INVESTIGATION OF A MODEL FOR OPTIMIZING PSYCHIATRIC HOSPITAL TREATMENTS

by

Robert Durrell Gifford

B.A., University of California, 1968

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in the Department

of

Psychology

SIMON FRASER UNIVERSITY

October, 1971

APPROVAL

Name: Robert Durrell Gifford

Degree: Master of Arts

Title of Thesis: Investigation of a Model for Optimizing Psychiatric Hospital Treatments

Examining Committee:

L. M. Kendall
Senior Supervisor

R. Koopman
Examining Committee

E. M. Coles
Examining Committee

I. Andrews
External Examiner
Department of Economics and Commerce
Simon Fraser University,
Burnaby, B. C.

Date Approved: December 21, 1971

Abstract

An attempt was made to construct a model which optimizes the efficacy of psychiatric treatments in a large mental institution.

Demographic, symptom and initial diagnosis information was collected for 4237 first admission patients in a large Canadian public mental hospital who were admitted and discharged within a three year period, 1965-1968. They were divided into groups based on actual treatments received. Multiple discriminant analysis was used to determine the separability of the groups. Seven significantly separable groups emerged, each with successful and unsuccessful subgroups, on a 36 item profile.

Explicit formulations were developed from the profiles of the successful subgroups in the analysis. These discriminant functions represented optimal treatment formulations for each of the several therapeutic patterns. They were developed on two-thirds of the sample, then used on the remaining one-third in a quasi-experimental "implementation" of the model to find out if efficacious treatments would be recommended. A patient was recommended to a given treatment pattern if his profile more closely resembled that of persons improved by that treatment pattern than any of the alternative ones.

The hypothesis, that when a patient's recommended and actual treatment patterns coincided he would have a more successful outcome than when they did not coincide, was not upheld. An analysis of the possible sources of error was undertaken, followed by a discussion of the merits of the general idea of the investigation.

A second set of analyses focussed on the prediction of the successful

outcome criterion instead of separation of groups. It was felt that higher order interactions of the profile items might account for more of the variance than the variables did when alone. Consequently a series of multiple regressions were performed which explored the potency of all possible first order interactions among the 36 variables. This general linear model approach isolated the variables and interactions which were most useful in predicting improvement. A discussion of the possible moderator effects of the assigned treatments followed.

The 28 variables and interactions were then used to repeat the experimental paradigm (developing treatment formulations on one sample and cross-validating them on another). Two variants of this procedure, divided-sample and whole-sample, produced results in the direction of the hypothesis.

Acknowledgments

I would like to thank the members of my committee, Dr. Lorne Kendall, Dr. Raymond F. Koopman and Dr. E. M. Coles, for their advice and assistance on this program. The staff of Riverview Hospital, Essondale, British Columbia, graciously permitted use of the hospital records. I owe a special debt to the numerous people who assisted with the statistical analysis, particularly Mrs. Misa Gratton and Mr. C. F. Macurdy.

Table of Contents

List of Tables

"No reason has presented itself to justify receding from the views presented for several years past, of the unsoundness and consequent uselessness of what are called the statistics of insanity. Every year's experience convinces me that those facts regarding this subject, which are capable of being authentically noted, are of too little moment to be worth recording at all, while those circumstances touching the duration, form, symptoms and event of cases, which would be truly important are, from their nature, incapable of being generalized tabularly into even an approximation to the truth...I still find insanity rarely produced from a simple cause, but by a combination or accidental coincidence of causes."

L.V. Bell, Annual Report of the McLean Asylum, 1844

"The real problem is not with inability to classify what is known but rather with the fact that not enough is known."

H. Brill, in Comprehensive Textbook of Psychiatry, 1967

Chapter 1: Introduction and Review

The most basic purpose of a psychiatric hospital is to assist in the patient's struggle to alleviate his mental turmoil. It is usually assumed, rightly or wrongly, that the patient who is hospitalized does not immediately have the capacity to help himself, and so procedures have been developed to give him help from outside himself. Historically, these procedures have taken a great number of forms. An approach has gradually developed which emphasizes determining the identity of the disorder, that it may be specifically and effectively treated. In the last six or seven decades this approach has come to center on a fairly well-defined and nearly universal method, fathered by Kraeplin, of designating the disorder. Once more, the aim of the designation is to provide a guide for apt treatment.

Diagnosis

Psychiatric diagnosis is an important and ubiquitous hospital routine, although its reliability, validity and even morality have been attacked by some investigators. Psychiatric diagnosis is an accepted and generally supported practice among medical practititioners, social welfare agencies and courts of law, as well as by most other workers in mental health. Besides for therapeutic assistance, this support comes for administrative purposes and for dispensing justice. Even those who criticize diagnosis usually agree with the principle of classifying people, if not the every-day practice of diagnosis (e.g. Nathan, 1967). A commonly encountered opinion is expressed by Plunkett and Gordon (1960), who say the present methods of psychiatric diagnosis are decidedly imperfect, but they are

the result of many years of careful observation and thought and are the best available approximation to a perfect system. They also maintain these methods provide a very practical and useful structure, one that could be gainfully used in developing a better one.

Psychiatric diagnoses, as names for disorders and erroneously as names for people, may have powerful effects on a person's life. It should not be forgotten that the strongest implications of the diagnostic epithet are encountered by the individual to whom it is applied, in terms of subsequent treatment and dealings with psychiatrists as well with social welfare agencies, courts and prospective employers. In this study a model for guiding treatment allocation is explored which is in some respects a radical departure from current standard practice.

In this model, the recommended prescription of therapy comes through a multivariate analysis of the characteristics of patients who have been successfully treated in the hospital. A formulation for optimal treatment is based on the multi-dimensional profile of a group of patients who were administered a given pattern of treatment and subsequently had a successful outcome. While it is recognized that the psychiatric treatments employed in this study are not always extremely or directly curative, they typify those presently available and are the ones in widest usage in North America.

In this system a patient would be assigned a treatment pattern by comparing his profile of characteristics to each of the standard successful profiles (one for each pattern of therapy) and recommending for him the pattern which his profile most nearly resembles.

This method is different from established conference methods of diag-

nosis in that it adds a quantified, computer-oriented dimension. A closer look at the meaning and purpose of diagnosis is justified in view of possible questions raised by the model. Of course, in another sense this system is not really different from the ideals of current practice: its aim, like the clinician's, is to produce an effective therapeutic prescription.

The Meaning of Diagnosis

The word itself comes from the Greek, literally meaning "through knowledge" (Skinner, 1961).

The official American Psychiatric Association glossary of terms does not include the word, referring only to the Diagnostic and Statistical Manual (American Psychiatric Association, 1969).

Various other dictionary definitions, including Webster's Third New International (Gove, 1969), English and English's (1958) and Drever's (1964) psychological dictionaries list one or more of the following elements of meaning attached to diagnosis: (a) the art or act of identifying a phenomenon from its signs and symptoms, (b) a technical description or classification of a taxonomic entity by its distinguishing characteristics and (c) an analysis of the nature or cause of the phenomenon.

These general technical definitions are not commonly dealt with in everyday psychiatric practice. While their implication is there, the emphasis is more on the practical aspects and the types of diagnosis, a discrimination which is referred to in several other, later works surveyed (e.g. Kolb, 1968 and Redlich and Freedman, 1966).

The first type of diagnosis Levine calls the clinical diagnosis, the "shorthand formulation of the broad general category in which the patient's

reaction belongs." The second is the dynamic diagnosis, which refers to the 'understanding of the forces that are currently operative...in the difficulty, the environmental pressures and internal pressures." The third type of diagnosis is the genetic diagnosis, in which "genetic" refers not to genes, but to the "understanding of the genesis...and development of those forces."

The three technical definitions and the three practical ones share certain basic orientations. In both there is a concern with the antecedents of the present condition and the development of the disease. By which historical facts (or, indeed, combinations of them) are we to know the mentally ill person and how he is different from one who is not mentally ill? Second, there is a shared concern for present classification. Once abnormality has been decided upon, the question becomes, which species of abnormality? Finally, an implication of the definitions' reference to nature, cause and development is that some idea of the future course of the disorder is implied by a good diagnosis.

Which of these concerns is foremost is hard to say: probably in practice each of them has been at one time or another. But the point is that, ideally, they should all be present and work together to synthesize a reliable and valid classification scheme.

Although none of the authors reviewed for definitions of diagnosis recommended the use of multivariate statistics, their conceptions of good diagnostic formulations could hardly demand the assistance of the methods more clearly. Redlich and Freedman (1966), for example, say "the diagnostic process involves stressing relevant data and eliminating irrelevant information." Levine (1961) specifically supports the inclusion of ecological and demographical factors: "The age of the patient, his level

of intelligence, his cultural background, his educational background, his geographical location...all are of importance." Aldrich (1967)notes that a satisfactory psychiatric diagnosis usually requires more information than can be obtained just from the direct psychiatric, psychological or physical evaluation of the patient. Information about the social, familial, occupational and educational experiences of the patient are also of importance.

The diagnostic conception outlined here would seem to be a very difficult task for the psychiatrist to carry out well for one patient, let alone a caseload that may extend to the hundreds in a large state mental hospital. Nevertheless, the ideal of looking at the whole case, stressing the relevant portions at the expense of the irrelevant portions, etc., seems worthwhile.

The memory and computational capacity of the digital computer could be of some assistance in the diagnostic task (Engelsman et al., 1968). Certain multivariate methods are perfectly suited to the task of objectively looking at all the information available and weighting individual parts of it according to instructions. These same methods can also be of assistance in developing the instructions.

In an article comparing the theoretical potential in this area of the computer and the clinician, Holtzman (1960) makes a relevant assessment. He asserts that diagnosis is a threefold process: information collection, processing and interpretation. In Holtzman's estimation the clinician is the better in the first and the third of these, and the computer is better in the second. He feels that in the future the computer may supplant the clinician in the third area, provided we are able to make some headway in specifying the rules by which the clinician operates. Bahn (1967), com-

menting in a symposium on psychiatric epidemiology, adds "multidimensional data, although more complex to handle, are more valid operationally [than single variables] and increase the amount of objective information... for the crystallization of syndromes and the eventual development of a valid typology."

The Purpose of Diagnosis

Why are diagnoses made? What purpose is served by assigning a word or two to a severely troubled person when the consequences can be so serious (vide Szasz, 1963 and Rubin, 1965)? Apparently diagnoses are made for a number of reasons, by no means all of which are directly related to the patient's welfare. Redlich and Freedman (1966) point out that diagnoses seem to shift according to the purpose for which they are made, and the situation in which they are made. Some of the purposes they mention are etiology, prognosis, therapy, institutional management and research. Szasz (1961) remarks that diagnoses may also vary according to the amount of welfare benefits attached to the particular diagnosis, especially in military hospitals in the United States.

While, then, diagnosis has several purposes and may vary according to the situation (in addition to its variability due to unreliability), these purposes are not the central one. At least for the clinician and his patient, the main reason for a diagnosis is to indicate the pattern of treatment which is most likely to result in the patient's improvement. Aldrich (1966) and others stress the close relationship which should exist between diagnosis and therapy. Kolb (1968) adds that "a comprehensive psychiatric diagnosis is analogous to the diagnosis which the internist aims to construct in order that his treatment may be rationally directed."

The primary task of a mental hospital is the rehabilitation of the

patient; all other professional work is, or should be, peripheral to rehabilitation and directly aimed at supporting it through improvement of its methods, techniques and knowledge. One statement that may be made with certainty in the mental health field is that knowledge of the cause, nature or cure of most of the maladies is meagre. Consider that there is actually no known cure for schizophrenia: it can often be controlled through phenothiazines, but if the patients stop taking the drugs the possibility of a relapse is high.

The conclusion of Redlich and Freedman on the purpose of diagnosis: "Obviously, we diagnose to satisfy our need for systematic scientific presentation, but, pragmatically, physicians diagnose to obtain guidelines for treatment. Diagnosis is implicitly and explicitly in the service of therapy. We agree with Karl Menninger that diagnosis and therapy cannot be separated from each other."

Criticisms of the Diagnostic System

That diagnostics, as practiced, is far from ideal is well documented. There are many papers in the literature which begin with a statement of the variability of diagnoses, inter-cultural definitional problems, blurring of categorical boundaries, etc. But systematic and specific accounts of the complaints against the diagnostic system are not often found. The following is an attempt to summarize the reported sources of error in diagnostics.

The most common charge is unreliability. The majority of studies on diagnosis focus on this problem, and most of them find unreliability. It is not necessary to recount the procedures and paradigms of these experiments in detail, although it should be noted that some authors in the

psychiatric journals have occasionally criticized them for the artificiality of the experimental setting.

In short, unreliability of psychiatric diagnoses may be attributed in roughly equal parts to variability of the observer, patient and disease (Nathan, 1967). In any disease, physical or mental, the disease may be more or less severe. Also, the patient's disease-fighting resources may be greater or lesser. The observer's vision may be colored or obscured by his training, culture, social biases and lack of time to observe as closely as he wishes.

The review of the literature by Nathan cites studies which find unreliability due to nearly every combination of these causes. The usual paradigm is to hold two of the elements in control and let the third vary.

Nathan finds three investigations in which experienced observers of psychopathology differed among themselves in diagnostic formulations based on the same sets of diagnostic cues "because of their own idiosyncratic weighings of these cues." He reports two studies demonstrating that "the same observer may diagnose the same disease entirely differently on different occasions, though presented essentially the same diagnostic data." Presumably, not only do clinicians vary between themselves on weighting of data, but they vary individually from one time to another.

Other illustrations of the observer source of variability come from two studies which found low reliability of diagnoses given the same stimulus-patient (seen personally in one study and a videotape in the other) by different observers. In a recent doctoral dissertation (Lee, 1968), observers were found to exhibit social class biases. A professional actor played "normal" and was given a script which avoided references to his socio-

economic status. The lower the patient's SES was supposed to be, the more often he was diagnosed mentally ill (rather than normal) and the poorer prognosis he was assigned.

With respect to the symptomatology of the patient, Nathan found three experiments. Two of these attributed low reliability to changing symptomatology and one attributed it to changes in the intensity and frequency of the same symptoms.

Another source of variability are the vicissitudes of the disease itself. Partly, of course, this is due to our ignorance concerning the functional psychoses in general. Example; "The twentieth century historian does not know whether the term schizophrenia comprises one entity with one etiology, one entity with several etiologies, a group of similar entities with the same etiology, or a group of similar entities with different etiologies" (Wender, 1963).

