WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

AN LP PLANNING MODEL FOR A MENTAL HEALTH COMMUNITY SUPPORT SYSTEM

H. Stephen Leff[†]
Maqbool Dada[*]
Stephen C. Graves[*]

WP 1520-84

December 1983

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

AN LP PLANNING MODEL FOR A MENTAL HEALTH COMMUNITY SUPPORT SYSTEM

H. Stephen Leff[†]
Maqbool Dada[*]
Stephen C. Graves[*]

WP 1520-84 December 1983

DRAFT

Please do not cite or quote without permission.

[*] Alfred P. Sloan School of Management, Massachusetts Institute of Technology.

[†] Laboratory for the Study of Community Support Systems and Management Research Associates, Cambridge Hospital Department of Psychiatry.
Harvard Medical School.

ABSTRACT

In this paper we develop a resource planning and allocation model for a mental-health community-support system. The intent of the model is to provide a tool to system managers and planners to address the relevant tradeoffs with regard to resource planning and allocation. More and more, mental health care is being delivered by a variety of community-based service agencies instead of total institutions. Furthermore, the service needs of a client population are continually changing, and depend upon the relative functioning of the population, which depends upon the efficacy of the services provided. Finally, the availability of resources is becoming scarcer, and is more closely scrutinized. Consequently, the system manager is faced with a very complex planning problem in which the manager must attempt to apply the available resources as effectively as possible.

To make the model operational we develop a conceptual framework for aggregating patients according to their functional level and service needs, and for specifying "service packages" that are available to the aggregate patient groups. We also estimate probabilistically the effect of giving a particular service package to a patient group, in terms of functional improvement or regression. We view the planning problem then as the assignment of service packages to aggregate groups over a multi-period planning horizon so as to not exceed the available resource levels while maximizing some measure of system welfare. For this framework we can then formulate the planning problem as a very large, linear programming problem. We are currently exploring the validity of the conceptual framework and the applicability of the LP model by doing small case studies for the mental health planning agencies of several states.

1. Introduction

In recent years, the mental health service system in the United States has undergone fundamental structural changes. The post World War II trend of increasing institutionalization of the mentally ill has been reversed. A series of federal laws, beginning with the Community Mental Health Center Act of 1963, mandates transferring responsibility for the mentally ill from the state to local communities. One way in which deinstitutionalization (the process by which patients are returned to live in the "least-restrictive" environment) has been put into operation is through the development of Community Support Systems (CSS's) that are responsible for providing care to the chronically mentally ill residing in their area. CSS's encompass therapeutic and psychosocial rehabilitation services such as community residences, day treatment programs, and sheltered workshops. They can also relate to or include facilities such as state hospitals, emergency and inpatient services of general hospitals, and nursing homes.

CSS administrators are responsible for designing and managing the treatment programs available to their patients. Depending on program availability, patients (or clients) are assigned a set of services that best satisfy their needs. As patients respond to treatment their needs change, necessitating a dynamic adjustment in program packages provided. Ideally, CSS managers would be able to design, with no restrictions, individual programs for their patients. However, the number of patients and the fact that patients compete for the same (limited) resources make this impossible. Consequently, for planning purposes at least, CSS managers have to think of their patients in aggregate terms to facilitate making trade-offs involved in allocating resources to programs and in assigning programs to patients. In this paper we present a multi-period resource planning and policy evaluation model to aid CSS administrators in making their resource-allocation decisions.

The organization of the remainder of the paper is as follows. In Section 2 we present the resource-allocation model. In Section 3 we review relevant literature and relate it to the model we propose. In Section 4 and 5 we make the model operational by first presenting a conceptual framework for categorizing patients and services, and then specifying the necessary data for the model for this framework. A hypothetical example using this data is presented in Section 6. In Section 7 we offer some concluding remarks.

2. The Resource-Allocation Model

CSS managers, with their limited resources, attempt to devise programs that facilitate the functioning and integration of the mentally ill into the community. To accomplish this objective, it is useful to provide a framework in which to structure the resource-allocation decision facing them.

The modeling of the resource-allocation problem has three components: 1) Patient Aggregation, 2) Design of Programs, and 3) Measurement of Performance.

For client aggregation, we assume that the population of patients can be clustered into categories such that patients within each category have similar needs and have similar responses to treatment. For this model application we employ a functional level categorization (see Section 4).

The mentally ill, as a class, are extremely heterogeneous and consequently have been very difficult to categorize. Yet for planning purposes it is essential to aggregate patients in some way to reduce the scope of the planning problem. We desire to classify the chronically mentally ill into relatively homogeneous groups called functional levels that are meaningful for program planning. In particular, the functional levels must be defined so that the progression (regression) of patients can be viewed as improvements (reductions) in functional levels. Furthermore, the functional levels must make clinical sense; patients in different functional levels will have different service needs and will receive different service and treatment packages. Being able to classify patients in this manner permits one to model a mental health care system as a dynamic process, in which patients move from one functional level (FL) to another depending upon the services provided.

For treatment programs, we assume that we can identify a set of service packages where a services package can consist of a residential assignment (bed in a supervised apartment or nursing home for example), and participation in specific social, psychotherapeutic, and rehabilitative programs. Service packages must meet certain minimal needs of patients. Because patients in different FL's have different service needs, not all service packages are appropriate for all FL's. Service package options are discussed more fully in Section 4.

