Capacity Management In Health Care Services: Review And Future Research Directions
Smith-Daniels, Vicki L;Schweikhart, Sharon B;Smith-Daniels, Dwight E
Decision Sciences; Fall 1988; 19, 4; ProQuest
pg. 889

Capacity Management in Health Care Services: Review and Future Research Directions*

Vicki L. Smith-Daniels
College of Business, Arizona State University, Tempe, AZ 85287-4706

Sharon B. Schweikhart
Carlson School of Management, University of Minnesota, Minneapolis, MN 55455

Dwight E. Smith-Daniels
College of Business, Arizona State University, Tempe, AZ 85287-4706

ABSTRACT

Health care has undergone a number of radical changes during the past five years. These include increased competition, fixed-rate reimbursement systems, declining hospital occupancy rates, and growth in health maintenance organizations and preferred provider organizations. Given these changes in the manner in which health care is provided, contracted, and paid for, it is appropriate to review the past research on capacity management and to determine its relevance to the changing industry. This paper provides a review, classification, and analysis of the literature on this topic. In addition, future research needs are discussed and specific problem areas not dealt with in the previous literature are targeted.

Subject Areas: Health Care Administration and Service Operations Management.

INTRODUCTION

Health care managers are confronted with a number of crucial decisions regarding the management of operations. Rising health care costs result in pressure on health care providers to reduce costs through more effective resource management across the entire health care system. In addition, trends toward growth and integration in health care organizations have rendered invalid much of the earlier research performed during a time when health care essentially was characterized as a cottage industry. Given these changes in the health care environment, it seems appropriate to assess previous research on health care capacity planning and management and to determine its relevance to this changing industry. In addition, it is necessary to consider those features of the new environment which should be addressed in future research.

Throughout the 1960s and early 1970s, national health policy fostered the expansion of the health care system in an effort to improve the availability of quality health care. During this time the federally sponsored Medicare and Medicaid programs were established, and employer-paid health benefits became the norm. The burden of health care costs was shifted to third parties who reimbursed health care providers under a cost-based system. Health care expenditures rapidly increased from 5 percent of the gross national product in 1965 to a current (1986) level of nearly 11 percent. The significant rise in health care costs also has been attributed to such factors as increased technological costs, an aging population with health problems, defensive medicine, excess capacity, and an increased number of well-trained specialists demanding higher wages. As a result of escalating health care costs, the government has enforced cost containment by implementing new fixed-price payment

---

*This research was partially funded by a grant from the Operations Management Center, University of Minnesota. The authors would like to thank Roger Schroeder and John Anderson for comments on earlier versions of this paper.

reimbursement methods that reward more efficient utilization of resources. Employers have attempted to limit health care spending through the use of increased insurance deductibles, health maintenance organizations (HMOs), and preferred provider organizations (PPOs).

Along with cost-containment pressures, health care managers also must contend with an extremely competitive marketplace. New types of health care organizations including freestanding surgical clinics, emergency facilities, and birthing centers compete directly with hospitals to provide ambulatory care services. Hospitals are joining investor-owned and nonprofit multi-institutional health care systems. Furthermore, because of the growing supply of physicians and other health care providers, there is increased competition for the patient market share as evidenced by extensive advertising and marketing activities.

While previous papers have reviewed the literature written on specific topics such as blood banks [85], admission scheduling [75], and surgical-suite management [71], these are somewhat dated and do not offer a comprehensive review of the literature across all facets of the capacity management problem in health care. This paper provides a review, classification, and analysis of the literature and suggests future research needs regarding capacity management in health care services.

