly important effects can be missed again and again, simply because they are never looked at. Each time, they are eliminated arbitrarily at the outset by the competition for the limited number of regressor "berths" that can be provided in a nonsearch analysis.

In passing, it should be noted that the arbitrary rejection of possibly important regressors from a nonsearch analysis can lead to strong upward biases in the error mean square. This in turn can enhance the advantage of the search procedure over the nonsearch approach.

Some other considerations of a different nature also are worth stressing. At any one step, and particularly at the last which is of most interest, the procedure has the full character of a multiple covariance analysis. Each of the fitted effects, including both the class effects and the concomitant regression effects, is an *adjusted* one, taking into account the effects of all other items fitted. Furthermore, the form of the analysis is *nonorthogonal*. This feature allows all of the adjustments to take place properly, whether or not the observations are distributed equally among subclasses.

Although the investigator, in the interests of efficiency, should plan as much balance in the design as possible, the nonorthogonal analysis capability can be a considerable convenience. Indeed, because of the way data often must be obtained, particularly in clinical situations, this approach may be an inescapable necessity. In such cases, the method described here may be a useful addition to the investigator's armamentarium for evaluating the data at hand.

APPENDIX 8B

A General Multiple Covariance Search Procedure: Guide to Use of Computer Program

The statistical considerations underlying the general multiple covariance search procedure have been described in Appendix 8A, which also includes a summary account of the computer program implementing the procedure. The following is a detailed practical guide to the computer program designed for use in conjunction with the preceding account.

Language and System

The program is written in Fortran II for the IBM 7094 with 32K core storage capacity. It has run successfully on the Fortran Monitor System (FMS) and on the Johns Hopkins University Applied Physics Laboratory version of the Share Operating System (SOS).

The sample setup shown in Figure 8B-1 includes system control cards appropriate for FMS. The operator should be notified to mount a scratch tape on the tape drive in the user's system corresponding to logical tape 2. The system control cards will

SAMPLE SYSTEM (FMS) AND PROGRAM CONTROL CARDS FOR 2 PROBLEMS

```
123456789 123456789 123456789 123456789 123456789 12345678 COLUMN
*      JOB     4388,C,*MEDIC,05,9999,9999   UHLENHUTH, MCOV
*     PLEASE MOUNT TAPE NUMBER 661616 ON A5 AND SAVE AFTER RUN.
*     PLEASE MOUNT SCRATCH TAPE ON B-3.
*     PAUSE
*     XEQ
*     LABEL  (USE THIS CARD IF YOU WISH OBJECT DECK LABELED.)
(PLACE ANY FORTRAN SOURCE DECKS HERE.)
(PLACE ANY BINARY (OBJECT) DECKS HERE.)
*     DATA
TOFRANIL, TOT WTD SCL, 4 WK CHANGE, SEARCH
 42  1 11  3 12  1118
          1        1  1  1     1
 4.08  3.23  2.85  2.62  2.48  2.36  2.30  2.23  2.19  2.14  2.12
(ERROR MEAN SQUARE CARD IS BLANK IN THIS EXAMPLE.)
116  8  1  2 12155 73 74 75 76 77 78
(ANALYSIS LIST CARD IS OMITTED IN THIS EXAMPLE.)
  2MEDICATION    2      1  2        1.       2.
  4SEX           4      1  2        1.       2.
  6ITEM 18,SCL1  6      1  3        1.       2.       3.
 13 2 × 3        2  3   2  2        0.       0.
 14 2 × 4        2  4   3
 15 2 × 5        2  5   2  2        0.       0.
 16 2 × 6        2  6   3
 17 2 × 7        2  7   2  2        0.       0.
 18 2 × 8        2  8   2  2        0.       0.
 19 2 × 9        2  9   2  2        0.       0.
 20 2 × 10       2 10   2  2        0.       0.
 21 2 × 11       2 11   2  2        0.       0.
 22 2 × 12       2 12   2  2        0.       0.
 23 4 × 8        4  8   2  2        0.       0.
 2414 × 8       14  8   2  2        0.       0.
TOFRANIL, TOT WTD T/S, 4 WK CHANGE, SEARCH
 42  1 11  3 12  1118
          1        1  1  1     1
 4.08  3.23  2.85  2.62  2.48  2.36  2.30  2.23  2.19  2.14  2.12
(ERROR MEAN SQUARE CARD IS BLANK IN THIS EXAMPLE.)
118  8  1  2 12155 73 74 75 76 77 78
(ANALYSIS LIST CARD IS OMITTED IN THIS EXAMPLE.)
  2MEDICATION    2      1  2        1.       2.
  4SEX           4      1  2        1.       2.
  6ITEM 18,SCL1  6      1  3        1.       2.       3.
 13 2 × 3        2  3   2  2        0.       0.
 14 2 × 4        2  4   3
 15 2 × 5        2  5   2  2        0.       0.
 16 2 × 6        2  6   3
 17 2 × 7        2  7   2  2        0.       0.
 18 2 × 8        2  8   2  2        0.       0.
 19 2 × 9        2  9   2  2        0.       0.
 20 2 × 10       2 10   2  2        0.       0.
 21 2 × 11       2 11   2  2        0.       0.
 22 2 × 12       2 12   2  2        0.       0.
 23 4 × 8        4  8   2  2        0.       0.
 2414 × 8       14  8   2  2        0.       0.
FINISH
123456789 123456789 123456789 123456789 123456789 12345678 COLUMN
```