Patients improve or regress based on the service packages assigned. In particular we assume that a Markov property applies to patient transitions. Namely, the probability that a patient in FL i makes a transition (improvement, regression, or no change) to FL j within a certain period of time depends only on the current state of the patient (i.e. FL i) and on the service package that is assigned to the patient. Inherent in this assumption is an underlying time period for modeling patient transitions as well as for making resource decisions. Furthermore for planning purposes we assume that patient transitions occur exactly at their expected value despite the probabilistic nature of individual patient movements. For instance, suppose the probability of improvement is .3 for each patient in FL i that receives a service package k; then if there are 100 patients in FL i receiving service package k, exactly 30 of these patients will improve in the immediate time period. We believe that this ignoring of transition randomness is appropriate for and consistent with the development of an aggregate resource-planning model.

The assignment of service packages to patients is restricted by a set of resource constraints. We assume that each service package consumes a certain amount of each resource per patient per period and that the system has limited resources in each time period. The model assigns service packages to patients in each period of the planning horizon, explicitly taking into account the end-of-period transitions and resource constraints, to optimize some multi-period (linear) measure of system performance or welfare.

Since CSS goals vary we propose using several measures of performance.

In Section 6 we illustrate various types of measures that might be used.

To present the linear program formulation we use the following notation:

T number of periods in planning horizon

t index for periods $(t=1,2,\ldots,T)$

M number of functional levels

i,j indices for functional levels $(i,j=1,\ldots,M)$

K number of service packages

k index for service packages $(k=1,\ldots,K)$

L number of resources

ℓ index for resources $(\ell=1,\ldots,L)$

n_i^t number of new patients entering functional level i at beginning of period t

$r_{k\ell}$ amount of resource ℓ consumed per period by one patient receiving service package k

R_ℓ^t amount of resource ℓ available for period t

p_{ijk} probability of transition from FL i to FL j after receiving service package k for one period

X_{ik}^t decision variable denoting number of patients in FL i receiving service package k in period t.

The problem is:

Problem (P):

$$\max f(\cdot) \qquad (1)$$

subject to

$$\sum_k X^1_{jk} = n^1_j \qquad j=1,2,\ldots,M \qquad (2.1)$$

$$-\sum_i \sum_k p_{ijk} X^{t-1}_{ik} + \sum_k X^t_{jk} = n^t_j \qquad t=2,\ldots,T;\ j=1,2,\ldots,M \qquad (2.t)$$

$$\sum_j \sum_k r_{k\ell} X^t_{jk} \leq R^t_\ell \qquad t=1,\ldots,T;\ \ell=1,2,\ldots,L \qquad (3)$$

$$X^t_{jk} \geq 0 \qquad \text{all } j,\ k,\ t.$$

Since not all service packages are appropriate for all functional levels, summations run over allowable combinations only. Equation (1) corresponds to any linear objective function. Constraints (2.1) ensure that all patients in a functional level are assigned to service packages in period 1. Constraints (2.t) are flow conservation constraints that allow transitions to occur at their expected values, i.e., if X^t_{ik} patients from FL i receive service package k in period t, then $(p_{ijk})X^t_{ik}$ of them move to FL j at the end of period t. (3) are the resource availability constraints.

It is important to observe that although (P) is a large scale optimization system it has special structure. In particular, the decision variables X^t_{jk} appear in two consecutive periods only; thus the problem exhibits the well-known staircase structure. Several decomposition schemes are available for solving staircase linear programs. We use the algorithm due to

Ho and Manne (1974) because it is easy to interface with existing LP codes and allows the use of an ε-optimal termination criterion.

Although we have modeled constraints in terms of resources only, the formulation provides for a rich variety in the constraint set. The illustration described in Section 6 demonstrates some of the flexibility of the model.

3. Literature Review

The proposed model is related to earlier work that relied on management-science methodology for program evaluation and resource planning in the health services area. We briefly review here this work. For a broader review of use of management-science/operations-research methods in the mental health field, see Kessler (1981a, 1981b).

Trinkl (1974) used a Markov model to evaluate proposed programs for the mentally retarded in Hawaii. Patients were classified into functional levels on the basis of two factors, their degree of independence in the living environment and their level of self-sufficiency in the work environment. Transition probabilities were estimated for each of five proposed treatment programs. Values or weights were assigned to each type of transition to indicate the desirability of the transition. Then a Markovian analysis was conducted for each of the five program choices to determine the expected outcome of that program. Constraints, such as capacity of sheltered workshop programs, were handled in an elementary manner by assessing penalties for constraint violations. Programs were evaluated by comparing ratios of value (benefit) to cost, where value was computed based on transition weights and costs consisted of program costs and the penalties assigned for violated constraints.

Meredith (1976) reports on a study of hospitalized mentally-retarded children. He classified patients into six functional levels on the basis of their scores on eight self-help categories. Rather than choose a single best program, his goal was to find, for each functional level, the optimal choice from a variety of programs not all of which were available at every functional level. Benefits were measured in dollars on a linear scale using the cost of the worst program in the worst state as a base; a constant (dollar) amount of benefit was added to each improved state. Howard's (1960) policy-iteration technique was used to find the set of programs that maximized the long-run expected discounted value, where the program values represented benefits minus costs. The solution method did not consider resource constraints such as budgets and capacity limits on bed types. If an unconstrained solution violated some constraint it was discarded. A heuristic was then employed to generate "the next-best" policy. If this policy violated some constraint, the heuristic procedure continued until a feasible solution was found.