CAPACITY MANAGEMENT IN HEALTH CARE

The management of capacity in health care involves decisions concerning the acquisition and allocation of three types of resources: work force, equipment, and facilities. Long-range capacity decisions involve the acquisition of facilities and major equipment. These decisions place physical constraints on both the quantity of services that can be delivered and the flexibility of the delivery system to significantly change its service mix in response to shifts in demand. Health care facilities can provide either inpatient or ambulatory care where inpatient facilities provide a greater and more complex mix of services than ambulatory care facilities. Thus the selection of a health care facility involves decisions on size and location as well as on the general mix of services to be offered. Other resource-acquisition decisions—those concerning work force, overtime, and subcontracting—are classified as medium-range decisions spanning a horizon of six to twelve months. Over this horizon, the appropriate resource input mix should be determined so that customer service goals are attained and budgetary limits are not exceeded. In the nearer horizon, after resources are acquired and demand can be more accurately forecasted, the scheduling process allocates available capacity to specific tasks and/or patients.

Table 1 integrates this framework of capacity decisions with the topics that have emerged in the literature. Past research has decomposed the capacity management problem into separate decision areas regarding acquisition and allocation decisions for facilities and work-force resources. The literature has examined these capacity issues in both inpatient and ambulatory care facilities. Subsequent sections will offer a review of articles in each of the categories in Table 1.

FACILITY ACQUISITION DECISIONS

Decisions involving the acquisition of physical capacity typically have addressed the issues of how much capacity is needed and where these additional resources should be located. In health care systems, facility size and location issues are particularly complex because of the interaction between the demand for services and the demographic characteristics of the population surrounding the facility. Health

Table 1: Capacity decision themes in the literature.

	Facility Resources	Work-Force Resources
Acquisition decisions	Facility location and aggregate capacity size	Hospital staffing
	Size of inpatient care units	Ambulatory care staffing
	Size of ambulatory care facilities	
Allocation decisions	Inpatient admissions scheduling	Assign workers to days and shifts
	Surgical facility scheduling	Assign workers to units
	Ambulatory care scheduling	Assign workers to tasks

care decision makers were motivated during the 1970s to use a variety of operations research techniques to analyze regional size and location planning problems as a result of government policies requiring the justification of capital improvements. As will be shown, the majority of research on facility location and size has concerned regional planning systems containing both ambulatory and inpatient health care facilities. Other facility acquisition models have studied the size of inpatient facilities, often referred to as the ''bed-allocation'' problem. We describe the bed-allocation literature in this section since this capacity problem has long-term implications for the size of an inpatient facility as well as for the types of services delivered. Research on the size of an ambulatory clinic also is included in this section, but is very limited in scope. Table 2 summarizes a representative sample of research on the topic of facility acquisition decisions.

Facility Location and Aggregate Capacity Size Decisions

Service location was first studied by Toregas, Swain, and ReVelle [110], where alternate locations were evaluated using total travel time as a measure of effectiveness. In a later study, Abernathy and Hershey [3] determined the locations of a specified number of primary health centers with the objective of maximizing utilization measured as a function of distance, socioeconomic factors, and personal preferences. Shuman, Hardwick, and Huber [96] integrated location and size decisions in locating HMO ambulatory clinics in a metropolitan area with the objective of maximizing the number of members who subscribed to the HMO. This research featured economies of scale by requiring each clinic to operate above a minimum capacity subject to a limit on the amount of capital expenditures.

Other authors proposed extensions for regional location and size decisions with regard to the level of care offered in each facility. Level of care is defined as the range of technology and subservices available. Parker and Srinivasan [83] determined the types and locations of facilities to be added to a regional health system based on a methodology that required consumers to rank attributes affecting their choice of health care facilities including a nurse practitioner's office, physician's office, or primary care center. Dokmeci [22] examined a hierarchical health care system where demand at one level determined the capacity requirements for higher levels of care. In that model, the number and location of facilities at each care level were selected based on the objective of minimizing patient travel costs and investment and operating costs (which differed according to type of facility).

The reallocation of capacity within an existing group of regional health care facilities is another acquisition decision. Ruth [89] studied the situation where beds in each hospital in a region were classified according to three care levels. The model minimized the costs of acquiring additional capacity and upgrading beds to a higher

Table 2: Literature on facility acquisition decisions.