FIGURE 8B-1.

not be discussed, since the requirements of the user's system, rather than this program, will determine these cards. "Control cards" hereafter will refer to program control cards, which the user's system will regard as data. "Data" henceforth will refer to the data proper (from subjects) to be analyzed.

Data Input

The program calls for data stored on a binary tape, that is, a tape produced by the Fortran statement WRITE TAPE *i*, which the program reads with a READ TAPE *i* statement. The operator should be notified to mount the data tape on the tape drive in the user's system corresponding to logical tape 9.

The first record on the tape should contain a linear array of data names, each consisting of two adjoining six-character words. This list serves as a header to identify the tape and to assure that the operator has mounted the correct tape. The data names appear in the print-out to identify the results attributable to the variables specified for analysis.

Each subsequent record on the tape should contain the data for a single subject. Each of these records may contain up to 500 data variables. A variable must appear in floating point mode in the same data location on every record.*

The tape record for each subject should include all the dependent variables and independent variables likely to be analyzed. The independent variables may include classifications and concomitant regressors.

The program reads the data from each record in turn into a linear array. From this array, it selects a dependent variable and independent variables for the analysis, as specified by the user, and stores these variables sequentially in another linear array. This array may be visualized as a list of data.

Now the list of variables is expanded to accommodate dummy regressors for each classification. Each group of dummy regressors numbers one less than the number of classes, as detailed in Appendix 8A. At this point it becomes more convenient to think

*A program for producing such data tapes from data cards is available on request.

in terms of regressor subsets, each subset of one or more regressors representing one of the original data variables.

The program next performs any transformations required. These are entirely under control of the user and are of three general types: (a) scoring dummy regressors for subsets to be treated as classifications, (b) transforming subsets in their original positions on the data list, and (c) generating new subsets to be stored in successive locations beyond the last datum originally on the list. Such new subsets include, for example, the scores for interaction between two of the original subsets, when both original subsets should be preserved as main effects in the analysis. Significance tests are performed on such new subsets jointly with the subsets from which they were generated. The data list now appears as below:

| Regressor subsets read from tape and transformed | Newly generated subsets |

This program regards a value of 9999. in a data location as a signal indicating a missing datum and not as a legitimate data value. It rejects the subject's entire data and automatically reduces by one the number of subjects for the analysis. The program, however, retains subjects with missing data not involved in the current analysis.

Restrictions

The number of subjects is limited to 999. The number of regressors is limited to 79. This set of no more than 79 regressors includes any new regressor subsets generated by the program. The dependent variable constitutes an additional subset of one variable, for a total of 80. Execution will stop and an alarm will appear on the print-out if the program attempts to fit an equation containing more than 30 of the 79 regressors in the pool.

The total number of transforms produced in a single analysis is limited to 80. The total number of constants used in these transforms is limited to 150.

The program will process first-order and higher-order interactions through the main multiple covariance analysis, including the search feature. The following additional results can be computed by hand as shown in Appendix 8A: for first-order interactions, sample lines for the interaction between concomitant regressors; for higher-order interactions, the same, plus all category effects.

Program Control Cards

1. TITLE CARD. This card may contain in columns 1 through 72 any information for identification of the problem. This title will appear in strategic locations in the printed output.

2. PARAMETER CARD. The parameters of the problem should be punched, right adjusted, in consecutive three-column fields on this card as indicated below. All numbers should be punched without decimal points, unless noted to the contrary.