Although useful for program evaluation, the models of Trinkl and Meredith are not totally satisfactory as resource-allocation tools because they do not explicitly handle resource constraints. Zarin (1979) developed a resource allocation model for use in a community mental health care system. Her model served as the initial basis for the work reported in this paper. Zarin assumed patient transitions are Markov and occur at their expected values at the end of each period. Each time period Zarin solved a one-period linear program that assigns patients to services, primarily residential, at minimum cost. At the end of that period patient trasitions were "simulated" at their expected value levels to provide the input for the next period's LP. Because the model is decoupled into a series of single-period allocation problems, considerable suboptimization can occur. However, Zarin's approach is a useful way to conceptualize the system behavior and to incorporate resource constraints that may change over time.

Schaifer and Schaifer (1981) make the same assumptions about system behavior as Zarin in developing a detailed deterministic simulation model to evaluate resource allocation decisions for the department of mental health of the state of Ohio. Administrators input a detailed plan including the maximum number of patients from each FL assigned to each service package, the service package-dependent transition probabilities, and the resources needed to provide each service package. The simulation provides as output the costs of the plan, the amounts of each resource required to meet the plan, and reports on patient transitions so that the plan's impact on patient welfare can be evaluated. However, the simulation is incapable of identifying an optimal plan. Different plans are proposed until a satisfactory program is determined, which may or may not be near the optimal plan.

Heiner, Wallace and Young (1981) developed a single-period goal programming model for allocating resources to provide service to the mentally retarded. As in the models discussed above patients are clustered into FL's. However, in a departure from other models, the goal programming model creates one service package for each FL by determining how much of each resource the "average" patient in that FL will receive. Since it is assumed that the expected impact of each resource on improvement in functional skills is measurable, it is possible to compute the expected improvement in functional skills in the "average" patient that results from a particular allocation of resources to that functional level. To ensure that all patients receive some minimal services, the authors specify goal constraints setting minimal expected improvements in functional skills for each FL. Since the goal programming model is a linear program, considerable flexibility is provided in representing resource constraints. Hence, Heiner et al. are able to specify goal constraints, cost constraints, and manpower constraints such as capacities in vocational counseling, occupational therapy, and physical therapy.

The resource allocation model presented in this paper may be viewed as a generalization of the approach used by Zarin. Besides allowing for a wider set of resource constraints, the model explicitly incorporates the dynamic system behavior by linking interactions between consecutive periods in the constraint set to create a large-scale multi-period linear program. The resource-allocation model is also realted to the goal programming model [Heiner et al., (1981)] to the extent that both allow similar cost, service, and resource constraints. However, whereas the goal programming model is a single-period model, our model is multi-period and permits consideration of dynamic behavior and constraints. We believe that it is essential to consider these dynamic consequences of any resource plan when conducting a planning exercise.

4. Development of a Conceptual Framework

To test the model we need first to make specific the conceptual framework. This requires the following tasks:

a) identify a taxonomy of categories between which patients can move in response to services;

b) for each category identify a set of service packages that are available to planners or clinicians to assign to patients.

In this section we report upon our efforts at developing a specific framework for CSS planning. We note that the resource-allocation model is not restricted to the specific framework given here but may be applicable for any framework for conceptualizing the interaction between patients and services.

Functional Levels

The category system we developed is for the chronically mentally ill and is shown in Table 1. The system consists of seven major functional levels (FL's) ranging from Dangerous to Fully Independent. We define the various FL's with respect to six dimensions of functioning:

(1) The ability and willingness to control behaviors dangerous to self or others and/or cooperate in own care.

(2) The ability to behave in a relatively symptom-free manner.

(3) The ability to perform activities of daily living (e.g. personal care and grooming).

(4) The ability to perform tasks necessary for community living (e.g. taking public transportation and shopping).

(5) The ability to engage in new problem solving.

(6) The ability to use natural support systems.

As Table 1 shows, "Dangerous" persons are those who do not manifest the ability to control behaviors dangerous to themselves or others. Such

Table 1: Dimensions of Functioning Associated with Functional Levels

Dimension of Functioning	FUNCTIONAL LEVEL						
	(I) Dangerous	(II) Acute	(III) Residual	(IV) Dependent	(V) Vulnerable	(VI) Recovering	(VII) System Independent
Ability and willingness to control impulses and cooperate in own care.	o	+	+	+	+	+	+
Ability to behave in relatively symptom-free manner.	NA	o	+	+	+	+	+
Ability to perform activities of daily living.	NA	o	o	+	+	+	+
Ability to perform community living skills.	NA	o	o	o	+	+	+
Ability to engage in new problem solving.	NA	o	o	o	o	+	+
Ability to use natural support systems.	NA	o	o	o	o	o	+

NA - not applicable

\+ - yes

o - no

persons may differ in their other characteristics, but these differences are considered irrelevant to their immediate service needs. "Acutely Disturbed" persons are not dangerous, but exhibit substantially high levels of symptomatic behavior. "Residually Disabled" persons are neither dangerous nor relatively symptomatic, but are unable to cope with the simple tasks of daily living. Needless to say they also lack the higher-level skills necessary to function in a community environment. The remaining FL's are defined similarly, with each succeeding level having greater skills.