	Methodology	Managerial Objective	Comments
Abernathy & Hershey [3]	Hooke-Jeeves pattern-search algorithm	Determine sites for a specified number of clinics according to one of four objectives: (1) maximize utilization, (2) minimize distance per capita, (3) minimize distance per visit, and (4) minimize percent degradation in utilization.	Health center location model recognizes negative effect of distance on utilization.
Cohen, Hershey, & Weiss [17]	Simulation	Determine effect of unit capacity levels in a progressive patient care hospital on utilization levels, percent of transfers blocked, and proportion of patient-days resulting from inappropriate use for each unit in the hospital.	Simulation methodology shows impact of capacity level in one unit on performance of other units.
Dokmeci [22]	Nonlinear, nonconvex programming	Determine number, size, and location of facilities to minimize facility costs and patient transportation costs subject to service requirement constraints.	Planning model for vertically integrated system provides optimal system configuration by considering provider and patient costs.
Esogbue & Singh [29]	Stochastic model	Determine number of beds in a unit to minimize total costs of refusing admission because of a bed shortage, holding an empty bed, and purchasing and operating a bed.	Model determines a cutoff occupancy value after which nonserious patients are not admitted, given a unit with a specified number of beds.
Geller & Yochmowitz [35]	Stochastic model	Determine cost of changes in maternity beds throughout region given a desired patient service level.	Methodology can be used to assess the cost of alternative bed-closings rules.
Gitlow [37]	Steepest ascent gradient search algorithm	Determine number of counseling and procedure rooms which maximizes clinic profitability subject to a physician overtime constraint.	Clinic objective function specifies revenue and cost functions, linking clinic design to overall profitability.
Goldman, Knappenberger, & Eller [38]	Simulation	Determine effects of alternative beds-to-service policies and room configurations on patient transfers, frequency with which capacity is exceeded, and number of days elective patients wait for a bed.	Model used to test performance of alternative bed allocations under varying overall occupancy levels, showing that waiting time increases significantly as percentage occupancy increases.
Hershey, Weiss, & Cohen [49]	Semi-Markov process	Determine mean utilization and service level for a class of capacity-constrained service network facilities with stochastic service times and stochastic patient movement through the facility.	Model results are exact for 2 cases; simulation with published data shows this method provides an accurate approximation in the general case.

Kao & Tung [58]	Marginal analysis	Determine assignment of beds to units to minimize expected total average unit overflows across all months of the planning horizon.	Methodology can be used to reallocate beds in an existing facility or to allocate beds to units in a new hospital.
Parker & Srinivasan [83]	Heuristic search algorithm	Determine set of additional facilities which maximizes consumer preferences subject to a budgetary constraint.	Model for expansion of an existing health care system which uses a multiple criterion consumer preference model to select additional facilities.
Ruth [89]	0-1 integer programming	Determine changes in the location, size, care level, and conformance of existing hospital beds in a region which minimize total regional modification costs subject to service demand constraints.	Regional planning model specifies increases in care levels and number of hospital beds.
Schmitz & Kwak [90]	Simulation	Determine effect of an increase in bed capacity on surgical and recovery room facility time and space requirements.	Model describes interaction between number of beds and surgical facility capacity.
Shuman, Hardwick, & Huber [96]	0-1 integer programming	Determine number, size, and location of clinics which maximize HMO enrollment subject to minimum clinic enrollment and capital constraints.	Planning model for multisite ambulatory care organizations, such as HMOs.
Sumner & Hsieh [104]	Simulation	Determine effect of the number of exam rooms, medical students, and physicians and patient demand, examination time, and physician arrival time characteristics on (1) exam room utilization, (2) patient waiting time, (3) physician waiting time, (4) clinic session duration.	Least-squares regression analysis used with simulation data to develop equations for predicting exam room requirements for given set of clinic objectives.
Thompson & Fetter [109]	Simulation	Determine effect of room configuration on occupancy, patient transfers, and bed shortages.	Model shows how hospital design (single- versus multiple-bed rooms) can affect performance.

care level while satisfying all demand for inpatient services. Geller and Yochmowitz [35] investigated the use of various bed closing rules in the reduction of maternity beds in a region. McClain [72] provided for additional flexibility in a model that derived regional obstetric bed requirements by permitting gynecological patients to be admitted to the obstetric unit.