Column

1- 3	Number of subjects.
4- 6	Data location containing the dependent variable. This number should indicate the position of the dependent variable on the list on card 6 below.
7- 9	Number of data variables to select for analysis from each record on data tape (exclude dependent variable).
10-12	Number of transforms which will *replace* existing subsets on the list on card 6 below.
13-15	Number of new subsets to generate. This refers to transforms which will be *added* to the existing list of subsets.
16-18	Number of cards containing list of critical F ratios (see card 4 below). Up to two cards may be employed.
19-21	Location on data tape of last datum required for analysis.
22-24	Number of subsets to force into the final equation. If this equals the sum of the entries in columns 7-9 and columns 13-15, then the search feature will be deleted. In this case, omit card 4 below. This number may not exceed 30.

3. OPTIONAL FUNCTION AND OUTPUT CARD. Select the features desired by punching 1 in the column indicated.

Column

3 Print complete data after generating dummy regressors and transforms.
6 Punch complete data after generating dummy regressors and transforms.
9 Print corrected sums of products and square roots of corrected sums of squares.
12 Continue computation after correlation matrix is produced. A blank or 0 in this column will terminate execution after the correlation matrix has been computed and before a search or a regression analysis is carried out.
21 Compute variance-covariance matrix for regressors in the final equation.
24 Compute effects for each category in classification data.
27 Compute value of dependent variable predicted for each subject by final equation.
30 Punch difference between predicted and observed values of dependent variable for each subject.
33 Compute correlations among regression coefficients in final equation.

4. CRITICAL F RATIO CARD(S). This card specifies a list of values for the critical F ratios for adding regressor subsets to the equation or dropping subsets from the equation. These values should be punched with a decimal point in 12 consecutive fields of six columns, so that the first field (columns 1-6) contains the critical F ratio for single-regressor subsets with df = 1 (concomitant regressors and dummy regressors for 2 classes), the second field contains the critical F ratio for two-regressor subsets with df = 2 (3 classes), the third field contains the critical F ratio for three-regressor subsets with df = 3 (4 classes), and so forth. Up to 24 F ratios may be listed on two cards. F ratios should be included for the degrees of freedom appropriate to subsets generated from multiple regressor subsets on the original data list.

Omit this card if search feature will be deleted (see card 2 above).

5. ERROR MEAN SQUARE CARD. This card specifies the error mean square for all significance tests in the search and final analysis. If the experimentally determined error mean square is desired, the card should be left blank. If a specified, fixed error mean square is used, the value should be punched, with a decimal point, in the first six columns of the card. If the value desired is the one determined experimentally in the immediately preceding problem, any negative number may be entered in columns 1-6.

6. DATA LIST CARD(S). This card specifies which data the program should select for analysis from each subject's record on the data tape. Entries on this card should be punched, right adjusted, without decimal points, in 24 consecutive fields of three columns. For example, to include in the analysis data variables in locations 1, 78, 155, and 378 on the data tape, punch on this card 155 in columns 1-3, 78 in columns 5-6, 378 in columns 7-9 and 1 in column 12.

The sequence of entries is entirely at the user's discretion. The total number of entries should be one more (to include the dependent variable) than the number entered in columns 7-9 of card 2 above. This total may not exceed 80.

The program will deal with a blank entry by entering the blank field's own position on the list. For example, if the total number of entries is 16 and columns 6-9 are left blank, then this will have the same effect as a 3 punched in column 9.

7. ANALYSIS LIST CARD(S). This card specifies which subsets should be forced into the analysis without being subject to search. Entries on this card should be punched, right adjusted, without decimal points, in 24 consecutive fields of three columns. These identification numbers refer to the locations of original subsets on the list of card 6 and the locations of newly generated subsets shown on the transformation cards below.

The sequence of entries is immaterial. The total number of entries may not exceed 30. Do not include the dependent variable. The program will deal with a blank field by entering the blank field's own position on the list, as on card 6.

Omit this card if the option is not used.

8. TRANSFORMATION CARDS. Prepare one card for each subset to be transformed and one card for each new subset to be generated. The order in which these cards follow one another is immaterial, except in reference to classifications, powers, and interactions (see below).

All identification numbers for original data subsets refer to their locations on the list on card 6, rather than to their locations on the data tape record. All numbers should be punched right adjusted, without decimals, except as noted.

Column

1- 3 Location to store transformed or generated subset. Newly generated subsets should be stored in consecutive locations beyond the last location on the list on card 6.

After a *new* subset has been generated and stored, it should not be replaced by further transformations since this results in the loss of indexing information essential to testing its significance in conjunction with the subsets from which it was generated. New subsets may be used to produce transformed or generated subsets in other locations, however.