Table 2 shows the relationship postulated between FL and service need. We define service need on three dimensions:

(1) need for residential security, i.e. physical security and/or intensive staffing;

(2) need for residential supervision where continuous need is 24-hour supervision, partial need is regular but less than 24-hour supervision, and intermittent need is as needed;

(3) need for treatment and rehabilitation services.

As Table 2 shows, we associate with every FL its own unique pattern of service need. Each pattern represents the amount of security, supervision, and treatment and rehabilitation services that patients at the FL are postulated to need in order to meet minimum standards of safety and clinical appropriateness.

The treatment and rehabilitation column shows that all clients except those who reach the point of being fully independent, need these services. As one moves from one FL to another the quality and the quantity of these needs change. In general we assume that the extent of treatment and rehabilitation needs diminishes as FL increases, although in certain areas (e.g. vocational rehabilitation) needs change rather than disappear. However, it should be noted that the definition of treatment and rehabilitation

Table 2: Relationship Between Functional Levels and Service Dimensions

Functional Level	SERVICE DIMENSIONS		
	Security	Supervision	Treatment and Rehabilitation
(I) Dangerous	Maximum	Continuous	More Extensive
(II) Acute	Moderate	Continuous	
(III) Residual	Minimum	Continuous	
(IV) Dependent	None	Partial-Regular	
(V) Vulnerable	None	Intermittent	
(VI) Recovering	None	None	Less extensive
(VII) System Independent	None	None	None

service needs also will be a function of system objectives or aspirations. We discuss this point in the subsection on service options.

One weakness of coupling this classification system and the assumption of Markov transitions is the inability to take account of client history. For example, we are unable to distinguish between patients entering an FL for the first time and patients remaining in or returning to an FL. The FL classification system assumes that the current status is a sufficient descriptor of all past history. This need not be the case. Research on recidivism supports the concept of different prognoses for "new", "long-stay", and "revolving-door" chronically mentally ill persons, particularly at the lower FL's. To correct for this limitation we subdivide both FL I (dangerous) and FL II (acute) into two subcategories each: one for patients with a good prognosis (I_G and II_G), the other for patients with a poor prognosis (I_P and II_P). Generally, new patients to the system will have a good prognosis, which shifts to a poor prognosis the longer they remain in the system. Patients with a good prognosis will have a higher likelihood for improvement than those with a poor prognosis. We make this distinction only at the lower FL's where it is most important to make this differentiation.

We do not intend for this FL system to be useful to clinicians, but rather have designed it to be meaningful for resource planning and management. However, we have found that this FL system can be related to clinical scales such as given in Endicott et al. (1976). In Leff (1983) we discuss this relationship.

Service Package Options

We developed a framework for conceptualizing service options for use with our model. As in the case of the FL system, we emphasize that other frameworks can be used with the model. The one we present reflects our own theory as to what is important to planning CSS's, corresponds to our FL system and is useful in categorizing services as they are described in the program evaluation literature.

We postulate that CSS clients require services in at least four dimensions: the residential dimension, the psychotherapeutic treatment dimension, the social rehabilitation dimension, and the vocational rehabilitation dimension. Case management ideally should integrate these services into service packages that are consistent with client FL's in the security and supervision (level of structure) they provide.

We also postulate that different service systems can have different objectives or aspirations for their clients. These aspirations play a role along with other factors in determining how client service needs are defined and what services are included in the service packages offered to clients. We refer to service systems that aspire primarily to meet clients' subsistence needs as having custodial aspirations. Alternatively, service systems may aspire to increase client adaptation to their current FL's. We describe such systems as having adaptive aspirations. Finally, we refer to systems that aspire to increase client FL's as much as possible, as having promotive aspirations.

Systems with custodial aspirations provide "room and board", but few treatment and rehabilitation services. Systems with adaptive aspirations provide both "hotel" and pschotherapeutic treatment services, but few rehabilitative services. Systems with promotive aspirations provide psycho-social rehabilitative services in addition to "hotel" and psychotherapeutic or counseling services.

Table 3 presents illustrative service package options (SPO's) for each structure and aspiration level. In total we have eighteen SPO's. As suggested by Table 3 there are specific SPO's for each FL. In particular we would

Table 3: Illustrative Custodial, Adaptive and Promotive Service Package Options for Each Demand and Aspiration Level*

Level of Structure	ASPIRATION LEVEL		
	Custodial (C)	Adaptive (A)	Promotive (P)
(I) Maximum Security Continuous Supervision	State Hospital	Private Mental Hospital	Quarter Way House
(II) Moderate Security Continuous Supervision	State Hospital	General Hospital	Crisis Hostel
(III) Minimum Security Continuous Supervision	Nursing Home	Supervised Community Res. Day Treatment	Supervised Comm. Res./ Day Treatment/ Prevoc. Training
(IV) No Security Partial, Regular Supervision	Board & Care	Foster Home Day Center	Foster Home/Day Center/ Sheltered Employment
(V) No Security Intermittent Supervision	Rooming House	Supervised Apartment Social Club	Supervised Apartment/ Social Club/ Transit. Employment
(VI) No Security No Supervision	Case Management	Case Management Alumnae Program	Case Mgt./Alumnae Program/ Job Development