Many of the approaches for facility size and location decisions use analytical techniques dominated by the assumption that the distance to be traveled acts as a barrier for individuals requiring health services. Numerous empirical studies, particularly those published outside the operations management literature [18] [27] [34] [73] [92] [103], have established the significance of distance and time factors on a hospital's utilization in both rural and metropolitan areas. Other factors such as service mix, size, price per day, amount of labor inputs, physician affiliation, community visibility, and patients' socioeconomic groups have been found to affect a hospital's market share [18] [27] [119]. Thus, while travel and distance attributes have been considered in facility acquisition decisions, there has been little integration with other factors having an impact on the demand for health care services and the development of location and size decision models.

Although regional planning models have considered facility costs in determining the size and location of various regional medical facilities, the concept of economies of scale has not been featured explicitly in most capacity research. Early empirical studies [9] [30] [31] [47] [66] of hospital costs showed conflicting evidence on the existence of economies of scale. A recent study by Banker, Conrad, and Strauss [6] demonstrated that economies of scale were not apparent until patients were separated according to age. Similar conclusions were reported by Hornbrook and Monheit [50], who suggested that complex medical services should be delivered through regionalization of services or some multi-institutional arrangement to benefit from economies of scale. While some authors [22] [83] [89] capture differences in the level of care offered at each facility, most size or location models do not directly consider the question of which services will be delivered at each facility, a question crucial to the economies-of-scale issue. Although a larger number of services increases market share [27], a large service mix at each location leads to increased investment and operating costs. A study of the Kaiser program [69] found that this large, multihospital HMO concentrated specific types of patient services in a small number of widely dispersed hospitals in order to reduce costly excess capacity. Recent studies [18] [73] indicated that a patient's willingness to travel a longer distance for medical care depends on the specific type of services sought.

Rather than using an integrated approach for facility size, location, and service mix decisions, researchers have decomposed the problem into the determination of the size of each service within a single facility. In particular, researchers have examined the size of medical units in both ambulatory care and inpatient hospital facilities. In the following sections, the size of service units in hospitals and outpatient clinics will be reviewed.

Size of Inpatient Care Units

The literature on the size of inpatient care units assumes a patient requires a single medical or nursing unit throughout the entire stay. Patient medical care units are grouped according to patients' diagnoses, nursing skill requirements, and available equipment. In most size models, researchers considered the interrelationship between the size of a facility and the use of policies for dealing with inadequate capacity. Patient movement between different medical units may occur when the appropriate treatment services are not available at the time of admission. In this case, a patient is assigned to an inappropriate care unit and, when circumstances permit, the patient is moved to the correct treatment area. Other responses to excess demand for services include the denial of admission to nonemergency patients (and their placement on an "on-call waiting list), the relocation or discharge of other patients to accommodate the needs of a newly arrived patient, or forfeiture of revenues through the transfer of a patient to another hospital.

Queuing theory has been used to model the behavior of hospital occupancy with respect to arrival rates and length of stay in order to determine the appropriate bed complement. Young [122], Shonick [93], and Shonick and Jackson [94] examined the improvement in utilization resulting from policies that maintained a fixed number of beds for emergency patients while nonemergency patients were placed on a waiting list when only emergency-designated beds were available. Parker [84] simultaneously sized two medical units based on the flexibility resulting from the sharing of bed capacity when one unit was fully occupied. Esogbue and Singh [29] offered an approach that explicitly considered the costs due to over- and underutilization of bed resources. A total cost function consisting of shortage and overage costs for each patient type (i.e., emergency and nonemergency) was used with a queuing model to derive the optimal unit size for each classification of illness. The costs of maintaining an empty bed were balanced against the costs of refusing admission to a patient—costs which might include lost sales, lost goodwill, and penalties for not meeting demand. Kao and Tung [58] derived the size of each patient care unit in a hospital where the number of hospital beds already had been determined. Beds were assigned to each service using a queuing model with forecasted average annual demand. A marginal analysis technique then was used to allocate the remaining capacity to specific services in an attempt to minimize the overflows occurring because of month-to-month variations in demand rate.