Poor reliability has been the major complaint, but not the only one. There has been some discussion recently of the validity of diagnoses. In this discussion the important question is, of course, validity for what?

Thomas Szasz's approach maintains that diagnoses are invalid because they connote nothing (Szasz, 1961). He says that mental illness, especially functional disorders, are labelled with words that "explain everything" and thereby "explain nothing" - exactly like a classic myth. Szasz believes that the whole construct of mental illness is merely an escape device: for the patient it is a way to assign away his "problems in living" to a powerful, mysterious entity, and for the doctor it is a way to avoid really attacking the problem, out of preference for concentration on classification and chemotherapy.

A second approach to the invalidity of diagnoses comes from the defini-

tion. One implication of the threefold aspect of the definition is that an estimate of the patient's prognosis is included. Nathan, however, notes a prevalent feeling that current practices in diagnosis are not, in general, prognostic indications. "Dissatisfaction with current diagnostic procedures on grounds of their doubtful validity centers on the inability of such procedures to permit either prediction of future course of illness and of its response to various kinds of treatment or delineation of its etiology."

A third complaint also stems from the definition. As we have seen, a gooddiagnosis is treatment-indicating. In practice, though, they are not. Pilot studies for the present investigation revealed only mild relations between diagnosis and treatment, with wide individual variation. Phenothiazines, for example, were given to 97% of the schizophrenics, 84% of the neurotics, 88% of the brain damaged and 85% of the affectives and involutionals. Group therapy was given to 42% of schizophrenics, 37% of the neurotics, 43% of the personality disordered, 43% of affectives and involutionals and even 31% of the brain damaged. The other treatments were only slightly more diversified by diagnosis. Certain notions seem to exist about the relationship of some diagnoses to some treatments and even more about the relationship of a treatment to a symptom (e.g. ECT for depression) but nearly all treatment is modified by "the individual aspects of the case." In effect, diagnosis does not strongly indicate a treatment or treatment pattern, but rather seems hardly more important than any other background fact such as marital status or disordered stream of thought in making the therapeutic decision.

In a paper which attacks the validity of diagnosis on this basis and proposes looking most closely at a patient's current behaviors from a learning theory framework, Kanfer and Saslow (1965) suggest that "an effec-

tive diagnostic procedure would be one in which the eventual therapeutic methods be directly related to the information obtained from a continuing assessment of the patient's current behaviors." They further suggest an outline for the collection of diagnostic information that attempts "to achieve definition of a patient's problem in a manner which suggests specific treatment operations."

Nathan criticizes Kanfer and Saslow on the ground that such a procedure might lead to a situation where there was one syndrome for each patient. Nathan, who "tentatively accepts" the medical model, feels the idea of relatively few categories is better than a plethora of quasi-diagnoses. However, Nathan agrees that the still current Kraepelinian emphasis on remote etiology and the idea of the disease having a "course" (Kraepelin, 1919) should be rejected in favor of looking at the patient's recent and current behavior and his environmental situation. In the end, he accepts the current diagnostic labels even though they are "not particularly appropriate or useful." Nathan's purpose, however, is not to create a new diagnostic scheme but to specify and standardize decision-making procedures for arriving at the existing one. His goal is to make possible "reliable categorization of separate psychiatric disease entities from systematically collected, reliably codified and consistently evaluated behavior samples."

Szasz (1957) notes that diagnoses are made by methods which are different from the purposes for which they are made. Diagnoses must be situation-relevant; that is, made and suited to the situation and purpose at hand. The formulations of most modern diagnoses were conceived in quite different settings from typical hospital settings today. A possible solution to this dilemma is an operational, treatment-oriented definition.

The Purpose of the Investigation

So far, an attempt has been made to show that, ideally, diagnoses are disorder designations which specify efficient treatments, and to list some of the deficiencies of the present system. The purpose of this investigation is to make a quasi-experimental test of the proposed treatment-allocation model. This model aims at fulfilling the ideals of diagnosis without falling into the snares of the present system.

From the definition developed earlier, the requirements for a good allocation system are that it be (1) descriptive, (2) classificatory, and especially, (3) treatment-indicating. Further, some indication of prognosis would be valuable. These requirements will be discussed as they relate to the proposed method.

In some physical diseases, "description" might nearly be reduced to the fact that a certain virulent organism had been observed. But in the case of the unknown and probably manifold causes of mental illness, a description must be a hope, in that it hopefully includes the cause of the disease. The best that can be done for an eventually true description of a mental disease syndrome is to begin by including each of the aspects of information concerning it which might have relevance for the condition. A description will probably comprise combinations of certain portions of this information.

Presumably people so identified will fall naturally into some number of categories, despite Nathan's fears of one syndrome per patient. This aspect, a seemingly successful one in the present study, will be discussed below.

A good diagnosis indicates a treatment. A corollary of this probably is that one is sure of the causative agent as well, so that the treatment

may be designed if possible to act directly on the agent, or more precisely, on the reaction to this agent. Nevertheless, it is assumed in this investigation the causes are not previously known. Consequently, it might appear premature to propose a treatment-indicating system. The criticism is probably valid. However, this investigation only seeks to make a true beginning toward the goal it sets. In this quasi-experiment, the allocation model is set up as treatment-indicating a priori, since the model is constructed from the characteristics of people who received the different patterns.

The requirement for an indication of prognosis would seem to be satisfied by the further stipulation that the formulations stem from the successful recipients of a given treatment pattern. Furthermore, the proposed model is not fixed, but the formulations would change with advances in diagnostic and psychotherapeutic knowledge and feedback from results with patients who were treated according to its recommendations.

So much for the definitional ideal side of the purpose. From the practical side, the purpose is to increase reliability and validity in treatment allocation. The information used in this formulation is easily objectifiable information. While any human processing of information is bound to be subject to some alteration, the proposed model uses minimally alterable information. Demographic data are among the best of the objectifiable data. Symptomatology is more difficult, but an attempt has been made (below) to increase the ease of its being rendered reliably.

The problem of validity is attacked essentially by making the formulations operational definitions. In the proposed system, the recommended patterns of treatments are a direct function of the characteristics of

patients.

One by-product of a system similar to the proposed model is that the resultant "labels" would be nothing more than the names of the therapies to be administered. The categories would be treatment group categories instead of the current nosological labels. The stigma of some of the labels might be removed. Very probably, any one treatment group would be much more heterogeneous to the casual observer than are the current diagnostic groups. There might well be patients of any of the former diagnostic categories in one of the new categories.

Through a system of rationally obtained treatment recommendations, some of the reported lack of preciseness in administering them might increase their effectiveness. In a monograph on the culture of a state mental hospital, Dunham and Weinberg (1960) point out that "while the hospital which we studied had in operation most of the accepted therapies, the results seemed to be meagre and unsatisfactory. This was precisely so because all therapies were interwoven with the strands of hospital culture with the result that their value was diminished considerably." Besides being impersonal and uncounselled, the administration of treatments may sometimes be subject to unproven ideas of what "kind of patient" should get which therapies.

In summary, the purpose is to assess the possibilities of a multivariate approach to treatment allocation, one which closely links relevant antecedent conditions to therapeutic choice. Ideal diagnoses, when the physical causative agent and a certain cure are known, are of this sort. The goal is a definitional ideal which is also patently practical because it is itself free of abstract or theoretical concepts and is directly tied to the tasks of the hospital. This investigation recognizes the probable pre-

maturity of such an attempt in this field, but it also realizes that if the results show promise the method can be developed and expanded. For example, as new forms of psychiatric treatment appear it is possible to assimilate them into this system. The formulations themselves would change with the intelligent use of feedback information from the subjects for whom the recommendations were implemented.

Chapter 2: Method

The method used in the first part of this investigation involves, in the most basic terms, employing weighted quantifiable information in a manner which best discriminates a number of groups according to a stated criterion. Multiple discriminant analysis is a sophisticated descendant of the simple weighted scoring key.

Previous Use of the Method

In one of the earliest related studies, a life insurance company used a personal history questionnaire in 1919-1921 to distinguish successful salesmen from unsuccessful ones (Russell & Cope, 1925). The investigators analyzed questionnaires completed by 500 salesmen. They found 12 items which, when weighted, strongly distinguished between the successful and unsuccessful men. Subsequently the company's training failure rate decreased from 90% to 30%.

Because of this initial success other insurance companies (but apparently no one else for a while) completed similar studies (e.g. Goldsmith, 1922). Later a standardized aptitude index with weighted background items was made available to other companies, one which has been extensively used and successfully cross-validated since 1938 (Barnette, 1969).

The technique, bolstered by more sophisticated statistical methods, has been used with positive results in a great variety of settings in recent years. Some uses have been: to select good drivers, to predict achievement in first year graduate school in psychology, to recognize successful parolees in federal prisons and to identity psychotherapy clients who will stay the full term of the treatment. Two recent books have dealt entirely

with the problem of personnel classification using this paradigm (Cronbach & Glesser, 1965 and Rulon et al., 1967). In an application similar to the present one, Ward and Davis (1965) discuss the possibilities of programing a computer to assist in making personnel decisions. They foresee the operator instructing the computer in the rules to decide by, observing the results, and adjusting the decision rules until a desirable set of decisions is being produced.

Typically, some criterion of success in an institution is employed. For example, W. J. Smith et al. (1961) used a 484 item personal history questionnaire to investigate the research ability of a group of petroleum scientists. In this case, three criteria of success were used, two sets of supervisorial ratings and the number of patent disclosures. An item analysis revealed that while the criteria themselves were only moderately intercorrelated, the concurrent validity of the scoring keys developed were all quite significant.

Many such studies have been done in the industrial area. Using another common criteria, Kirchner and Dunnette (1957) employed a weighted application blank to screen office workers for long term vs. short term employment.

Stein (1963), in reviewing the literature in the area of academic prediction noticed certain differences in the strategy of investigators who used this method. The most obvious of these was the type of information used to develop the prediction system. His review showed that four classes of information are commonly used, some much more than others: (a) the subject's previous achievement in a similar setting (e.g. high school grades to predict college success), (b) social background and demographic information, such as in most of the industrial situations described above, (c) psychological information in the form of inventories and other psychological information in the form of inventories and other psychological

tests, intelligence tests, etc. and (d) "transactional" information, data which reflects person-environment relationships.

A direction in which most recent work has been moving is toward multivariate treatment of the weighted scoring key. In the early industrial studies, of course, such an approach was nearly impossible because of the size and number of the calculations involved. It should be noted that it was not difficult because the idea of multivariate statistics was lacking: multiple regression analysis goes back to the beginning of the century and Fisher (1936) introduced the discriminant function nearly 35 years ago. It was rather for the development of the larger digital computers that the technique was forced to wait, especially if a large sample or many variables were contemplated. If the results of the analysis were needed within a certain time period (as for example in the model proposed in the present study, where treatment recommendations are needed early in the hospital stay), the technique was rather unusable.

The successful use of apparently unrelated (either to other sections of itself or to the criterion) information is interesting. This successful use of "unrelated" information lends an optimistic note to the present effort, since some of the information (and more especially, combinations of the information) have not been firmly tied to the success criterion. In a study like that of Safford (1967), in which school performance was successfully predicted from object sorting and concept formation tasks, relationships are being tapped which may have been previously unknown (or even unsuspected) and perhaps indirect, but which undeniably exist. Through such studies an entirely new light may be cast on the theoretical conceptions of the nature of phenomena. Another example, pedestrian but useful and illustrative, is the work of J. T. Smith et al. (1967). These authors

used a combination of sociometric, academic and personality inventory data to predict elementary school success. Among these variables, the most important in the resultant regression equations was IQ, as might be expected, but some of the other items were found to significantly contribute to the predictions.

In a slightly more striking example, Anastasi, Meade and Saunders (1960) were commissioned by the College Entrance Examination Board to investigate the relative value of different types of information for prediction of college success. The authors found the criterion to be more highly correlated with some of the biographical items than with the CEEB verbal and mathematics scores. In agreement with this finding, Stern, Stein and Bloom (1956) found that activity preferences and personality inventory items, etc., could be more important in prediction of college success than more "direct" variables.

Several other authors have found similar relationships between the success criterion and various forms of supposedly indirect information using multivariate techniques (Stone, 1958; Bishton, 1958; Nason, 1958; Mellinger & Haggard, 1959; Michael, Jones & Tremblay, 1959; McGuire et al., 1961; Stein, 1963; Tucker, 1963; Creelman, 1964; McGrath, 1960).

The area of mental health prediction is a difficult one. One of the few places where fruitful work has been accomplished is in the length of patient stay. Lindemann et al. (1959) found five demographic variables (marital status, diagnosis, severity, legal competence and alcoholism) which were important in distinguishing long stay patients. Seaman (1968) found that three common inventories (MMPI, Edwards Personal Preference Scale, Revised Beta Examination) administered during hospitalization had low correlations with length of stay. She was unable to derive multiple

regression equations from them which were of use. However, she discovered that marital status, living arrangement at the time of admission and work history were related to length of stay. She concluded that important determiners of length of stay were external: the demands of home, family and job.

In a review of many attempts to predict prognosis, Fulkerson and Barry (1961) also conclude that psychological tests are of little value. They noted that certain background items, such as severity and duration of the precipitating episode have been more consistently related to prognosis. Their review concludes by stating that single predictors have generally been much less successful combinations, and even when a single one looked good at first it usually failed to cross-validate.

Bucklew (1967a, b) has in a series of multivariate studies employed discriminant functions to show that prediction of various kinds of early hospital information is possible. For example, he predicted symptomatology from case history information in one study. The other study reversed this and showed that case history information is also predictable from symptoms.

Bucklew envisions a mental health clinic where the counsellors employ "personalized" discriminant functions to guide their practice. That is, the items and the assigned weights are chosen by the counsellor in accordance with his beliefs about their value. They are changeable with his changing understanding of their relative worth and to suit different individual clients' cases.

In a symposium on the evaluation of psychiatric treatment, Cole (1964) reports a study of the effectiveness of phenothiazines with acute schizophrenics. Nine hospitals co-operated in the study to standardize treatment practices and pool data. In this case, the discriminant analysis was employed

to determine whether staff ratings could distinguish patients given placebos from those given various phenothiazines, and if they could, to assess which observational and behavioral rating items were most important in making the distinctions.