* This table draws upon commonly held ideas about the typical aspiration levels of services in different types of facilities and programs. However, it is conceivable that a state hospital could offer a promotive service package, while a supervised apartment would be a part of a custodial one.

expect that patients in each FL would be placed with the SPO's that at least satisfy their needs for security and supervision (structure). For instance, patients in FL II (acutely disturbed) would be eligible only for the three SPO's that provide moderate security and continuous supervision; these patients could then receive SPO's with custodial, adaptive or promotive aspirations. We do not consider combinations that provide patients with more or less structure than is warranted by their FL's. Providing clients with SPO's that are inadequately structured is what has been called "dumping". Supplying clients with SPO's that provide unnecessary security and supervision violates the principle of placing the patients in the least restrictive environment. In our modeling activities to date we have only explored the impacts of offering clients SPO's that vary in their aspiration levels, but not SPO's that would be considered "dumping" or "overly restrictive". Nevertheless, such strategies might be modeled if such options are being exercised or considered.

5. Model Specification

To test the model we need not only to develop a conceptual framework for the model, but also to perform the following data-related tasks:

a) specify a time period for the model;

b) specify the transition probabilities (p_{ijk}) for all relevant combinations;

c) specify the resource requirements for each SPO as well as the constraints on resource usage;

d) specify the initial census and the number of arrivals to the system each perio for each FL;

e) specify the objectives to be optimized.

This section reports on our efforts at model specification.

Model Time Period

We model the dynamic CSS with a multi-period model in which patients arrive for service at discrete epochs, services are rendered in unit "dosages" defined by the period length, and patients improve or regress at periodic intervals in direct response to the services assigned. The choice of the model time period is critical for the model to properly represent the underlying dynamic system. With a time period that is too large we cannot fully represent a patient's movement through the system; however, the computational complexity and storage requirements for the model increase quickly as the time period is reduced. We choose a time period of one month with a planning horizon of 36 months. The critical factor in selecting this time period is the rate at which CSS services can affect patients. For a number of the services that we consider one month seems to be the largest allowable time period for capturing patients' movements between FL's.

Transition Probabilities

Associated with every permissible combination of an SPO and an FL is a probability distribution giving the percentages of patients in the given FL that move to other FL's after receiving the SPO for one time period. These transition probabilities (TP's) may be considered measures of SPO effectiveness. Presumably the greater the probability that an SPO moves a patient to a higher FL the greater is its effectiveness, and vice-versa.

Estimating these TP's is an enormous task. To our knowledge no comprehensive data base on CSS services exists for estimating these TP's. Nor could we expect localities to have such data bases to be used in specific model applications. As a consequence of this, we attempted to create our own data base by reviewing and then synthesizing the recent evaluation literature for CSS programs. We first identified a large set of published and unpublished articles and reports describing the functioning of CSS programs. Once an article was identified as potentially having relevant data, trained coders extracted TP data by means of the following operations.

(1) Code SPO's. First, the coders identified the different SPO's offered to clients. Most often these would be a "model" program(s) and a comparison program(s) representing the "usual" mix of services offered. In some cases a sequence of "model" programs would be describe. Each SPO was then categorized with regard to demand and aspiration level (i.e. Table 3) based on the descriptions of the services provided in the reports.

(2) Code Initial FL's. Next, coders categorized each cohort of clients assigned to an SPO according to the average or modal FL of the cohort prior to receiving the SPO. These codings were done on the basis of the ways clients in the cohort were described in the reports prior to their assignment to SPO's.

(3) Map Client Pathways. Following the coding of initial client FL's, coders mapped the transitions to other FL's made by clients after receiving SPO's. Mapping a transition required coding the FL of a client cohort based on its "post-treatment" description and specifying the time period for the transition indicated. If a sequence of SPO's was provided, then multiple transitions were mapped. In addition the number of clients involved in both the origin FL and destination FL's were specified.

(4) Determine Transition Probabilities. Once the client pathways had been mapped, transition probabilities could be calculated. Since the transition probability is linked to the time period for transition and since the time periods were described in different ways in different articles, methods for converting different descriptions of "treatment time" into a common metric had to be adopted. To simplify this calculation we assumed that the unit of time for a transition is one month and that the transition probabilities depended only on the current state of the patient, and not on the length of time that the patient has been in the current state.

Using these methods for estimating TP's from individual studies, it was possible to summarize the information in each report in terms of how many clients moved from origin FL's to destination FL's after receiving various SPO's for one month periods. We derived data from 175 cohorts containing approximately 8000 clients. Once the data from individual reports were standardized in this form it was then possible to aggregate across studies. Data were acquired for all but three FL and SPO combinations: FL I_G clients in custodial SPO's; FL I_G clients in adaptive SPO's; and FL III clients in adaptive SPO's. These data were interpolated based on expert opinion.

Similarly we used expert opinion to estimate the rate at which clients exited from the system due to death or disappearance.

Table 4 shows the TP's for clients at each FL receiving appropriately structur custodial (C), adaptive (A), and promotive (P) SPO's. It also shows the numbers of clients from which the TP's were estimated.

The method described above can be considered a limited form of meta-analysis. Meta-analysis employs techniques for aggregating the results from different studies and analyzing these results as a function of study level variables (Glass et al., 1981). We believe there is great potential in applying meta-analytic techniques to developing a modeling data base. Leff (1983) provides a more detailed description of these efforts for estimating the TP's.