One of the major limitations of queuing theory approaches is the assumption of a Poisson arrival rate. In an early study, Young [122] tested the use of his queuing model against the results of a simulation model where scheduled admissions arrived at approximately the same time of day. It was found that significantly different results were obtained by the two techniques primarily because of the queuing theory assumption that arrivals were equally likely throughout the day. Swartzman [106] measured the hourly variation in daily arrival patterns and found that unscheduled arrivals could be represented by Poisson processes with time-varying arrival rates for different periods of the day. However, as discussed in [44] the assumption of a Poisson arrival rate for nonemergency patients substantially overstated the required number of beds since admissions could be controlled through the use of a scheduling system. Hancock, Magerlein, Storer, and Martin [43] studied the interrelationships between admissions scheduling, hospital occupancy, and facility size with a simulation model that assumed Poisson emergency arrival rates.

Simulation models also were used to evaluate the effects of several different size policies on the number of patients placed in higher or lower quality inpatient care units. Goldman, Knappenberger, and Eller [38] investigated the performance of different policies for allocating beds to medical care units. The simulation model assumed that nonemergency patients waited for admission when a medical unit was fully utilized and that emergency patients were admitted to an inappropriate unit during periods of inadequate capacity. Goldman et al. also investigated beds-to-rooms policies to determine the effect of allocating beds to private rooms, semi-private rooms, four-bed wards, and a combination of each. Thompson and Fetter [109] examined size issues related to the number of single- and double-patient rooms and found little difference in overall utilization levels between hospitals with all single-person rooms and ones with single-bed and multiple-bed rooms.

Researchers also have examined the balancing of capacity in situations where patients move from one medical care unit to another as their physical condition changes. For instance, patients requiring services associated with coronary care treatment may move between intensive care, coronary care, surgical care, and other nursing units. As patients progress through the hospital, they may need to be transferred to a unit already at full capacity. Analytical models have investigated the relationship between capacity levels and progressive patient care movement by representing the stochastic flow of patients through the use of Markov chains [79] [80] [101] and semi-Markov process models [117] [108] [54] [55]. The major limitation of these analytical approaches for sizing progressive patient care facilities is that the models assume infinite capacity in all medical services. However, the semi-Markov models of Kao [56] and Hershey, Weiss, and Cohen [49] include a capacity limit on the number of available beds in the most critical medical unit. Simulation studies [33] [17] [88] also have been performed to examine the impact of progressive patient care movement on size decisions.

Capacity decisions regarding the size of a specific medical service can affect the utilization of other units in an inpatient facility. Schmitz and Kwak [90] constructed a simulation model to determine the increase in operating room and recovery room capacities resulting from an additional complement of hospital beds. Thomas and Stokes [107] developed cost functions for each support service (such as laundry, laboratory, and dietary) in order to derive the size of each service based on the number of hospital beds. Thus, interactions between units of the facility also have been studied.

Size of Ambulatory Care Facilities

The size of an ambulatory care facility typically is measured by the number of examination rooms. Sumner and Hsieh [104] determined the capacity of an orthopedic outpatient clinic by simulating the effects of different numbers of examination rooms on the clinic's performance on several conflicting objectives. Using the date generated from a simulation model, regression equations were developed to predict the room requirements based on management goals regarding total clinic hours, idle examination room time, average patient waiting time, and average physician idle time. In other ambulatory care work, Gitlow [37] selected the number of procedure and counseling rooms that maximized profits for an abortion clinic. Other authors [91] [113] examined the interaction between work-force staffing and floor-space requirements in an ambulatory care facility.