In the same symposium, Clyde (1964), in giving a final overview to the proceedings, noted that (1) many factors are involved in the evaluation of psychiatric treatment and (2) that dissatisfaction had often been expressed concerning the inadequacy of the statistical methods which had been used in the reported studies. Clyde pointed out the feasibility of discriminant analysis for many problems, that it was appropriate to the complicated interrelationships involved in psychiatric treatment investigations, that several worthwhile computer programs existed which were written specifically with biomedical problems in mind, and finally that while its purpose was primarily to discriminate groups, it was possible to compute recommendations and other criterion scores for individuals.

Three papers are particularly relevant to the present problem. In a speculative article, Kleinmuntz and McLean (1968) propose a system for large-scale psychodiagnostic interviewing by computer. They give a sample algorithm for such a system, using the MMPI. In discussing the possibilities of computer diagnosis, the authors concede certain limitations, at least in the near future. One is the rudimentary state of person-machine communication. Due to the linguistic infancy of computer science, responses must be very simple - usually yes or no, or some other common, expected word. Second, little is known about the process clinicians themselves go through to diagnose disorders, which makes it difficult to reconstruct that process for the computer.

A second study, recently performed in Czechoslovakia (Engelsman et al.,

1968), used multivariate analysis to test the relationship of functional psychotic disorders to items of a rating scale. The rating scale was composed of symptoms and background information. The authors found they could classify 70% of their sample of 180 patients into the diagnosis which they had actually been assigned.

Finally, in an encouraging clinical study, a method of weighted vector differences has been used in the diagnosis of various cardiac diseases. In this study (Klingman & Pipberger, 1967) 8 measurements of the subjects' hearts were obtained and classified by multiple discriminant analysis. Eighty-four percent of the cases were correctly diagnosed (agreed with independent judges who were cardiac specialists). Residents intending to become heart specialists classified an average of 67% correctly.

Multiple Discriminant Analysis

Multiple discriminant analysis is a way of finding the maximum possible separation of several groups on a given criterion from the available information about the groups. The computations are usually sufficiently complex to require a computer. The information must be in continuous variable form. The method also assesses the relative importance of each variable to the resultant dimension which represents maximum separation.

In the beginning there are the scalable scores of individuals. The individuals comprise *a priori* groups. The mean of each variable in each group is calculated. For each group, a vector is computed which is the product of the vector of the group means and the variance-covariance matrix. This vector is of the form $y = b_1x_1 + b_2x_2 + \ldots + b_nx_n + c$. In this equation y is the criterion's value and b is the computed estimate of the coefficient-weight for variable x. The c is a constant.

One discriminant function of this form is obtained for each group. The

n discriminant functions represent centroids in hyperspace which maximize the separation of the groups. Various tests of significance are available to determine whether the separation of groups is such that the null hypothesis that their mean vectors come from the same populations may be rejected (Hope, 1968).

If the discriminant functions thereby appear to be representative of groups which are not from the same population, then newly observed and previously unclassified events or subjects may be classified. The scores of the new subjects on each variable are multiplied by the discriminant weights of each of the functions. He is assigned to the group with the profile most nearly resembling his own profile.

The greater certainty associated with the hypothesis that the several groups do indeed come from different populations, the greater percentage of correct classifications of new observations will be made.

The particular program used in this investigation was developed at UCLA (Dixon, 1967) and revised in May 1969. Slight modifications of it were made during the course of this investigation. These will be discussed later.

The program uses a step-wise method. That is, variables are added or deleted one at a time in this analysis. A variable is included in the set of discriminating variables from a pool of potential discriminating variables according to the following criteria: (a) the one of the pool with the largest F-value, (b) the one which when partialed on the previously entered variables has the highest multiple correlation with the criterion, and (c) the one which gives the largest decrease in the ratio of within to total generalized variances. These three criteria are equivalent. Finally, if in the course of the step-wise procedure a variable's F-value becomes too

low it is deleted.

As this step-wise procedure has come under some suspicion recently, a program option allowing the user to specify the tolerance level for inclusion and deletion of the variables was exploited. Setting this tolerance at a very low level for both inclusion and deletion probably removed the possibility of erroneously eliminating variables which appear weak at one step but would be strong if they were included later. By setting the tolerance level very low (at .00001), the last step of the analysis, when all the variables have entered, becomes equivalent to an analysis performed using all variables simultaneously.

The Samples

The complete sample is composed of first-admission patients admitted to Riverview Hospital between 1965 and 1968.

The sample was fairly heterogeneous, and seemed in most respects to resemble a normal British Columbia population. The modal decade of birth was 1930-40. Four percent were born before 1900, more than a dozen in the 1880's. Three percent were born after 1950. Fifty-five percent are female, forty-five percent are male. Half were married at the time of admission, one-third had never been married and the rest were formerly married. About 85% were born in Canada, and 15% were born in the United Kingdom, Australia, New Zealand and the United States. One-third listed a denomination other than the established Protestant or Catholic faiths, which were represented in a 5:2 ratio. However, the sample is a poorly employed one. Almost two-thirds were unsettled, unemployed or retired. Nevertheless, their educational level is not far below a normal distribution. The average (mean, median and mode) completed educational level was grade 9-11. One percent had no formal education and one percent had

a university degree.

Approximately half these first admissions came to the hospital voluntarily, and the other half were committed by court order.

The information available on these patients which seemed potentially relevant (see Table I of the Appendix) from the survey of related previous studies and a pilot study was coded onto data processing cards from hospital records. It includes several categories: background, admission, mental status examination results, treatments administered, initial diagnosis.

An effort was made to arrange the information contained in a given variable in continuous form. This was relatively easy in some cases, such as educational level, and somewhat more difficult in others, such as religion. However, in each case an effort was made to spread the information along an implicit dimension in the variable. Nevertheless, the difficulty encountered in this task in a few cases led to dichotomous variables: no three or four point scale could be developed. This is frequently the case in multivariate analysis, though Kendall (1957) speaks of the practice as being "a little rough." The symptomatology of the patient, in the form of the mental status examination, was dichotomized. In the exam, the patient received in each of seven spheres, a rating of normal or one of many words used to connote abnormality. In an attempt to render these ratings more simply and perhaps more reliably, they were summarized as normal or abnormal. The religion of a subject was expressed in two ways, to allow two possibly useful dimensions to emerge. In one, the dimension employed was "established" (i.e. Protestant, Catholic and Jewish) religions vs. "unestablished" ones, and in the other it was

Protestant vs. Catholic, the very few Jewish people in the sample being set aside in this latter variable.

The diagnoses were not arranged as a continuum but as a series of variables, one for each diagnosis. The sample included people with diagnoses of organic and functional disorders, as well as addiction and personality disorder.

The sample was randomly split into two sections for this investigation. About two-thirds, 2820, were used for the experimental sample to develop the standard successful profiles and the remainder, 1417, for the cross-validation sample.

Division into Groups

The sample was divided into groups based on the treatment pattern received. Seven categories of treatments are available to the patient at Riverview, and he may conceivably receive any combination, permutation or amount of them. In this study information on permutations and amounts were not available and therefore could not be considered. No distinction could be made between one who received ECT and then group therapy and one who received these in reverse order. Nor could any be made between those receiving few electro-convulsive therapy treatments and those receiving many.

Despite these simplifications, there are many possible combinations of group therapy, ECT, anti-depressants, tranquilizers, insulin shock therapy, behavior therapy and drugs for physical disorders (e.g. anti-convulsants, anti-Parkinsons). However, less than ten instances of the reported use of either behavior therapy or insulin shock therapy prompted the deletion of these treatments from consideration.

A computer program was written to search the sample and list all the combinations of treatments administered. Altogether 33 patterns of the five

treatment categories emerged, including one in which none of the treatments was given. Some of these patterns occurred very infrequently, especially combinations of three or more treatments. The twelve most popular patterns accounted for 85% of the sample.

Therefore, initially, the samples were divided into twelve homogeneous groups and one catch-all group for the 15% who received unusual treatment. Later this division was modified - this will be discussed below.

Inspection of a matrix of diagnoses and assigned treatments from the treatment-listing program showed that each one of the treatment groups is quite heterogeneous in terms of standard diagnosis. The heterogeneity of the current diagnostic system to reliably indicate any given treatment or pattern of treatments.

Construction of Indexes

Two indexes were constructed from already known information. Table 1 summarizes the construction.

The first was an index of social class. In a pilot study of first admission males moderately high correlations between social class and improvement were noted. A score on the index was calculated for each subject. The index is composed of two factors: occupational level and educational level. A somewhat arbitrary decision was made that present occupation has more to do with present social class than educational level attained. Consequently, a person's social class index score was defined as twice his occupational level score plus his educational level score.

The second index was of improvement or successfulness of hospital outcome. A rating of improvement is made for each patient at discharge by the attending physician. However, a necessity was felt to broaden the index. From a social point of view, improvement in the mentally ill is not only a

Table 1

Construction of the Indexes

Index	Component Variables	Weight	Index Range
Social Class	Educational Level: 0 - no formal education 1 - grades 1-4 2 - grades 5-8 3 - grades 9-11 4 - grades 12-13 5 - some university 6 - university degree	1	2-14
	Occupational Level: 1 - unskilled 2 - service, semi-skilled 3 - clerical, sales, skilled 4 - professional, managerial	2	
Improvement	Improvement Rating: 1 - symptoms present, uncontrolled 2 - symptoms present, controlled 3 - symptoms free, no insight 4 - symptoms free, some insight	2	
	Medical Referral: 0 - none 1 - family doctor 2 - psychiatrist, clinic 3 - other hospital	1	2-14
	Employment Status: 0 - unemployed, unemployable 1 - employable, unemployed 2 - to new job 3 - to former job	1	

matter of symptom alleviation, but also should be reflected in the ability of the patient to function well in the social world after discharge.

Consequently a three factor index of improvement was made. The rated level of improvement of the patient, being most important to the meaning of the index, was given a double weight. Single weight value was given to the other two factors, the "employability" of the patient at discharge and the degree of post-hospital medical care he required. This choice of components for the index was aimed at strengthening the validity of the improvement measure from the single four point rating to a 13 point multi-factor index with some more certainly objective components and some reflecting the social aspects of improvement. The correlation between the rating alone and the three factor index was .87. The correlations of the two other components with the index are .53 for the "employability" variable and .28 for the medical referral variable. The intercorrelations among these variables are quite small (rating of improvement-"employability" .13 and the other two below .01).

A further check on the weighting system was performed to determine if variances of the variables were comparable, so that the intended weightings were actually reflected in the indexes. In the social class index (where the components were correlated .25, with education and occupational level correlated with the index .63 and .91 respectively), the variances of the two components were 1.00 and .89. Those of the improvement index components were also in the same range, although that of the medical referral variable was lower than the other two: improvement rating, 1.01; medical referral, .46; and employability, 1.05.

Selection of the Variables and the Separation of Groups

The main criterion for the final selection among the variables was efficiency for the discriminatory purposes of the investigation. Table 1 of the Appendix lists the complete original pool of variables and the final selection. After the main sample had been divided into the 13 groups, it was further subdivided into successful and unsuccessful outcomes on the basis of the improvement index. Decisions were made as to the cut-off scores for success and failure by examination of the constituent variables of the index. On each one an assessment was made of the score which reasonably indicated a division point of success and failure. Then the variables were combined in the manner of Table 1. The range for the index is 2-14 and the cut-off score was set at 8. When this division had been accomplished, approximately two-thirds of the sample had been defined as improved. This, incidentally, was very close to the same proportion which got a discharge rating of "symptom free," as opposed to "symptoms present".

A variable was retained for the analysis if it materially assisted in distinguishing the 13 pairs of groups resulting from the divisions of treatment pattern and success. The original pool of variables consisted of 39; 17 demographic, 7 mental status, and 15 initial diagnoses. Simple correlations among all these original variables and the criterion are presented in Appendix Table 3.

Two kinds of analysis were then carried on simultaneously. They answered the questions: (1) which variables are efficacious in differentiating the groups? and (2) which of the groups can be significantly differentiated? First all the variables were used to try to distinguish all the groups. Some groups would not be distinguished and some variables were useless in trying to do so. A process of amalgamating groups and choosing "good" vari-

ables ensued. Sometimes when two groups had been joined together, rechecking showed that a variable which had been discarded was now valuable in distinguishing the remaining groups.

After several arrangements of groups and variables had been tried, a final combination evolved which seemed to be successful in recognizing and dividing a maximum number of groups with a maximum number of variables.

The final result of the process was seven pairs of groups and 36 variables. These seven groups (see Table 2) were composed of six which were amalgamations of former groups and the catch-all group. The six amalgamations shared a common feature: all of them were combinations of (a) a given pattern of treatments and (b) that same pattern plus group therapy. In other words, group therapy patterns and non-group therapy patterns were not distinguishable, using the final selection of variables. The division into main and cross-validation samples was quite well representative over all groups with respect to proportions of improved and unimproved subjects, although some individual groups were moderately differently represented. A significance test of the difference is, however, inappropriate since the groups by definition came from the same population.

Modification of the Program

The purpose of the investigation was to develop discriminant function profiles for the seven treatment-indicating formulations, profiles of 36 items each.

The main program used for the analysis was modified so that the profiles developed on the main sample could be imposed on the cross-validation sample. When the final selection of groups and variables was complete, a discriminant analysis was performed over the seven pairs of groups in the main

TABLE 2

The Final Treatment Groups

Name		Main Sample N	Main Sample % improved	Main Sample % not improved	Cross-Validation Sample N	Cross-Validation Sample % improved	Cross-Validation Sample % not improved
A	1 or 1+5	898	73.8	26.2	443	80.6	19.4
C	1+3 or 1,3+5	335	75.2	24.8	166	75.9	24.1
D	1+2 or 1,2+5	357	70.0	30.0	182	60.4	39.6
E	1+4 or 1,4+5	305	79.7	20.3	149	69.1	30.9
F	none or 5	263	73.4	26.6	137	67.2	32.8
K	1,2+4 or 1,2,4+5	174	68.4	31.6	92	59.8	40.2
X	all others	488	64.8	35.2	248	75.8	24.2
Totals		2820	71.8	28.2	1417	72.7	27.3

[a] where 1 = tranquilizers
2 = anti-depressants
3 = drugs for physical disorders
4 = electro-convulsive therapy
5 = group therapy

sample. The standard profiles, to be used in making recommendations for "future" patients (the cross-validation sample), were the means standard deviations and co-efficients of discriminant functions for the seven successful subgroups.