Resource Requirements and Constraints

The model permits the inclusion of multiple resources for each SPO. For our test study we consider two resources, the per-patient cost of the SPO, and the requirement for a community residential bed by the SPO. Then our resource constraints are per capita budget constraints on total expenditures, and limitations on the number of beds in community programs.

In model applications SPO costs will be constructed based on data provided by local planners. For the hypothetical illustration presented here, we first attempted to extract costs from the evaluation literature that we had used to estimate TP's. However, this literature either did not report costs or was highly inconsistent. Consequently, we estimated costs based on a small number of reports that contained complete cost data [primarily Brewster vs. Dukakis (1978), but also Ashbaugh (1983), Weisbrod et al. (1980) and Rubin (1978)].

The most detailed and complete information was contained in Brewster

Table 4: Transition Probabilities for Clients at Each Functional Level Receiving Custodial, Adaptive, and Promotive Service Package Options

INITIAL		OUTCOME FUNCTIONAL LEVEL AFTER ONE MONTH										Number in Cohorts
FL	SPO	I_P	I_G	II_P	II_G	III	IV	V	VI	VII	EXIT	
I_P	C	.86	0	0	0	.03	0	.08	.02	0	.01	1126
	A	.86	0	0	0	.04	.01	.08	0	0	.01	273
	P	.80	0	0	0	.15	.04	0	0	0	.01	138
I_G	C	.13	.51	0	0	0	.35	0	0	0	.01	0
	A	.13	.51	0	0	0	.35	0	0	0	.01	0
	P	.11	.48	0	0	0	.40	0	0	0	.01	86
II_P	C	0	0	.52	0	.31	.09	.04	.03	0	.01	109
	A	.06	0	.56	0	.08	.01	.21	.07	0	.01	284
	P	.03	0	.36	0	.05	.05	0	.50	0	.01	60
II_G	C	0	0	.01	.15	0	.82	0	.01	0	.01	453
	A	0	0	0	.09	.25	.27	.28	.10	0	.01	1761
	P	0	.11	.14	0	.01	0	.35	.32	.05	.01	350
III	C	0	0	.04	0	.94	.01	0	0	0	.01	645
	A	0	0	.03	0	.87	.01	.04	.04	0	.01	0
	P	0	0	.02	0	.80	.01	.08	.08	0	.01	303
IV	C	0	0	.04	0	0	.94	0	.01	0	.01	584
	A	0	0	.05	0	0	.94	0	0	0	.01	569
	P	0	0	0	0	0	.86	.13	0	0	.01	101
V	C	0	0	.01	0	0	0	.98	0	0	.01	162
	A	0	0	.01	0	0	0	.97	.01	0	.01	443
	P	0	0	0	0	0	0	.98	.01	0	.01	229
VI	C	0	0	.03	0	0	0	0	.96	0	.01	312
	A	0	0	0	.02	0	0	0	.96	.01	.01	1130
	P	0	0	0	0	0	0	0	.97	.02	.01	165

vs. Dukakis, in which costs had been very carefully constructed. We estimated costs by matching our SPO codes to the services described in Brewster vs. Dukakis. It was necessary at times to scale certain costs to reflect the fact that not all clients in an FL would need all services in the SPO for the entire time period. We checked these cost estimates against the costs presented in the other studies, taking into account differences associated with geographic location and year of study. In the few instances in which discrepancies could not be explained in terms of geographic or temporal factors, an average of the disparate costs was taken. Table 5 shows the costs developed.

We also had limits on the number of beds in supervised apartments or community residences. This constraint reflects the difficulty of setting up apartments and houses for the chronically mentally ill stemming from factors such as low vacancy rates and community resistance. For our set of SPO's we postulated that only IP, IIIA, IIIP, IVA, IVP, VA, VP would require beds in supervised apartments or houses (see Table 3).

System Census and Arrival Rates

For each FL the model needs an initial census of patients as well as an estimate of the number of new arrivals from outside of the system to the FL in each time period in the planning horizon. Arrivals to the system can consist of persons just becoming mentally ill (new incidence), those ill for some time but outside the service system (latent demand), and persons reentering the system after "dropping out" of care for some reason. In model applications these estimates could be derived from expert opinion, special surveys, or management information systems. For our model illustration we derived estimates for a hypothetical locality with a population of one million from a small number of reports dealing with the prevalence and incidence of

Table 5: Monthly SPO Costs for Each FL in Dollars

FL	Custodial	Adaptive	Promotive
I	1650	2223	3380
II	1650	4611*	2374
III	928	1682	2064
IV	671	1489	1830
V	351	809	1258
VI	249	626	867

* II A is more expensive than II P because of its exclusive use of general hospital care.

Table 6: Estimates for each FL of both numbers in care at CSS startup and monthly arrival rates for a hypothetical locality with a population of one million.

FL	Number at CSS Startup	Monthly Arrival Rate*
I_P	67	.06
I_G	20	8.16
II_P	42	14.37
II_G	243	14.37
III	812	21.26
IV	1419	4.91
V	1449	17.47
VI	1561	19.35

* May depend on service availability.

chronic mental disorder and the demand for CSS services [e.g. Brewster vs. Dukakis (1978) and Gerhard (1980)]. The method for estimating these data was similar to the method used to estimate costs. We applied the FL taxonomy to reported estimates of prevalence and incidence or demand, combined with expert judgement. Table 6 shows for each FL, the population-based estimates of numbers in care of CSS at startup and the monthly arrival rates used in the modeling analyses reported below.