It is not possible to produce a transform from two subsets and eliminate both the original subsets from the analysis. Problems such as this are best handled by writing a special purpose subroutine (see below).

4-15 Name of transformed or generated subset.

16-18 Location of first subset required for transform.

19-21 Location of second subset required for transform.

22-24 Transformation code indicating type of transformation or generation desired (see below).

25-27 Number of constants needed for transform.

28-36 Value of first constant, with decimal.

37-45 Value of second constant, with decimal. Continue entering constants in this fashion through column 72 and then on another card, up to the total number of constants needed for this transform.

If several analyses (problems) using the same data tape should be included in the same run, prepare additional sets of items 1 through 8 and stack them serially.

9. FINISH CARD. Punch FINISH in columns 1-6 and place this card last in the control deck.

Transformation Codes

CODE 1. Classify subset specified in columns 16-18 into the number of classes specified in columns 25-27. The maximum number of classes is 12.

Enter constants in columns 28-36, 37-45, and so forth, to specify the upper limiting data value for each class. Enter as many constants as there are classes. If the entries go beyond column 72, continue on another card in nine-column fields. Punch all values with a decimal point.

This transform generates the dummy regressors, one less than the number of classes desired. After dummy regressors have been generated, the subset should not be submitted for further transformations, except to produce interaction scores, as described under code 2 below.

CODE 2. (Subset specified in columns 16-18 plus constant specified in columns 28-36) multiplied by (subset specified in columns 19-21 plus constant specified in columns 37-45) of the transform card. Punch 2 in column 27. If such a product represents an interaction term, it should be stored as a newly generated subset.

This transform effectively handles subsets of more than one regressor in order to generate interaction scores for classification subsets. Two such subsets should be submitted *first* to the classification transform (code 1), which generates the dummy regressors. Then the same subsets should be submitted to this multiplication transform, with both constants specified as 0.

This transform should *not* be used to raise a subset to a power.

Special considerations apply in generating interactions above first order. Suppose subsets A, B, C, and D are selected from the data tape, and a third-order interaction ABCD is required. For this purpose, the entire set of related first- and second-order

transforms should be specified, each by an appropriate transform card. A higher-order transform always should be produced by multiplying a lower-order transform by one of the original subsets selected from the data tape. For instance, the transform ABCD may be produced by multiplying the transform ABC by the original subset D, but not by multiplying the transform AB by the transform CD. However, if B is a quantitative variable, then the transform ABCD must be produced by multiplying the transform ACD by the original subset B. Furthermore, the transforms related to ABCD should be stored in the list of data for analysis in order of increasing complexity, that is, with all first-order terms coming first, followed by all second-order terms, finally followed by the third-order term. The diagram below will serve as an illustration.

	Selected from data tape with card 6	*Newly generated with one transform card for each*														
Location on list of data for analysis	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Subset	A	B	C	D	Y	AB	AC	AD	BC	BD	CD	ABC	ABD	ACD	BCD	ABCD

Code 3. Raise (subset specified in columns 16-18 plus constant specified in columns 28-36) to power specified in columns 37-45. Punch 2 in column 27. In general, such transforms should be stored as newly generated subsets, with higher powers following lower ones.

Code 4. Subset specified in columns 16-18 multiplied by constant specified in columns 28-36. Punch 1 in column 27.

Code 5. Subset specified in columns 16-18 plus subset specified in columns 19-21.

Code 6. Subset specified in columns 16-18 plus constant specified in columns 28-36. Punch 1 in column 27.

Code 7. Subset specified in columns 16-18 minus subset specified in columns 19-21.

Code 8. Absolute value of (subset specified in columns 16-18 minus subset specified in columns 19-21). In general, such a transform should be stored as a newly generated subset.

Code 9. Natural log of (variable specified in columns 16-18 plus constant specified in columns 28-36). Punch 1 in column 27.

Code 10. Subset specified in columns 16-18 divided by subset specified in columns 19-21.

Special Purpose Subroutine

This program calls a special purpose Subroutine TRNA immediately after reading each subject's record from the data tape. The user may write this subroutine to perform data manipulations not otherwise provided prior to the analysis proper. For example, subjects may be eliminated from the analysis by setting JUMP = 1 under specified conditions within the subroutine. Transforms which are complex or involve other special problems also may be produced within this subroutine. Data should be indexed according to their locations on the data tape. A branch should be included to avoid transforming any missing data code (9999.). A dummy Subroutine TRNA is supplied with the program.