This procedure was used so that in the differentiations, both dimensions of separation would be represented (success-failure and one group from another). An alternate method would have been to use only successful groups in the construction of profiles, since only successful profiles were desired. This was tried, and the results were quite the same; less than one percent of the recommendations were different from the two-way profile recommendations.

The investigation must be termed a quasi-experimental one, after Campbell and Stanley (1963). These authors discuss 16 possible research paradigms, together with the advantages, disadvantages and suitability of each for differing types of investigations. According to criteria suggested by Campbell and Stanley, several of these designs are not truly experimental. The principal manner in which the present study does not meet these criteria is that the subjects in the study (the cross-validation sample) are not randomly assigned to treatment. It was impossible to actually apply the treatment recommendations made for the patients by the study.

However, Campbell does not single out the quasi-experimental design in order to condemn it. Rather, he maintains it is a necessary and desirable technique for approaching questions which are otherwise unapproachable with conventional empirical designs.

It is believed this investigation legitimately falls into this category. A truly empiric test would have necessitated the assignment of patients in

the hospital to treatment under a completely untried method. Such a proposal would have certainly been opposed by the medical authorities. Yet the basic idea of the study seemed worthwhile, and even potentially valuable. Apparently the only way to test the idea was to use one of the pseudo-experimental designs for a preliminary test of the general hypothesis. Strong confirmation then might be support for a small field trial.

The program was modified, then, to enable a pseudo-experimental test of the formulations. Specifically, the program was changed so that it would employ means, standard deviations and co-efficients of discriminant functions from the main sample in calculating which group individuals of the cross-validation sample "should" have been in. It was as if the profiles had been developed on a standardizing sample and then were implemented in hospital practice for the 1417 members of the cross-validation sample. As each new patient was "admitted" his profile was compared to the standard successful profiles of the various possible treatment patterns and he was assigned to the group he most nearly resembled. Of course the recommendation would leave the question of group therapy open, and if the patient was assigned to the catch-all group his recommendation would be open - except for the first six patterns (where 80% of all patients are). A patient assigned to the catch-all group would be known to need an unusual pattern of therapy.

Although each patient of the cross-validation sample had of course actually been through the hospital and received some pattern of treatment, this actual pattern was ignored for the moment in making a recommendation according to the proposed system.

The Hypothesis

In general the hypothesis is that cross-validation sample members who

received the same pattern of treatments in the hospital as was recommended for them by this analysis would have higher scores on the improvement index than those who did not receive the same pattern. It would be for these patients as if the proposed system had been implemented. For others, it would be as if the recommendations had been made but not heeded.

The general hypothesis may be extended to an examination of each of the treatment groups independently. Since members of the cross-validation sample had improvement index scores, it is only necessary to compare the means on the index of the "recommended" group and the "non-recommended" group.

Chapter 3: Results

Separation of the Groups

The greater statistical separation of the treatment groups that can be achieved, the greater percentage of correct predictions will be made. This does not mean that simply if significant differentiations are achieved, all predictions will be correct. A certain amount of overlap in the multidimensional distribution is very difficult to avoid. In the course of selecting variables from the original pool the discovery was made that variables not quite significantly differentiating the groups still will have some beneficial influence on the number of correct predictions. For this reason variables were allowed to participate in the discriminant analysis if they had an F ratio of 1.00 or greater, where the value necessary for significance was in this instance 2.03.

Consequently the 36 variables used in the final runs had this minimal value. Table 3 shows their individual values. High F ratios are not always a good indication of the variable's value. In the case of the social class index and its components, complete interdetermination of the variance resulted in spuriously high values. An idea of the true value of these variables was obtained in other runs in which one of them was omitted. The true value of all three of the variables is probably in the area of 2.5, slightly above the 1% significance level. The spuriously high F ratios do not affect the actual predictions. It is clear from the table that no one set of variables demonstrated clear superiority in utility over the others. There are good discriminators among demographic, symptom and diagnostic information. Of the 36 variables, 29 were differentiating the groups at the 1% level or better. Some of the items which performed particularly well were sex, employment status, duration of the precipitating episode, presence or absence of epilepsy, among the back-

Table 3

Contributions of Individual Variables to the Discrimination of Treatment Groups in the Main Sample

Variable	F-ratio[a]		F-ratio[a]
Sex	4.78	Thought Content	5.47
Year of Birth	2.22	Motor Activity	3.80
Marital Status	1.43	Behavior	2.04
Place of Birth	1.26	Orientation	2.15
Established Religion	1.55	Memory	2.20
Occupational Level	28.61	Brain Damage	5.60
Employment Status	4.00	Involutional Psychosis	9.71
Educational Level	28.61	Affective Reaction	4.64
Involuntary Admission	2.36	Chronic Schizophrenia	4.13
Source of Admission	2.43	Acute Schizophrenia	7.29
Previous Care	2.00	Paranoid Schizophrenia	9.41
Duration of Episode	4.62	Other Schizophrenias	3.46
Precipitating Factor	2.33	Depressive Neurosis	15.70
Living Arrangement	1.16	Trait Disturbance Personality Disorder	2.32
Epilepsy	6.75	Sociopathic Personality Disorder	3.11
Social Class	28.61	Transient Situational Personality Disorder	1.70
Affect	2.92	Other Diagnoses	3.34
Stream of Thought	1.00	Addiction to Alcohol or Drugs	3.20

[a]Where df=13,2771 and 2.03 is the 1% probability level.

ground variables, thought content among the mental status variables and brain damage, involutional psychosis, affective reaction, chronic, acute and paranoid schizophrenias and depressive neurosis (which was the best single item) among the diagnoses.

Taken together, these variables distinguished the groups quite significantly. The overall F ratio for the 14 group separation was 6.45 (df = 468,33988), where the value for the 1% level of significance is 1.16.

Values were also computed for the degree of separation of each group from every other group (improved and not improved). Table 4 presents these inter-group comparisons. Of the 91 total comparisons of pairs of treatment groups, 84 are significant at the .001 level and three more at the .01 level on the F test. Four comparisons were not significant, two of which were the improved vs. not improved sections of single treatment groups.

In general the groups were apparently very well separated. It would seem Nathan's fear of one syndrome per patient when groupings are done on the basis of observed and background behavior is ungrounded.

Obtaining Successful Profiles

The actual formulations of successful treatment groups which constitute the basis for making the recommendations were extracted from the separation of groups analysis. The weights for the successful subgroups, together with these standard groups' means and standard deviations on the 36 items were collected. These defining profiles are presented for each of the seven groups in Table 2 of the Appendix.

Imposing the Profiles

The profiles were then imposed on the cross-validation sample through the modified multiple discriminant program. In effect, the members of the

Table 4

Separation of Groups in the Main Sample

Group		A Imp	A Nimp	C Imp	C Nimp	D Imp	D Nimp	E Imp	E Nimp	F Imp	F Nimp	K Imp	K Nimp	X Imp	X Nimp
A	Imp	-													
	Nimp	4.23[a]	-												
C	Imp	5.64	6.21	-											
	Nimp	4.13	3.11	3.06	-										
D	Imp	9.33	7.74	11.49	6.04	-									
	Nimp	4.41	3.24	6.65	3.58	1.35	-								
E	Imp	12.01	8.47	11.89	6.03	16.07	8.42	-							
	Nimp	4.80	3.21	5.72	3.93	6.91	4.67	2.50							
F	Imp	6.54	7.34	7.14	5.94	10.27	6.88	17.34	8.31	-					
	Nimp	3.62	3.31	5.50	4.03	5.57	3.84	8.94	5.17	1.97	-				
K	Imp	7.99	5.44	9.23	5.01	2.75	1.67	8.43	4.56	10.96	6.09	-			
	Nimp	3.82	2.70	5.51	3.12	2.46	1.47	5.37	3.32	5.97	3.91	1.00	-		
X	Imp	7.05	4.50	8.10	3.69	3.72	1.72	9.87	4.17	9.63	4.90	2.54	1.79	-	
	Nimp	4.16	2.48	4.60	1.38	4.52	2.15	8.14	3.49	7.00	3.51	3.99	2.59	2.22	-

[a]F-ratio value where df=36,2771 and 1% level of significance is 1.63.

cross-validation sample were being individually assessed and treatments were recommended for them just as if the model were in operation at the hospital.

The program multiplied the score of an individual on a given variable by the co-efficient of the corresponding variable on the discriminant function of one of the groups. For the individual the sum of these products over all 36 variables was computed for each potential treatment pattern. The individual was assigned by the program to the group for which this criterion score was the highest; that is, to the group of improved patients he most nearly resembled.

From the criterion score two further pieces of information were computed. One was Mahalanobis' D^2, a measure of the geometric distance a subject is from each of the groups. D^2 is the monotonic inverse of the criterion score, so that assignment based on it would be to the group with the smallest score. The posterior probability of the subject being assigned to each of the groups was also computed. This probability is based on the criterion score and is standardized so that the sum of the probabilities of being in the 7 groups is 1.0. So for every member of the cross-validation sample the criterion score was computed and its derivatives, D^2 and the posterior probability, were printed out for each of the seven groups, together with the name of the group to which he was assigned.

Parenthetically, the modification of the program to impose profiles did not affect the normal discrimination routine. The statistical separation of the cross-validation sample groups was carried out with the indigenous figures. The imposed profiles only affected the assignment to groups. Consequently it was possible to assess the degree to which the 36 items distinguished the cross-validation groups. The seven categories

(improved and unimproved sections were not separated in the cross-validation sample because the assignment to treatment had to be done "blind") were differentiated even better than the groups of the main sample. Table 5 summarizes these distinctions. The overall separation of groups was also highly significant ($F = 5.21$, $df = 210,8162$). Finally, 28 of the 36 variables helped in this separation with F ratios of 1.0 or better, and most of the variables previously mentioned as especially important in discriminating the main sample were also important in this analysis.

The cross-validation sample had, of course, already received some pattern of treatment. In order to test the hypothesis, those subjects for whom the actual pattern of treatment was the same as the prescribed pattern of treatment were separated from those in whom it differed. The prediction was that those receiving the recommended treatment would have more successful outcomes than those who did not.

The prediction was not borne out. In the cross-validation sample 438 people received the recommended treatment and 979 did not. The mean scores of these groups on the improvement index were 9.17 and 9.57, the values being in the opposite direction from the prediction. The F ratio for this difference was 7.90 ($df = 1,1415$), significant at the 1% level.

An examination of the individual groups' means revealed where the principal areas of this reversal of the prediction were. Group C and the catch-all Group X are the prime sources of differences in the sample. None of the other five differences are significant, as may be seen in Table 6.

Table 5

Separation of Groups in the Cross-Validation Sample

Group	A	C	D	E	F	K	X
A	-						
C	4.04[a]	-					
D	6.76	8.30	-				
E	7.67	8.25	10.80	-			
F	4.89	5.04	7.46	11.85	-		
K	6.06	7.10	1.91	6.04	8.18	-	
X	5.12	6.02	2.70	6.80	7.02	1.97	-

[a]F-ratio value where df=35,1376 and 1% level of significance is 1.63.

Table 6

Improvement Index Means for Subjects Receiving and Not Receiving Recommended Treatment

Treatment Group	Recommended	Not Recommended	F	p
A	9.57	9.60	.02	-
C	8.80	9.87	6.28	.025
D	9.45	9.38	.04	-
E	8.94	9.47	1.74	-
F	9.11	8.57	1.57	-
K	9.27	9.35	.02	-
X	8.57	9.87	5.20	.05
Total	9.17	9.57	7.90	.01

Chapter 4: Discussion

The results indicate a system developed on two-thirds of a sample does not generalize to the other one-third of the sample. In fact, the prediction was significantly negated in two of the seven groups. This comes in spite of the very significant, in statistical terms, separation of the main and cross-validation samples. In addition, a post-mortem comparison of the main and cross-validation samples using the 36 variables was undertaken, with negative results; they could not be differentiated, which suggests they were, in fact, fairly randomly divided. There is the further knowledge that the same variables acted in about the same way in differentiatiating each of the samples.

It therefore appears the group separations were very nicely cross-validated, but the predictions in the case of the average individual patient were not.

Numerous possibilities exist. One is that despite the large statistical separation of the centroids of the groups, much overlapping of distributions occurred. This is suggested by a scan of some typical D^2 scores. Here is one subject's D^2 values for the seven groups: A,25.35; C,26.78; D,21.25; E,26.16; F,24.01; K,24.20; X,24.01. There is an indication in these values that the group centroids are nearer to one another than to many of the data points. Another subject's values are all between 6 and 7. It resembles an example from astronomy in which persons on opposite sides of the earth seem to be far apart, but if the reference system includes their distance from a group of stars, they are both essentially in the same position and both of them would report being very nearly the same distance from each of the group of stars. In this way, the usual pattern of D^2 values for a given subject was that they were all in the same range, e.g. 6-8 or 25-27, etc.

A drop in predictive ability from the standardizing to the cross-validating samples is traditional, though not always so severe as in this case. One example is the study of Neithercutt (1969). His attempt to predict successful federal prison parolees saw a decrease in accuracy from 80-90% to 50-70%. Drops, of course, can be even more severe when a difference of one sort or another exists between the standardizing and cross-validating samples. Smith et al. (1968) report a study which discovered that a weighted scoring key developed on a sample of chemists failed on a sample of physicists in the same institution. Another study (Wernimont, 1962) found that a weighted key for selecting long term office employees from application forms did not cross-validate after a five year period. The author suggested that certain of the items had changed in social meaning.

To counteract these prediction drops several schemes have been suggested. Goldstein and Linden (1967) discuss an "indecision model." They maintain that of the overlap in the distributions, where error is likely to be very high, a portion should be set aside for "no decision" cases. If the subject falls, for example, into the central two-thirds of the distribution overlap, a prediction for him is not made, but he is set aside for further tests. Of course, when the overlap becomes large, many non-predictions are made. The authors claim, however, that even if this is true, fewer errors in prediction will be made with their model.

A second approach to bettering mediocre prediction is that of Tesser, Starry and Chaney (1967). Their approach is to stack separating devices. They propose beginning with a factor analysis, obtaining factor scores for individuals, doing a hierarchical profile grouping and finally performing an item analysis for the final groupings. Both these papers were, however,

illustrated with "nice" data in small samples. One wonders at the practicality of their solutions (and even if they really would produce significantly better results considering the cost of the process) with relatively large samples.