Model Objectives

One goal of CSS's is to increase client FL in a cost-effective manner. This is the one that we have adopted in our conceputal framework. The goal of increasing patient FL can be approached in several different ways. One way is to seek service delivery strategies that maximize patient forward movements, where a transition from FL i to FL i+j counts as j steps or movements (j > 0). Alternately, one could individually weight each forward transition, and then maximize weighted forward movements; for instance, a transition from FL I to FL III might be weighted more than a transition from FL II to FL IV. A second approach is to seek service delivery strategies that minimize (weighted) backward movements, while a third is to seek strategies that maximize net forward movements (forward minus backward movements, possibly weighted). Yet another approach would be to seek strategies that maximize or minimize the number of patients in particular FL's. Examples of such approaches would be ones seeking strategies to minimize the number requiring secure, supervised SPO's (e.g. patients requiring hospital beds) or to maximize the number of patients reaching full independence. Here again one might also approach the objective by seeking a mixed strategy that both minimizes the number of clients in some FL's and maximizes the number in others.

In our modeling we have used the following three objectives:

(1) maximize net forward movements (Z1)

(2) minimize census in FL's I and II at the end of the planning horizon (Z2)

(3) minimize total census at the end of the planning horizon (36 months), where the census is the number of patients in FL's I, II, ..., VI (Z3).

6. Illustrative Example

Using the model specification of the previous section we analyzed 14 different scenarios using three different objective functions of Section 5. The scenarios reported in Tables 7, 8, and 9 only constrain the average monthly per capita budget.

The last three columns of Tables 7-9 show the values of the actual objective functions evaluated on the optimal solution. Examination of these three columns reveals that the system performance improves sharply as the per capita budget is increased from $600/month to $800/month. Although the system performance continues to increase as the budget is increased to $1300/month, the change is relatively small. This suggests that there is a threshold budget level around $800/month below which system performance deteriorates sharply.

The linear programming solution determines what SPO's to assign to each FL in each period. By examining the assignments of SPO's over time, we can determine the dominant SPO assignment for each FL. Columns 2 through 8 of Tables 7, 8, and 9 describe this symbolically. We believe that providing the dominant SPO for each FL is an effective way for CSS managers to allocate available resources. The dominant assignments can determine what service package mix should be available at the CSS. Although clinicians might use these assignments as guidelines when prescribing treatment for individual patients, we would not expect that the assignments from the model to be of great value to the clinician.

By looking at the solutions in Tables 7-9 we can examine the sensitivity of the solution to the choice of objective function. Using objective functions Z1 and Z3 gives similar assignments. When optimizing Z2, the significant difference is that more clients in FL I and FL II are assigned promotive SPO's at the expense of clients in higher FL's.

Table 7: Maximize Net Forward Movement Rates Subject to Average Per Capita Budget

Budget	DOMINANT ASSIGNMENTS										
	I_P	I_G	II_P	II_G	III	IV	V	VI	Z_1	Z_2	Z_3
600	C	C	P	P	C,P	C	C	C	9620	225	6266
700	C	C	P	P	P	C	C	C,P	12,117	184	5857
800	C	C	P	P	P	C	C	P	13,319	148	5521
900	C	C	P	P	P	C	C,A	P	13,901	141	5360
1000	C,P	C	P	P	P	C	A	P	14,201	124	5303
1100	P	C	P	P	P	C	A	P	14,304	107	5281
1200	P	C	P	P	P	C	A	P	14,326	106	5277
1300	P	C	P	P	P	C	A	P	14,326	106	5276

Table 8: Minimize Final Census in FL I and FL II Subject to Average Per Capita Budget

Budget	DOMINANT ASSIGNMENTS										
	I_P	I_G	II_P	II_G	III	IV	V	VI	Z_1	Z_2	Z_3
600	C	C	P	P	C,P	C	C	C	9600	183	6277
800	P	C.P	P	P	P	C	C	C	11,473	110	5319
1000	P	P	P	P	P	C	A	C,P	12,800	86	5354
1200	P	P	P	A	P	P	P	C.P	11,460	71	5708

Table 9: Minimize Final Total Census Subject to Average Per Capita Budget

Budget	DOMINANT ASSIGNMENTS										
	I_P	I_G	II_P	II_G	III	IV	V	VI	Z_1	Z_2	Z_3
600	C	C	P	P	C,P	C	C	C	8976	279	5824
800	C	C	P	P	P	C	C	P	11,363	171	5165
1000	P	P	P	P	P	C	A	P	11,727	162	5078
1200	P	P	P	P	P	C	A	P	11,736	162	5076

Quantitatively, when the system objective is to maximize net forward movements (Z1), there are approximately 20% more clients in FL I and II than when we minimize the number of clients in FL I and II (Z2). However, maximizing net forward movements gives a total census (Z3) that is within 5% of optimal. When the system objective is Z2, the net forward movements are within 10% of the optimal and the total census is within 5% of the optimal. An intuitive explanation for this is that to minimize the census in FL I and FL II we want to move clients to FL's as far away from FL I and II as possible. Pushing clients away moves them forward or out of the system. When the system objective is to minimize the total census, the net forward movements are within 5% of optimal. However, the census in FL I and FL II can be more than twice the optimal. The system census is minimized at the expense of the lowest functioning clients, the ones least likely to leave the system.