Standardizing profiles may be very delicate. Especially when a substantial overlap exists in the distributions, they may be so sensitive to the characteristics of the standard sample they are not transferable even to very similar samples.

Information comprising the profiles is also important. A sometimes forgotten requirement is that variables must be continuous. Attempts to render certain variables continuous may not be successful. Further, even when the variable is continuous, the particular dimension exploited may not be the only one possible. While it may be a sensible dimension to exploit, it may not be the one which best accounts for the variance of the criterion. Perhaps a careful study of the "mode of continuity" of selected successful variables would yield some general rules as to methods for selection of relevant dimensions in variables.

In the present investigation simplifications of the information may have lessened the chances for better prediction. For example, it will be remembered that order and amount of the treatments was ignored. And, of course, it is possible some kinds of vital information were just not included in the study. Although psychological inventories have not seemed a good source of predictive items (according to the Fulkerson and Barry review mentioned earlier), perhaps intelligence tests or certain items from among the vast repertoire of tests would be useful. Among other classes of information (cf. Stein, 1963), only one seems potentially useful in this area; information relating to person-environment interactions

could be helpful, especially if by that one meant person-staff relations as well as person-architecture relations. Patients and staff certainly have social interactions like other people, and the nature of these might well have some bearing on a patient's improvement (and his rated improvement).

It is interesting to note that some of the demographic items were as important or more important than symptoms or diagnoses, the traditional correlates of treatment and improvement. Although the social class index did not perform particularly well, sex, employment status and duration of the precipitating episode did. There seems to be support here for Seaman's contention that various external demands are very important in the hospital career of the mental patient.

Much support and many exhortations have issued from health organizations of all types in recent times concerning the necessity of collection of epidemiological data on patients. While the general call has been heeded in many hospitals, the possibility remains of a slip between the idea and the practice of data collection. Hospital records are not always taken with the extreme care necessary for perfect accuracy; errors of omission may creep in as well as simple clerical errors. The author does not mean to suggest blame in the present study, and he accepts full responsibility for the quality of the data. Nevertheless he was not present at the taking of the hospital records of so many patients, nor at the conversion of them onto data processing cards.

One further intriguing possibility exists. It could theoretically have been possible that in some cases of "wrong" prediction, the recommendation made actually would have improved the patient even more than did the treatment he received which mildly improved him. There is no way in this quasi-

experiment to find out if this was ever the case. It seems reasonable, though, that more than one pattern of treatment could result in the patient's improvement, although the basis of the whole study rests on the proposition that an optimal combination of treatments exists for each patient.

Despite these possible sources of error and the failure of the cross-validation, a system similar to the present one should work. This is not stated out of hope alone. To the degree the present procedures of treatment administration work, there are rules implicit in its operation. This study is an attempt to discover and make explicit such rules, in the form of the profile formulations. There is a possibility that no uniform or stable rules are being used in the hospital, explicitly or implicitly. In the belief that there might be and that there should be, an approach such as the present one could be successful with further similar attempts. Such an attempt will be reported in the next chapter.

Nevertheless, the problems and confusions of the present diagnostic system may hinder development of a better system. It is the old problem of trying to create a good model from poor or questionable or unstable data. As a recent participant in a conference on psychiatric epidemiology put it, "the core of the dissatisfaction with the present classification is that the typology implies disease entities for which there is as yet no rigorous scientific evidence" (Bahn, 1967). The current labels may never be reliable when they are built on entities which, since they are not defined, are subject to varying opinions on their nature and relation to individual cases. Since their use as "shorthand formulations" may affect the whole cast of a patient's relations with the staff, his treatment and the perceptions of the staff and himself, the continued use of them may obscure attempts to

find generalizable hypotheses.

As long as etiology remains a mystery in psychiatry, the search for information relevant to the nature of the disorder must continue. Sources of information which have not been previously examined must be considered. Of particular interest may be complex combinations of events in a person's life which are associated with mental disturbance.

The need to search for the peculiar combinations of events which culminate in mental disorder is recognised in some areas. Leighton (1967), for example, proposes focussing not on symptoms (which he considers a "disease-contaminated" word), but on "behaviors of psychiatric interest." These are "recurrent psychological, biochemical, social, cultural and situational events." This is a broad conception, but it is not clear whether Leighton would advocate that any particular single behavior or event is necessarily of psychiatric interest. Clearly, such will not be the case for all people. Rather, it is patterns and combinations of Leighton's behaviors of psychiatric interest, perhaps not recurrent, perhaps somewhat idiosyncratic, that are most important.

A wide-ranging search for interactions of events and behaviors related to successful outcome is the subject of the next chapter. This search, though extensive, is limited to first-order combinations of variables already mentioned.

Chapter 5: A General Linear Model Approach to Prediction of the Criterion

The earlier portion of this investigation was concerned primarily with the attempt to construct an optimal treatment model based on the discrimination of groups of patients receiving certain patterns of treatment. The results indicated that while the groups were "significantly" separable, in practical terms (i.e. apt predictions for individuals) they apparently were not separable enough.

These findings prompted the initiation of a slightly different approach to the problem. One of the causes of insufficient group separation was probably the relative inability of the zero-order variables to account for enough of the variance in the improvement criterion. Consequently the approach evolved toward an investigation of strategies which might yield information more closely related to the criterion. If such strategies produced variables with a higher multiple correlation with the criterion a return to attempting to separate the treatment groups might be feasible. In general, the strategies involved were variations of the general linear model of prediction.

Method

Earlier results suggested the original selection of variables alone have limited predictive power, but additionally, discussion in the last chapter brought attention to the possibility of patterns or combinations of behaviors and events being related to hospitalization and outcome. The search for predictive information led naturally to the joint interactions of the original variables as a beginning.

Among the variables used, then, were the 36 variables in the final selection from the earlier part of the investigation, together with two new sets of variables. The five treatment variables were initally included

(they could not be used earlier since the groups were divided according to received treatments, while in this chapter the undivided sample will be examined).

The second new set of variables comes from the first-order interactions of the 41 already mentioned variables. The total number of these interactions, if both forms of them (a x b and b x a) are considered, and they are as will be seen, amount to 1681. These 1681 possible interactions may be visualized as the elements of a 41 X 41 matrix.

One set of rows of this matrix was eliminated. Interactions between diagnoses and other variables (of the form diagnosis x other variable) would not be meaningful due to the mutually exclusive nature of the coding of the diagnostic variables; if a person received diagnosis A (coded 1), the other diagnoses were not free to vary, but had to be coded 0. In this manner the matrix shrank to 28 x 41, or 1148 possible interactions.

When the step of evaluating the predictive power of the variables and their interactions was reached, several stumbling blocks appeared. The first, obviously, was the sheer size of the number of potentially valuable interactions. Furthermore, as Cohen (1963) points out, the best way to evaluate the interactions is to use them all at once in an equation. The procedure is to establish a baseline of predictive power, represented by the multiple correlation of the criterion with the variables before they are combined, to then add the interactions to the set of predictors (in this case now 1189 in number) and finally to note the resulting increment in the multiple correlation. Among other reasons, this is the ideal procedure because it recognises that interactions may not meaningfully be regressed against the criterion without the presence of the zero-order components.

It is not presently practicable, however, to perform a multiple regression over 2820 subjects with 1189 predictors.

Rather than abandon this approach altogether, or submit to the tempting impulse to guess which of the interactions might be useful, a systematic method of evaluating all the interactions was developed, as follows. An executable though still large regression equation might involve around 100 of the variables or so. With this and the form of the 28 x 41 matrix in mind, together with the necessity for the components of the interactions to be present in the equation, a practical and methodic compromise was devised.

The interactions could be examined in groups corresponding to rows of the matrix. With this basic idea, the interactions were all examined, in a series of 28 regression analyses. The reason for leaving both forms of the individual interactions in consideration now becomes understandable; each one had chances in two different regression equations to prove its worth. So, while the method devised was not quite ideal, it allowed a fairly good look at a very broad range of interactions in a systematic manner.

Each of the 28 regression equations would, then, have consisted of 41 zero-order variables plus 41 joint interactions, or 82 predictors of the improvement criterion, if it had not been for the second stumbling block. In the first runs extremely small values were noticed for the determinants of the correlation matrices on which the regressions are based. These values meant that the matrices were essentially singular; trustworthy evaluations of the interactions could not be expected.

The matrices were singular because several rows or columns in them had very nearly parallel sets of values, a condition arising in this case

from the presence in the matrix of two or more variables which were very highly, but spuriously correlated. For example, the variable "age" might be in this relation to the interaction "age x educational level." In order to free the matrices from singularity it was necessary to examine each of them for such spurious and repetitive interactions and to eliminate them.

The actual practice of this precaution resulted in the substantial reduction of the number of interactions. Approximately 50% of the interactions were actually used in the 28 regression analyses:

A baseline against which to measure the power of the interactions was determined. The multiple correlation between the 41 zero-order variables and the criterion was .367. When analyses involving the interactions were done, any increase in this value would be explicable in terms of the group effort of the interactions included. Among other items of information about the regression, STEPPR, the program used, provides a measure of the significance (an F test) of each of the predictors in the equation.

The best of the interaction variables in each of the analyses were noted. The criterion for this, receiving an F value of 5.0 or greater, roughly corresponded to a decrement in the multiple correlation (i.e. the decrease expected if the variables had not been included among the predictors) of .0016.

An inspection of Table 7 reveals that while a fair number of interactions (23) met this modest criterion, they and the other, less useful interactions in the equations did not collectively increase the baseline multiple correlation very much. Nevertheless, evaluation of the interactions was continued to the final stage.

In this stage the interactions which had shown promise were combined

Table 7

Selection of Useful Interactions

Analysis No.	Interactions[a]	Multiple Correlation
0	(Baseline - none included)	.367
1	sex plus source of admission, thought content, brain damage, drugs for physical disorders, group therapy	.389
2	age plus marital status	.374
3	marital status plus age	.381
4	none	.376
5	none	.377
6	catholic/protestant plus stream of thought	.381
7	none	.371
8	employed status plus drugs for physical disorders	.382
9	none	.367
10	involuntary admission plus depressive neurosis	.375
11	none	.381
12	none	.377
13	none	.369
14	none	.371
15	none	.377
16	none	.372
17	none	.382
18	none	.368
19	thought content plus trait personality disorder	.379
20	motor plus catholic-protestant and thought content	.382
21	none	.371
22	orientation plus other diagnoses	.377
23	none	.382
24	none	.369
25	anti-depressants plus stream of thought and thought content	.386
26	drugs for physical disorders plus sex and employment status	.382
27	ECT plus chronic undifferentiated schizophrenia, transient situational personality disorder, anti-depressants	.384
28	group therapy plus sex and sociopathy	.384

[a]Listed are those receiving an F-value greater than 5.0.

into one regression analysis for a final evaluation. As before, the correlation matrix was examined for spurious values. Twenty-three interactions began this final stage, but four were duplications and eleven had to be rejected on the false-high inter-correlation grounds.

The eight interactions which survived these various tests were the following paired combinations: sex and thought content, sex and brain damage, sex and drugs for physical disorders, sex and group therapy, age and marital status, motor activity and stream of thought, Catholic-Protestant and stream of thought, Catholic-Protestant and motor activity. These eight remaining interactions raised the baseline established by the 41 zero-order variables from .367 to .386 in the subsequent regression analysis.

One of the objectives of this chapter was to find the centers of predictive utility among the zero-order variables as well as among their interactions. As yet no elimination of superfluous variables in the zero-order set has been made. Perhaps the elimination of these would yield a compact set of predictors accounting for nearly all the variance accounted for by the larger set.

The 28 regression equations were scanned for guidelines to choose zero-order variables by.. The established was merely that the variable reach the F value of 5.0 in any of the equations. In the selection process, however, it was noticed that a quite strong but not absolutely rigid regularity in the relative strengths of the variables existed; those eventually selected were over 5.0 in most of the runs and those not selected rarely budged from their lowly significance levels.

A final series of regression runs then were completed to isolate the smallest set of predictors accounting for the largest proportion of the variance. The result was that 19 of the zero-order and two of the first-

order variables (both those involving the Protestant-Catholic dimension) were found to be expendable in this new context with a multiple correlation drop of only .01 from the baseline, to .376. Thus 28 predictors remained (see Appendix Table 4).

Certain of these remaining predictors are clearly the bulwark of the relationship to improvement in the hospital. Four of them stand out: sex, employment status at admission, duration of the precipitating episode and receiving group therapy. If their individual correlations with the criterion faithfully reflect the direction of improvement, one should be a male employed at admission with a short episode who is given group therapy.

These variables stand in direct individual relation to the criterion except for sex. The zero-order effect of sex in the context of the last equation is negligible. Its simple correlation with the criterion is also the lowest of the four. In the twenty-eight runs, however, sex had always been a very significant individual predictor. The explanation lies in this variable's tendency to hand its ability over to interaction terms. It will be recalled that four of the six finally useful interactions contain the variable as a member.

Other variables contributed slightly less substantially than these four. These include age, diagnoses of brain damage and sociopathy and receiving drugs for physical disorders. Although the simple correlations with the criterion were quite mild here, the improvement direction is toward youth, brain damage and drugs for physical disorders and away from sociopathy.

A third order of contributing variables had modest but definite and regular effects. These include involuntary admission, source of admission, stream of thought, memory, diagnoses of depressive neurosis, trait person-

ality disturbance and addiction, and treatments of anti-depressants and ECT.

The final set of 28 predictors was employed to develop a new system of treatment formulations. The paradigm remained the same; the formulations were developed on two-thirds of the sample and a cross-validation was attempted on the other third. The difference was that the tool in this case was multiple regression analysis instead of discriminant analysis.

This procedure took two roughly equivalent forms, representing exploration in the face of uncertainty concerning their relative value. The first method involved replacing the specific treatment variables with a dummy variable for treatment group membership. The improvement criterion was predicted seven times, each time over the undivided sample using the regular set of predictors plus one of the seven columns of the dummy variable. This yielded treatment formulations associated with each of the seven treatment groups.

In the second variant, the sample was divided as previously into the seven treatment groups. The dummy variable was not used. The criterion was predicted for each group, and the raw score weights of the regression equations, as in the first variant, were used to construct seven treatment formulations.

Results

The multiple regression technique involving a dummy variable did not prove practicable. The seven raw score regression weight formulations developed were extremely similar; this resulted in nearly random assignment of the patients to treatment.

However, the divided-group method produced results in the direction of the hypothesis. When the cross validation sample had been recommended to treatment according to divided-group recommendations (see Appendix Table 5),

133 of the recommendations coincided with the actual treatment received. The regression weights comprising these formulations may be seen in Appendix Table 5. The treatment variables were not used in this variation because they were redundant, since the variation depended on division of the groups by treatment group. An exception to this is group therapy, a treatment which was not employed in the division. Curiously, in the present case no one from the biggest group (A) was recommended to his own group.

The mean improvement of the people recommended to their own group was 9.42, while that of those who were not was 9.33. The F-ratio for the difference of means was not significant, however.

Discussion

An overall look at the performance of classes of variables makes it clear that background and treatment variables are more important in predicting improvement than the mental status examination or diagnoses as measured here. It is doubtful, however, that these latter classes may be completely dismissed without harming the efficacy of the prediction equations.

As was noted in earlier discussion, Seaman's "external demands," or at least items which refer to extra-hospital behavior and events seem to be very important. It might almost seem the hospital does not affect the patient one way or the other as much as the individual himself and any people close to him.

The multiple correlations obtained in this investigation have room for improvement, certainly, but they are respectable considering the size and diversity of a sample which includes the full range of diagnoses from addicts and organically impaired people to the usual functional diagnoses. Nevertheless, it is possible the criterion could have benefited from another,

equally reasonable choice of components or weightings. For example, the post-hospital occupational status of the patient was determined solely from what the patient said his plans were at discharge. Similarly, some of the variables might have served the criterion better with slight or subtle shifts in the exploited implicit dimensions.

One set of interactions could have been expected to perform better. The failure of the treatment-diagnosis interactions to have much effect in predicting improvement would seem to be indicative of a poor or vague inter-relationship just where a strong and clear one is needed. The blame could as well be laid the infancy of psychiatric therapeutics as the shortcomings of the diagnoses. In this study some attention has been given to the phenomenon of psychiatric diagnoses, not so much to challenge them as to explore their meaning in relation to therapy and treatment allocation and their potential as pure predictive information.

This exploration led to one particular series of regression analyses worthy of mention. The observation was made that the criterion had a higher multiple correlation with essentially the same body of predictors in certain instances. Specifically, the baseline of .367 was exceeded by a surprising margin (ranging up to .54) in each of the treatment groups when they were individually examined. In fact when the treatment variables were removed from the equation for the collective sample so that the body of predictors was identical, the R dropped to just below .32.

At first these findings suggested the treatment variables, as the basis for the division into groups, were acting as moderator variables (Saunders, 1956). This seemed especially evident when it was determined that the variables were less predictive when they were in zero-order or joint interaction form than when they were acting indirectly, as the basis for the

divisions.

However, the sample sizes of the groupings were naturally much smaller than that of the undivided sample. An alternative possibility to the moderator situation is that the increases in R simply reflected differences in the two analyses in subject-to-variable ratio. This hypothesis seemed more likely when it was noticed the correlations were fairly directly related to sample size even among the treatment groups themselves.

The results of the second variant of the multiple regression technique were encouraging, as they mildly supported the hypothesis despite the obstacles already mentioned. A group mean multiple correlation of .42 for prediction of improvement was obtained (the range was .39 to .51). Though few in number, those persons who "received" the recommended treatment, based on a set of predictors ranging from the strictly demographic to mental status at admission, conventional diagnosis and combinations of these classes, did improve more than those patients who received another treatment.

Specific areas which are now receiving attention, with the object of lessening their weight as obstacles to a more clearly-focussed prediction system, are the mental status observations, the recording of treatments and the criterion of improvement. A trial period of six months has been authorized for a new treatment recording form which will include very detailed treatment information (specific drug, dosage, frequency) for the patient on a sample-day basis (i.e. approximately every 10 days of his stay). Personal interviews with released patients are planned to clarify the conception of improvement.

References

Aldrich, C. K. An introduction to dynamic psychiatry. New York: McGraw-Hill, 1966.

Aldrich, C. K. Social service information. In Freedman, A. M. & Kaplan, H. I., (eds.) Comprehensive testbook of psychiatry. Baltimore: Williams and Wilkins, 1967, 541-542.

American Psychiatric Association. A psychiatric glossary (3rd ed.). Washington: American Psychiatric Association, 1969.

Anastasi, A., Meade, M. J. & Schneiders, A. A. The validation of a biographical inventory as a predictor of college success. New York: College Entrance Examination Board, 1960.

Bahn, Anita K. Some methodologic issues in psychiatric epidemiology. In Monroe, R. R., Klee, G. D. & Brody, E. B. (eds.), Psychiatric epidemiology and mental health planning. Research report #22 of the American Psychiatric Association. Washington: American Psychiatric Association, 1967, 69-83.

Barnette, W. L. (ed.). Readings in psychological tests and measurements. Homewood, Illinois: Dorsey Press, 1968.

Bell, L. V. 26th Annual report of the McLean Asylum, 1844. James Soring Press. Quoted from Plunkett R. J., & Gordon, J. E. Epidemiology and Mental illness. New York: Basic Books, 1960, 18-19.

Bishton, R. A study of some factors related to achievement of intelligence superiority in eighth grade children. Journal of educational research, 51, 1958, 203-208.

Brill, H. Nosology. In Freedman, A. M. & Kaplan, H. I. (eds.). Comprehensive textbook of psychiatry. Baltimore: Williams and Wilkins, 1967, 585.

Bucklew, J. Prediction of psychiatric symptoms from case history information. Psychological record, 1967, 12, 1-12.

Bucklew, J. The use of symptoms to assess case history information. Multivariate behavioral research, 1967, Special issue, 157-168.

Campbell, O. T. & Stanley, J. C. Experimental and quasi-experimental designs for research on teaching. In Gage, N. L. (ed.) Handbook of research on teaching. Chicago: Rand-McNalley, 1963.

Clyde, D. J. Discussion. In Hoch, P. H. & Zubins, J. (eds). The evaluation of psychiatric treatment. New York: Grune and Stratton, 1969, 309-311.

Cole, J. O. Evaluation of drug treatment in psychiatry. In Hoch, P. H. & Zubin, J. (eds). The evaluation of psychiatric treatment. New York: Grune and Stratton, 1969, 14-31.

Cooley, W. W. & Lohnes, P. R. Multivariate procedures for the behavioral sciences. New York: Wiley, 1962.

Creelman, A. G. The prediction of physics grades at the university level from previously recorded data. Master of Arts thesis, University of British Columbia, 1964.

Cronbach, L. J. & Gleser, G. C. Psychological tests and personnel decisions. (2nd ed.). Urbana: University of Illinois Press, 1965.

d'Heurle, A., Mellinger, J. C. & Haggard, E. A. Personality, intelligence and achievement patterns in gifted children. Psychological monographs, 73, 1959.

Dixon, W. J. BMD biomedical computer programs. Los Angeles: University of California Press, 1967.

Drever, J. A dictionary of psychology. Baltimore: Penguin, 1964.

Dunham, H. W. & Weinberg, S. K. The culture of the state mental hospital. Detroit: Wayne State University Press, 1960.

Engelsman, F. et al. A contribution to the diagnosis of functional psychoses by means of a computer. Ćeskoslovenska Psychiatric, 1968, 64 (6), 368-377. (Psychological abstracts, 43, 17619).

English, H. B. & English, A. C. A comprehensive dictionary of psychological and psychoanalytical terms. New York: McKay, 1958.

Fisher, R. A. The use of multiple measurements in taxonomic problems. Annals of eugenics, 1936, 7, 179-188.

Fulkerson, S. E. & Barry, J. K. Methodology and research on prognostic use of psychological tests. Psychological bulletin, 1961, 58, 177-204.

Goldsmith, D. B. The use of personal history blanks as a salesmanship test. Journal of applied psychology, 1922, 6, 149-155.

Goldstein, S. G. & Linden, J. D. Paradigm to maximize efficiency in clinical prediction. Proceedings of the American Psychological Association, 1967, 2, 197-198.

Gove, P. B. (ed.). Webster's third new international dictionary, Springfield, Mass: Merriam, 1969.

Holtzman, W. H. Can the computer supplant the clinician? Journal of clinical psychology, 1960, 16, 119-122.

Hope, K. Methods of multivariate analysis. London: University of London Press, 1968.

Kaelbling, R. & Patterson, R. M. Eclectic psychiatry. Springfield, Ill: Charles Thomas, 1966.

Kanfer, F. H. & Saslow, G. Behaviorial analysis. Archives of general psychiatry, 1965, 12, 529-538.

Kendall, M. G. A course in multivariate analysis. London: Hafner, 1957.

Kirchner, W. K. & Dunette, M. D. Applying the weighted application blank technique to a variety of office jobs. Journal of applied psychology, 1960, 44, 325-326.

Kleinmuntz, B. & McLean, R. S. Diagnostic interviewing by digital computer. Behavioural science, 1968, 13(1), 75-80.

Klingman, J. & Pipbeger, H. V. Computer classifications of electrocardiograms. Computers and biomedical research, 1967, 1, 1-17.

Kolb, L. C. Modern clinical psychiatry (7th ed.). Philadelphia: Saunders, 1968.

Kraepelin, E. Dementia praecox and paraphrenia. Edinburgh: Livingstone, 1919.

Lavin, D. E. The prediction of academic performance. New York: Russell Sage Foundation, 1965.

Lee, S. D. Social class bias in the diagnosis of mental illness. University of Oklahoma Ph.D. Dissertation, 1968. In Dissertation Abstracts, 1968, 28, 11-B, 4758-4759.

Leighton, A. Is social environment a cause of rental disorder. In Monroe, R. R., Klee, G. D. & Brody, E. B. (eds.) Psychiatric epidemiology and mental health planning. Research report #22 of American Psychiatric Association. Washington: American Psychiatric Association, 1967, 337-345.

Levine, M. Principles of psychiatric treatment. In Alexander, F. & Ross, Helen (eds.). The impact of freudian psychiatry. Chicago: University of Chicago Press, 1961, 200-259.

Lindemann, L. et al. The use of demographic characteristics in predicting length of neuropsychiatric hospital stay. Journal of Consulting psychology, 1959, 23, 85-89.

McGrath, J. J. Improving credit evaluation with a weighted application blank. Journal of applied psychology, 1960, 44, 325-326.

McGuire, C. et al. Dimensions of talented behavior. Educational and psychological measurement, 1961, 21, 3-38.

Michael, W. B., Jones, R. A. & Trembly, W. A. The factored dimensions of a measure of motivation for college students. Educational and psychological measurement, 1959, 19, 667-671.

Nason, L. J. Academic achievement of gifted high school students. Los Angeles: University of Southern California Press, 1958.

Nason, P. E. Cues, decisions and diagnoses. New York: Academic Press, 1967.

Neithercutt, M. G. Predicting outcomes of federal parolees. Doctoral dissertation, University of California, Berkeley, 1968. In Dissertation Abstracts, 1969, 29(9-B), 3512.

Plunkett, R. J. & Gordon, J. E. Epidemiology and mental illness. New York: Basic Books, 1960.

Redlich, F. C. & Freedman, D. X. The theory and practice of psychiatry. New York: Basic Books, 1966.

Rubin, S. Psychiatry and criminal law. Dobbs Ferry, N.Y.: Oceana Publications, 1965.

Rulon, P. J. et al. Multivariate statistics for personnel classification. New York: Wiley, 1967.

Russell, W. & Cope, G. V. Method of rating the history of achievements of applicants. Public personnel studies, 1925, 3, 202-219.

Safford, P. L. Prediction of academic performance from an object sorting concept formation recognition task. Proceedings of the American Psychological Association, 1967, 2, 291-292.

Seaman, Carolyn M. The prediction of length of hospitalization from personality test scores, demographic and social variables according to the process-reactive continuum. Doctoral Dissertation, Ohio University, 1968. In Dissertation Abstracts, 1969, 4-B, 1513.

Skinner, H. A. The origin of medical terms. Baltimore: Williams and Wilkins, 1961.

Silverberg, N., Iversen, I. & Silverberg, M. Predicting end of first grade developmental reading test scores from Gates reading readiness test scores administered in kindergarten. Proceedings of American Psychological Association, 1967, 2, 291-292.

Slater, E. & Roth, M. Clinical psychiatry (3rd. ed.). Baltimore: Williams and Wilkins, 1969.

Smith, J. T. et al. Academic, sociometric and personality variables in the prediction of elementary school achievement. Proceedings of American Psychological Association, 1967, 2, 339-340.

Smith, W. J. et al. The prediction of research competence and creativity from personal history. In Barnette, W. L. (ed.). Readings in psychological tests and measurements. Homewood, Illinois: Dorsey Press, 1968, 270-274.

Stein, M. I. Personality measures in admissions. New York: College Entrance Examination Board, 1963.

Stern, G. C., Stein, M. I. & Bloom, B. S. Methods in personality assessment. Glencoe, Illinois: The Free Press, 1956.

Stone, S. The contribution of intelligence, interests, temperament, and certain personality variables to academic achievement in a physical science and mathematics curriculum. Dissertation Abstracts, 1958, 18, 669-670.

Szasz, T. The problem of psychiatric nosology. American journal of psychiatry, 1957, 114, 405-413.

Szasz, T. The myth of mental illness. New York: Harper, 1961.

Szasz, T. Law, liberty and psychiatry. New York: MacMillan, 1963.

Tesser, A., Starry, A. R. & Chaney, F. B. Toward better prediction: a sub-grouping approach. Proceedings of American Psychological Association, 1967, 2, 261-262.

Tucker, L. R. Formal models for a central prediction system. Psychometric Monograph Number 10.

Van Naerson, R. F. Application of the decision theoretical approach to the selection of drivers. In Cronbach, L. J. & Gleser, G. C. Psychological tests and personnel decisions. Urbana: University of Illinois Press, 1965, 273-290.

Ward, J. H. & Davis, K. Teaching a digital computer to assist in making decisions. In Cronbach, L. J. & Gleser, G. C. Psychological tests and personnel decisions. Urbana: University of Illinois Press, 1965, 268-273.