Based on these scenarios, minimizing the census in FL I and FL II is the most robust and minimizing the system census is the least robust objective function.

The scenarios reported in Table 10 include a constraint on the number of supervised beds available in apartments and houses in addition to the per capita budget constraint. We assume that, initially, there are 1500 beds available and that 50 additional beds become available each quarter. By examining Table 10 we once again identify the budget threshold of $800/month. By comparing with Table 7 we conlucde that in this instance the bed constraint does not have a significant impact on system performance. Table 10 is representative of results for the other objective functions. However, we note that expert opinion suggests that the bed constraints used were unrealistically high. In future studies we plan to explore this and other constraints (e.g. manpower) more thoroughly.

Table 10: Maximize Net Forward Movements Subject to Per Capital Budget and Apartment Beds

Budget	I_P	I_G	II_P	II_G	III	IV	V	VI	Z_1	Z_2	Z_3
600	C	C	P	P	C,P	C	C	C	9620	225	6266
700	C	C	P	P	P	C	C	C,P	12,117	184	5857
800	C	C	P	P	P*	C	C	P	13,311	148	5521
900	C	C	P	P	P*	C	C,A	P	13,884	141	5349
1000	P	C	P	P	P	C	A*	P	14,089	123	5317
1100	P	C	P	P	P	C	A*	P	14,120	113	5314
1200	P	C	P	P	P	C	A*	P	14.120	113	5314

* Assign custodial SPO when bed constraint is tight.

7. Discussion

In this paper we have proosed a general model for supporting resource allocation and planning in a mental health community support system. To make this model operational required the development of a conceptual framework for planning. This framework entailed classifying patients into aggregate groups and characterizing possible services provided by the community support system. Based on this conceptual framework we have estimated the necessary parameters for the planning model. We then attempted to demonstrate the potential value of the planning model with an illustration. Although we are unable to report on the results of the use of the model at this time, we are currently engaged in exploratory applications in several states. We hope to report on our experience in the near future.

ACKNOWLEDGEMENT

This research was supported by the National Institute of Mental Health Grant No. 1R01-MH33581-02, and the Whitaker Health Sciences Fund. The authors would like to give thanks to Nancy Zarin for her contributions to this research effort, to P. Clark Santos for help with the computational results, and to Mindy Schimmel, Michael Senger, Judith Natkins, and Marion Block for help in preparing the data.

References

Ashbaugh, J. W. and Allard, M. A. 1981. A Plan for the Closure of a State Mental Hospital in the State of Connecticut: Final Report. Prepared for the Connecticut Department of Mental Health by Human Services Research Institute.

Brewster et al. vs. Dukakis et al. 1978. Consent Decree in the United States District Court for the District of Massachusetts. Civil Action No. 76 4423-F.

Endicott, J. et al. 1976. The Global Assessment Scale. Archives of General Psychiatry. 33, pp. 766-771.

Gerhard, R., Dorgan, R., and Miles, D. 1981. The Balanced Service System. Responsive Systems Associates, Clinton.

Glass, G., McGaw, B., and Smith, M. 1981. Meta-Analysis in Social Research. Sage Publications, Beverly Hills.

Heiner, K., Wallace, W. A., and Young, K. 1981. A Resource Allocation and Evaluation Model for Providing Services to the Mentally Retarded. Management Science 27, pp. 769-784.

Ho, J. K. and Manne, A. S. 1974. Nested Decomposition for Dynamic Models. Math Prog. 6, pp. 121-140.

Howard, R. A. 1960. Dynamic Programming and Markov Processes. M.I.T. Press, Cambridge, MA.

Kessler, L. G. (ed.) 1981a. Operations Research and the Mental Health Service System, Vol. I, Report of Ad Hoc Advisory Group, DHHS Publication No. (ADM) 80-1018, Superintendent of Documents, U. S. Government Printing Office, Washington, D.C. 20402.

Kessler, L. G. (ed.) 1981b. Operations Research and the Mental Health Service System, Vol. II, Critical Review and Annotated Bibliography. DHHS Publication No. (ADM) 80-1019. Superintendent of Documents, U. S. Government Printing Office, Washington, D.C. 20402.

Leff, H. S. 1983. A Computer Model for Community Support System. Final Progress Report, NIMH Grant No. 1R01-MH33581-02. North Charles Mental Health Research and Training Foundation, Inc., Cambridge, MA.

Meredith, J. 1976. Selecting Optimal Training Programs in a Hospital for the Mentally Retarded. Oper. Res. 24, pp. 899-915.

Rubin, J. 1978. Economics, Mental Health, and the Law. Heath, Washington, D.C.

Schaiffer and Schaiffer. 1981. Personal Communication.

Trinkl, F. H. 1974. A Stochastic Analysis of Programs for the Mentally Retarded. Oper. Res. 22, pp. 1175-1191.

Weisbrod, B., Test, M. A., and Stein, L. 1980. Alternative to Mental Hospital Treatment: II. Economic Benefit-Cost Analysis. Archives of General Psychiatry 37, pp. 400-409.

Zarin, N. 1979. A Mathematical Model for the Deinstitutionalization of the Cambridge-Somerville Mental Health District. Unpublished S.M. Thesis, M.I.T., Cambridge, MA.