Wender, P. H. Dementia praecox: The development of the concept. American Journal of psychiatry, 1963, 119, 1143-1151.

Wernimont, P. F. Re-evaluation of a weighted application blank for office personnel. Journal of applied psychiatry, 1962, 46, 417-419.

Further References

Cohen, J. Multiple regression as a general data-analytic system. Psychological bulletin, 1968, 70, 6, 426-483.

Darlington, R. B. Multiple regression in psychological research and practice. Psychological bulletin, 1968, 69, 161-182.

Saunders, D. R. Moderator variables in prediction. Educational and psychological measurement, 1956, 16, 209-222.

APPENDICES

Table 1

Original Pool of Potential Profile Variables

Name	Range	Means	Standard Deviation	In Final Selection?
1 Sex	1 - Male 2 - Female	1.54	.50	Yes
2 Year of Birth	(last 2 digits)	31.48	17.52	Yes
3 Marital Status	1 - Single 2 - Formerly Married 3 - Married	2.20	.89	Yes
4 Place of Birth	1 - Canada 2 - U.S.,U.K., Australia, New Zealand 3 - Other	1.12	.33	Yes
5 Religious Sects	0 - Established Protestant, Catholic,Jewish 1 - Other	.32	.47	No
6 Established Religions	0 - Protestant 1 - Catholic	.27	.45	Yes
7 Occupational Level	1 - Unskilled 2 - Semiskilled 3 - Clerical, Sales Skilled 4 - Professional, Managerial	2.35	.95	Yes
8 Employment Status at Admission	0 - Unsettled, Unemployed Retired 1 - Regular Part-Time,Seasonal 2 - Full-Time, Military	.77	.97	Yes
9 Educational Level	0 - None 1 - Grades 1-4 2 - Grades 5-8 3 - Grades 9-11 4 - Grades 12-13, Vocational 5 - Some University 6 - University Degree	2.96	1.01	Yes

10	Involuntary Admission	1 - Informal, Alcoholic 2 - Order in Council Criminal, Court Order	1.50	.50	Yes
11	Source of Admission	1 - Private Physician 2 - Penal, Hospital, Other Institution	1.25	.43	Yes
12	Previous Care	0 - None 1 - Psychiatrist 2 - Clinic, Other Hospital	.33	.88	Yes
13	Duration of the Precipitating Episode	1 - Under 1 week 2 - 1-2 weeks 3 - 3-4 weeks 4 - 5 weeks-4 months 5 - 5 months-1 year 6 - over 1 year	4.35	1.67	Yes
14	Precipitating Factor	1 - Physical: Illness, Pregnancy 2 - Individual: Sexual,Marital, Relative's Death 3 - Social: School, Job, Unemployment, Alcohol	2.04	.58	Yes
15	Living Arrangement at Admission	1 - Separate or with Spouse, Young Children 2 - With siblings, adult children, other relatives 3 - With Parents 4 - Group Living Home or foster home 5 - Penal or hospital	1.69	1.22	Yes
16	Epilepsy	1 - Not present 2 - Present	1.02	.14	Yes
17	Social Class	(see Table 1 in test)	7.66	2.36	Yes
18	Mental Status: Affect	1 - Normal,appropriate 2 - Flattered,anxious, inappropriate, euphoric,depressed	1.89	.32	Yes

9	Mental Status: Stream of Thought	1 - Normal 2 - Slowed, speeded, disconnected, bizarre, clouded sensorium	1.57	.50	Yes
0	Mental Status: Thought Content	1 - Normal 2 - Delusional, hallucinated, obsessed, phobic	1.49	.50	Yes
1	Mental Status: Motor Activity	1 - Normal 2 - Overactive, slow, bizarre, inadequate	1.54	.50	Yes
2	Mental Status: Behavior	1 - Normal 2 - Compulsive, anti-social, aggressive	1.77	.42	Yes
3	Mental Status: Orientation	1 - Normal 2 - Disoriented	1.10	.30	Yes
4	Mental Status: Memory	1 - Intact 2 - Impaired	1.14	.35	Yes
5	Brain Damage		.082	.27	Yes
26	Involutional Psychosis		.033	.18	Yes
27	Affective Reaction	(For each patient,	.073	.26	Yes
28	Chronic Undifferentiated Schizophrenia	1 - his initial diagnosis	.032	.18	Yes
29	Acute Schizophrenia	0 - all other diagnoses)	.054	.23	Yes
30	Paranoid Schizophrenia		.099	.30	Yes
31	Other Schizophrenias		.022	.15	Yes
32	Depressive Neuroses		.245	.43	Yes
33	Other Neuroses		.079	.27	No
34	Personality Disorder: Trait Disturbance		.062	.24	No
35	Personality Disorder: Pattern Disorder		.035	.18	Yes
36	Personality Disorder: Sociopath		.038	.19	Yes

37	Addiction	.011	.10	Yes
38	Personality Disorder: Transient Situational Reaction	.061	.24	Yes
39	Other Diagnoses	.076	.27	Yes

Table 2

Profiles of Successful Groups

Variable	Means[a]	Standard Deviation[a]	Coefficient of Function[a]
Sex	1.435	.496	7.997
	1.385	.488	8.192
	1.648	.479	8.212
	1.436	.497	7.934
	1.446	.498	8.257
	1.714	.454	8.638
	1.665	.473	8.644
Year of Birth	30.730	15.559	.162
	28.587	18.055	.159
	28.652	16.585	.160
	32.946	11.144	.172
	34.109	15.103	.163
	28.479	18.561	.163
	34.013	20.494	.179
Marital Status	2.198	.889	6.169
	2.175	.871	6.282
	2.488	.756	6.260
	2.062	.954	6.280
	1.990	.890	6.026
	2.420	.818	6.131
	2.329	.872	6.275
Place of Birth	1.113	.317	9.629
	1.111	.315	9.590
	1.092	.290	9.096
	1.095	.293	9.598
	1.145	.353	9.846
	1.193	.397	10.071
	1.146	.353	9.671
Established Religion	.303	.460	1.536
	.294	.456	1.346
	.232	.423	1.393
	.362	.482	1.794
	.223	.417	1.093
	.235	.426	1.457
	.212	.409	1.204
Occupational Level	2.306	.944	1.106
	2.310	.965	1.127
	2.556	.989	1.100
	2.239	.975	1.077
	2.332	.898	.888
	2.471	.872	.939
	2.446	.973	1.082

[a]For the groups in vertical order: A,C,D,E,F,K and X

Employment Status	.896	.988	1.714
	.643	.923	1.510
	1.000	.998	1.707
	.840	.985	1.688
	.907	.985	1.710
	.924	1.001	1.679
	.807	.972	1.598
Educational Level	2.871	.499	2.029
	2.921	.487	2.134
	3.084	.446	2.129
	2.905	.456	2.094
	3.109	.496	2.184
	3.101	.458	2.157
	2.968	.500	2.043
Involuntary Admission	1.533	.430	7.186
	1.619	.474	7.083
	1.272	.339	6.642
	1.707	.430	7.222
	1.570	.492	7.365
	1.294	.390	6.332
	1.478	.418	7.099
Source of Admission	1.243	.838	8.130
	1.337	.895	8.534
	1.132	.896	7.835
	1.243	.955	7.933
	1.404	.843	8.994
	1.185	.843	8.229
	1.225	.822	8.233
Previous Care	.294	1.744	-.318
	.349	1.872	-.262
	.340	1.529	-.189
	.374	1.653	-.231
	.301	1.992	-.306
	.277	1.506	-.356
	.320	1.497	-.266
Duration of Episode	4.317	1.744	2.389
	3.984	1.872	2.284
	4.512	1.529	2.480
	3.959	1.653	2.245
	3.824	1.992	2.263
	4.479	1.506	2.459
	4.399	1.497	2.456
Precipitating Factor	2.103	.543	8.865
	2.155	.688	9.047
	2.016	.522	8.700
	2.012	.564	8.618
	2.052	.528	8.603
	1.983	.537	8.692
	1.987	.622	8.689

Living Arrangement	1.695	1.274	2.541
	1.778	1.362	2.625
	1.348	.866	2.501
	1.893	1.329	2.640
	1.938	1.387	2.616
	1.370	.822	2.473
	1.538	1.040	2.545
Epilepsy	1.008	.087	58.390
	1.063	.244	60.741
	1.004	.063	58.530
	1.004	.064	59.035
	1.005	.072	57.881
	1.008	.092	59.134
	1.032	.175	60.070
Social Class	7.484	2.346	.488
	7.540	2.414	.503
	8.196	2.399	.538
	7.383	2.540	.493
	7.772	2.220	.591
	8.042	2.334	.552
	7.861	2.307	.534
Affect	1.856	.351	17.216
	1.802	.400	16.695
	1.956	.206	17.746
	1.967	.179	17.324
	1.679	.468	15.926
	2.000	.000	17.464
	1.927	.260	17.308
Stream of Thought	1.473	.500	-.873
	1.552	.498	-.716
	1.524	.500	-.542
	1.848	.360	-.521
	1.295	.457	-.622
	1.782	.415	-.490
	1.604	.490	-.761
Thought Content	1.481	.500	7.260
	1.512	.501	7.126
	1.276	.448	6.615
	1.881	.325	7.910
	1.155	.363	6.429
	1.580	.496	7.704
	1.475	.500	7.258
Motor Activity	1.424	.495	.637
	1.516	.501	.908
	1.520	.501	.912
	1.786	.411	1.427
	1.238	.427	.368
	1.840	.483	2.040
	1.623	.498	1.373

Behavior	1.753	.431	6.778
	1.683	.466	6.249
	1.804	.398	6.941
	1.922	.269	6.928
	1.554	.498	6.146
	1.924	.266	7.016
	1.772	.420	6.502
Orientation	1.089	.285	6.297
	1.210	.408	6.537
	1.040	.196	6.664
	1.206	.406	7.258
	1.047	.211	6.440
	1.076	.266	6.592
	1.073	.260	5.875
Memory	1.121	.326	3.784
	1.282	.451	4.062
	1.036	.187	3.354
	1.193	.396	3.329
	1.078	.268	3.684
	1.076	.266	3.205
	1.149	.356	4.319
Brain Damage	.092	.289	5.905
	.302	.460	6.595
	.032	.176	8.988
	.025	.156	7.139
	.098	.299	6.692
	.000	.000	5.149
	.038	.191	5.302
Involutional Psychosis	.014	.117	5.905
	.016	.125	6.595
	.060	.238	8.988
	.016	.128	7.139
	.010	.102	6.692
	.118	.324	10.782
	.076	.265	9.440
Affective Reaction	.074	.261	5.597
	.056	.230	5.937
	.056	.230	6.725
	.062	.241	6.354
	.021	.143	5.618
	.160	.368	8.324
	.085	.280	7.255
Chronic Undifferentiated Schizophrenia	.035	.184	4.934
	.036	.186	5.644
	.020	.140	5.735
	.045	.208	6.203
	.000	.000	4.685
	.034	.181	6.207
	.022	.147	5.854

Acute Schizophrenia	.047	.213	5.031
	.052	.222	5.796
	.004	.063	5.531
	.206	.405	8.960
	.005	.072	4.761
	.034	.181	5.971
	.066	.249	6.719
Paranoid Schizophrenia	.095	.293	4.706
	.067	.251	5.236
	.016	.126	5.251
	.321	.468	8.284
	.005	.072	4.589
	.050	.220	5.639
	.092	.289	6.132
Other Schizophrenias	.020	.140	5,858
	.024	.153	6.802
	.004	.063	6.302
	.054	.225	8.445
	.005	.072	5.914
	.000	.000	6.005
	.022	.147	7.309
Depressive Neurosis	.152	.359	9.900
	.111	.315	10.309
	.584	.494	13.005
	.070	.256	10.793
	.197	.399	10.577
	.471	.501	13.044
	.373	.484	12.350
Personality Disorder, Pattern Disturbance	.078	.269	10.224
	.048	.213	10.265
	.052	.222	11.035
	.008	.091	10.328
	.124	.331	11.327
	.000	.000	10.855
	.051	.220	11.231
Personality Disorder, Sociopathy	.038	.192	18.946
	.028	.165	19.197
	.012	.109	19.726
	.004	.064	19.532
	.114	.319	20.937
	.008	.092	20.599
	.006	.079	19.749
Personality Disorder, Transient Situational	.081	.273	15.112
	.032	.176	14.919
	.040	.196	15.566
	.012	.111	15.058
	.207	.406	17.551
	.000	.000	15.480
	.051	.220	15.907

Addiction	.015	.123	17.639
	.036	.186	20.239
	.004	.063	17.881
	.000	.000	17.970
	.010	.102	17.393
	.008	.092	19.250
	.016	.125	19.386
Other Diagnoses	.063	.243	7.255
	.083	.277	8.101
	.036	.187	8.343
	.119	.325	9.021
	.088	.284	8.001
	.084	.279	9.343
	.051	.220	8.459

Table 3

Simple Correlations Among the Original Pool of Variables and the Criterion[1]

	1	2	3	4	5	6	7	8	9	10
1	-									
2	-.005	-								
3	.255	-.230	-							
4	.048	.018	.089	-						
5	-.021	-.008	-.005	-.043	-					
6	-.057	.005	-.047	-.031	-.419	-				
7	.139	-.054	.126	.022	.006	-.028	-			
8	.015	-.041	.142	.024	.020	-.034	.136	-		
9	.010	.014	-.010	.077	.012	-.095	.251	.108	-	
10	-.167	.020	-.158	-.035	-.053	.070	-.091	-.071	.102	-
11	-.220	.017	-.184	.002	-.008	.087	-.076	-.115	-.057	.218
12	-.027	-.011	-.073	-.006	-.026	.035	-.025	-.076	.028	.052
13	.002	-.011	-.004	.005	.022	-.028	-.022	-.075	-.025	-.148
14	-.172	-.053	-.086	-.037	-.023	.031	.014	-.028	.086	.019
15	-.203	.148	-.495	-.048	-.003	.043	-.110	-.093	-.046	.123
16	-.502	.002	-.046	-.009	-.024	.018	-.038	-.053	-.038	.026
17	.155	-.038	.097	.051	.010	-.064	.910	.155	.630	-.117
18	.144	-.013	.057	.007	.027	-.025	.014	.009	.020	-.042
19	.034	-.051	-.019	-.010	-.017	.002	-.018	-.033	-.042	.135
20	-.053	-.089	-.060	-.042	-.041	.052	-.062	-.066	-.033	.182
21	.066	-.037	-.010	-.012	-.042	-.017	.006	-.044	-.026	.084
22	.047	-.019	-.011	-.019	-.008	-.004	.006	-.003	-.024	-.004
23	-.098	-.019	-.028	-.015	-.016	-.033	-.074	-.068	-.113	.214
24	-.099	-.037	-.019	.009	-.051	-.005	-.052	-.091	-.107	.191
25	-.175	-.085	-.032	-.010	-.040	.065	-.060	-.071	-.070	.160
26	.105	-.107	.067	.089	-.010	-.045	.029	.019	-.040	-.035
27	-.020	-.090	.109	.006	-.014	-.022	.072	.018	.015	.048
28	-.015	.009	-.077	-.038	-.008	.007	-.082	-.076	-.046	.062
29	-.015	.071	-.103	-.032	-.045	.049	-.032	.032	.041	.098
30	-.035	-.073	-.051	-.034	-.006	.016	-.015	-.038	-.022	.175

Table 3

Simple Correlations Among the Original Pool of Variables and the Criterion[1]

	1	2	3	4	5	6	7	8	9	10
31	-.022	.046	-.112	-.012	-.015	.017	-.076	-.074	.009	.040
32	.210	-.024	.169	.042	.050	-.059	.106	.067	.060	.268
33	.050	.014	.097	.030	.023	-.016	.009	.064	.028	-.145
34	-.008	.049	.020	.019	.000	.034	.003	-.001	.044	-.038
35	-.042	.009	.000	-.020	-.008	.001	-.009	-.034	-.026	-.006
36	-.117	.091	-.126	-.046	.000	.030	-.044	-.071	-.018	.020
37	-.067	-.025	-.008	-.009	-.014	-.026	.029	-.014	-.006	-.023
38	.034	.097	-.117	-.019	.049	-.035	-.022	.056	.002	-.020
39	-.073	.066	-.054	-.007	-.016	.011	-.027	.006	-.036	.102
40	-.087	-.064	.095	-.005	.008	-.021	.049	.195	.016	.011

	11	12	13	14	15	16	17	18	19	20
11	-									
12	.118	-								
13	-.073	.023	-							
14	.091	.078	-.040	-						
15	.188	.042	.019	.045	-					
16	.009	-.015	.029	-.065	-.002	-				
17	-.086	-.008	-.028	.048	-.108	-.047	-			
18	-.034	.016	-.003	-.037	-.035	-.034	.020	-		
19	.019	.039	-.020	-.047	-.003	.000	-.032	.330	-	
20	-.016	.056	.025	-.011	.039	-.033	-.064	.185	.373	-
21	.000	.033	-.021	-.067	-.000	.014	-.007	.305	.653	.283

Table 3

Simple Correlations Among the Original Pool of Variables and the Criterion[1]

	11	12	13	14	15	16	17	18	19	20
22	-.071	-.017	.037	-.022	.040	-.006	-.005	.238	.332	.207
23	.043	-.002	-.034	-.043	.041	.082	-.108	.071	.251	.243
24	.076	.004	-.052	-.050	.038	.119	-.088	.046	.228	.192
25	.031	-.035	-.056	-.014	.024	.158	-.078	-.092	.050	.077
26	-.064	-.029	.015	.002	-.065	-.027	.006	.060	.055	-.051
27	.005	.045	-.015	.013	-.064	-.022	.064	.037	.131	.119
28	.016	.093	.058	.044	.027	-.026	-.086	.053	.109	.132
29	.023	-.017	-.116	.001	.061	-.034	-.008	.076	.162	.201
30	.001	-.006	.071	.000	.030	-.048	-.021	.063	.132	.325
31	-.003	.039	.051	.070	.052	-.022	-.057	.024	.076	.094
32	-.049	-.020	-.006	-.055	-.151	-.036	.111	.124	-.096	.271
33	-.081	-.009	.044	-.046	-.046	.004	.019	.039	-.083	-.029
34	.042	-.006	-.039	.030	.013	.025	.021	-.073	-.150	-.162
35	.058	-.006	.091	.051	.010	-.001	-.018	-.045	-.073	-.043
36	.080	.020	.023	.052	.112	-.002	-.043	.227	-.163	-.157
37	.018	-.005	.006	.052	-.001	-.015	.021	-.080	-.073	-.070
38	-.048	-.041	-.071	-.008	.084	-.005	-.017	-.067	-.100	-.170
39	.038	.017	-.004	-.046	.067	.034	-.037	-.052	.045	.033
40	-.054	-.046	-.137	.034	-.055	-.035	.046	-.005	-.051	-.053

	21	22	23	24	25	26	27	28	29	30
21	-									
22	.359	-								
23	.222	.114	-							
24	.189	.073	.644	-						

Table 3

Simple Correlations Among the Original Pool of Variables and the Criterion[1]

	21	22	23	24	25	26	27	28	29	30
25	.053	-.047	.353	.362	-					
26	.070	.037	-.030	.010	-.055	-				
27	.093	.054	-.033	-.018	-.084	-.052	-			
28	.063	.055	.084	.064	-.054	-.033	-.051	-		
29	.113	.072	.064	.069	-.071	-.044	-.067	-.043	-	
30	.090	.079	.028	.023	-.099	-.061	-.093	-.060	-.079	-
31	.071	.058	.044	-.013	-.045	-.028	-.042	-.027	-.036	-.050
32	-.017	-.026	-.172	-.181	-.170	-.105	-.160	-.103	-.135	-.189
33	-.051	-.016	-.065	-.063	-.087	-.054	-.082	-.053	-.070	-.010
34	-.136	-.048	-.078	-.089	-.077	-.047	-.074	-.047	-.061	-.085
35	-.083	-.034	-.059	-.057	-.057	-.035	-.054	-.035	-.046	-.064
36	-.161	-.054	-.067	-.070	-.059	-.036	-.055	-.036	-.047	-.065
37	-.052	-.008	-.025	-.015	-.031	-.019	-.030	-.019	-.025	-.035
38	-.101	-.087	-.072	-.083	-.076	-.047	-.071	-.046	-.060	-.084
39	.170	-.016	-.100	.115	-.086	-.053	-.081	-.052	-.068	-.010
40	-.023	-.032	-.016	-.044	.036	-.018	-.002	-.048	.026	-.018

	31	32	33	34	35	36	37	38	39	40
31	-									
32	-.085	-								
33	-.044	-.166	-							
34	-.039	-.146	-.075	-						
35	-.029	-.109	-.056	-.049	-					
36	-.030	-.112	-.058	-.051	-.038	-				

Table 3

Simple Correlations Among the Original Pool of Variables and the Criterion[1]

	31	32	33	34	35	36	37	38	39	40
37	-.016	-.060	-.031	-.027	-.020	-.021	-			
38	-.038	-.145	-.074	-.065	-.049	-.050	-.027	-		
39	-.043	-.164	-.084	-.074	-.055	-.057	-.030	-.073	-	
40	-.045	.082	-.005	-.047	.005	-.089	.040	.016	-.047	-

[1]See Appendix Table 1 for names of variables. Variable 40 in this table is the improvement criterion.

Table 4

Simple Correlations Among a Selection of Original and First-Order Variables Which Maximize Prediction[1,2]

	Anti-Despressives	Drugs for Physical Disorders	E.C.T.	Group Therapy	Improvement Criterion	1 + 20	1 + 25	1 + Drugs for Physical Disorder	1 + Group Therapy	2 + 3	19 + 21
1	.167	.001	.060	.035	-.086	.657	-.081	.665	.671	.169	.052
2	-.057	-.015	.020	.012	-.064	-.073	-.072	-.007	-.001	.713	-.044
3	.166	-.009	-.022	-.020	.096	.142	-.006	.162	.147	.457	-.016
6	-.074	-.030	.006	-.006	-.019	-.006	.044	-.058	-.040	-.018	-.003
8	.047	-.067	.008	.032	.194	-.040	-.078	-.035	.027	.049	-.038
10	-.234	+.064	.073	-.030	.012	+.027	.148	-.068	-.120	.048	.120
11	-.111	+.040	-.042	-.073	-.053	-.158	.026	-.120	-.186	-.096	.010
13	.084	.004	-.009	.028	.140	.022	-.056	.018	.014	-.006	-.024
19	.022	.042	.259	.024	-.052	.290	.064	.058	.039	-.064	.890
20	-.113	.037	.301	.029	-.052	.681	.070	-.005	-.011	-.108	.355
21	.082	.049	.256	.009	-.024	.245	.068	.077	.048	-.036	.910
24	-.124	.173	.032	-.050	-.045	.055	.389	.048	-.093	-.043	.227
25	-.129	.184	-.134	-.085	.035	-.074	.941	-.011	-.158	-.088	.060
26	.166	-.013	.093	-.035	.018	.115	-.052	.060	.029	-.047	.067
32	.391	-.043	-.042	-.009	.081	-.066	-.160	.102	.130	.100	-.065
35	-.050	-.028	-.120	-.002	-.047	-.109	-.072	-.021	-.013	.062	-.150
36	-.084	-.048	-.094	-.038	-.089	-.169	-.055	-.107	-.095	-.017	-.170
37	-.037	.050	-.051	.007	.040	-.079	-.030	-.020	-.040	-.015	-.066
Anti-depressives	-	-.048	.096	-.001	.069	.033	-.118	.070	.095	.070	.053
Drugs for physical disorders		-	-.003	.008	.002	.031	.172	.711	.019	-.006	.050
E.C.T.			-	.081	.037	.264	-.126	.049	.080	.002	.281
Group Therapy				-	.166	.049	-.070	.042	.729	-.024	.017
Improvement Criterion					-	.107	.012	-.046	.052	.019	-.037
1 + 20						-	-.009	.464	.462	.039	.287
1 + 25							-	.057	-.092	-.061	.076
1 + Drugs for Physical Disorders								-	.468	.111	.071

Table 4

Simple Correlations Among a Selection of Original and First-Order Variables Which Maximize Prediction[1,2]

	Anti-Depressives	Drugs for Physical Disorders	E.C.T.	Group Therapy	Improvement Criterion	1 + 20	1 + 25	1 + Drugs for Physical Disorder	1 + Group Therapy	2 + 3	19 + 21
1 + Group Therapy									-	.085	.045
2 + 3										-	-.053
19 + 21											-

[1]This Table excludes correlations already reported in Appendix Table 3

[2]See Appendix Table 1 for variable names corresponding to numbers in this Table.

Table 5

Treatment Formulations Derived From a Successful

Multiple Regression Technique

Variable or Interaction	GROUP													
	A		C		D		E		F		K		X	
Age	.088	.016[1]	.109	.018	.001	.000	-.199	-.042	.233	.040	.185	.034	-.022	-.004
Sex	.068	.334	.306	1.503	-.267	-1.343	-.435	-2.107	.278	1.403	-.087	-.506	.375	2.046
Marital Status	.091	.246	.327	.910	.113	.342	.032	.081	.354	.979	.252	.813	.014	.042
Religion	.036	.191	.063	.343	-.024	-.136	.075	.379	-.106	-.611	-.002	-.010	.010	.063
Employment at Admission	.191	.479	.041	.108	.179	.438	.133	.328	.102	.268	.194	.520	.216	.612
Voluntary-Involuntary Admission	.119	.585	.030	.150	.049	.263	-.006	-.031	.159	.802	-.048	-.288	.016	.083
Source of Admission	-.064	-.366	-.042	-.231	-.010	-.068	-.093	-.524	.006	.030	.007	.043	-.028	-.173
Duration of the Precipitating Episode	-.105	-.151	-.196	-.265	-.151	-.082	-.114	-.163	-.147	-.190	.024	.046	-.010	-.016
Stream of Thought	-.114	-.561	-.183	-.902	-.120	-.581	-.029	-.186	.014	.076	-.341	-2.061	-.255	-1.408
Thought Content	.034	.168	.166	.812	-.244	-1.207	-.172	-1.250	.286	1.893	-.442	-2.397	-.168	-.897
Motor Activity	.020	.100	.068	.335	-.123	-.599	.190	1.068	.072	.404	-.118	-.714	-.090	-.488
Memory	-.008	-.062	-.134	-.741	-.098	-1.011	-.032	-.198	.086	.736	-.077	-.767	-.035	-.258
Brain Disease	.186	1.518	.014	.072	.139	1.992	.099	1.646	.290	2.516	.366	7.67	.189	1.831
Involutional Psychosis	-.019	-.412	.035	.611	.036	.367	-.034	-.565	.013	.316	.007	.057	.054	.627
Depressive Neurosis	.030	.210	.065	.501	.099	.479	.018	.174	.192	1.259	.039	.206	.197	1.146
Pattern Personality Disorder	-.021	-.213	.022	.323	.016	.197	.052	1.026	.040	.365	-.064	-1.631	.012	.255
Sociopathic Personality Disorder	-.097	-.995	.001	.011	-.107	-1.902	-.087	-2.194	.033	.247	.077	2.78	.054	.880
Transient Situational Personality Disorder	.032	.303	-.027	-.303	.132	1.814	-.045	-1.139	.058	.376	-.094	-3.39	.078	.978
Group Therapy	.248	1.257	.390	1.961	.132	.650	.032	.157	.119	.630	.595	3.181	.665	3.728
Sex X Thought Content	-.129	-.299	-.242	-.540	.234	.531	.200	.467	-.423	-1.288	.479	1.11	.265	.641

Table 5

Treatment Formulations Derived From a Successful

Multiple Regression Technique

	GROUP													
Variable or Interaction	A		C		D		E		F		K		X	
Sex X Physical Drugs	-.129	-.759	..048	.192	-.049	-.487	-.116	-2.540	-.216	-1.457	-.354	-5.243	-.165	-1.017
Sex X Group Therapy	-.185	-.443	-.317	-.734	.029	.069	.087	.221	-.060	-.142	-.581	-1.582	-.803	-1.977
Stream of Thought X Motor Activity	-.007	-.013	.102	.177	.183	.318	-.086	-.179	-.252	-.528	.604	1.312	.360	.723
Age X Marital Status	-.018	-.001	-.336	-.014	-.062	-.003	.269	.013	-.461	-.022	-.148	-.007	.012	.001
Constant		7.902		6.818		12.27		13.66		3.376		8.775		5.629

The first number is the standardized regression weight and the second is the raw score regression weight.