FACILITY ALLOCATION DECISIONS

The previous discussion on facility acquisition decisions demonstrates the strong relationship between the size of health care facilities and the policies regarding scheduling of facility resources. However, the research presented in this section examines the areas of hospital admission and surgical scheduling under the assumption that the size of each medical service has already been determined. We call these scheduling topics facility allocation decisions since the literature assumes that the resources which typically constrain admission and surgical procedures are beds and operating rooms. Milsum, Turban, and Vertinsky [75] reviewed research on inpatient admission scheduling up to 1973; surgical-suite scheduling research up to 1980 was reviewed by Magerlein and Martin [71]. Our subsequent discussion of inpatient facilities will complement these two reviews. Ambulatory care services are reviewed as a separate topic area in this section. Table 3 lists major research efforts on facility allocation problems.

Inpatient Admission Scheduling

The complexity of the admission scheduling problem arises from the high degree of physical patient contact inherent in the delivery of health care services. A high contact environment [14] creates more uncertainty in daily operations due to variability in arrival times and customer requirements. Admission scheduling models attempt to maximize bed occupancy levels subject to three sources of variability: emergency admissions, patient length of stay, and patient service-mix requirements.

While hospitals desire high bed occupancy, slack capacity must be maintained for emergency admissions. As discussed earlier, a buffer for the uncertain arrival of emergency patients might have been included in the size of an individual medical unit. In other situations, admissions models accommodate emergency-bed requests in the implementation of the scheduling system. Models of this type vary in their assumptions and methodologies. Kolesar [62] used a Markov decision model to allow the number of elective admissions to vary each day according to the occupancy level and an estimate of the number of future emergency patients. Unlike Kolesar, Barber [7] did not assume that an infinite queue of nonemergency patients was available and instead captured the stochastic nature of elective admission requests. Swain, Kilpatrick, and Marsh [105] scheduled elective admissions with a model that predicted occupancy levels while allowing for modification of the length-of-stay and unscheduled-arrival-rate distributions. Hancock, Warner, Heda, and Fuchs [45] used a simulation technique to evaluate an allowance policy that reserved a fixed number of beds for emergency patients. A rotational schedule of emergency admissions was proposed by Handyside and Morris [46] where hospitals or different medical units admitted emergency patients only on specified days in a cycle.

Length-of-stay (LOS) estimates for all patient types are needed in order to schedule admission dates for elective patients. A patient's LOS varies according to severity of illness and the admitting physician's treatment preference. Authors have estimated LOS using empirical data [46] [62] [88] and with the gamma [10], Pascal [10], geometric [105], and normal [7] probability distributions. Efforts in this area are justified by the findings of Robinson, Wing, and Davis [88] that compared the performance of several scheduling approaches under various levels of LOS accuracy using a simulation model. Robinson et al. found that better LOS estimates offered significant advantages over more complex scheduling models with poorer LOS data accuracy.

Patient service mix constitutes another source of variability in the process of scheduling patient admissions. A patient requires a bed and the use of other resources such as nursing care, surgical rooms, and support services; the demand for these resources is dependent on patient service mix. The models described above all focus on maximizing the utilization of bed resources, which can lead to extreme

Table 3: Literature on facility allocation decisions.

	Methodology	Managerial Objective	Comments
Barber [7]	Chance-constrained mathematical programming	Determine maximum number of admissions that can be booked each day to maximize average occupancy over the planning horizon.	Methodology can be used for static and dynamic admission scheduling and for long-term occupancy maximization.
Barnoon & Wolfe [8]	Simulation	Determine effect of various numbers of operating rooms, anesthetists, and nurses on the utilization of OR facilities and staff.	Model used to determine performance of scheduling techniques under various facility configurations.
Charnetski [13]	Simulation	Determine method for surgical procedure time estimation which balances idle time of surgeons and of OR facilities/staff.	Establishes procedure time estimates based on the ratio of surgeon waiting time costs to OR staff idle time costs.
Connors [20]	Heuristic	Determine admission date for each patient to minimize the costs of deviations from the patient's desired date and the desired hospital occupancy level.	Model jointly considers patient preferences and hospital operating performance.
Kolesar [62]	Markov chains and linear programming	Determine number of admissions to schedule to maximize average occupancy with an overflow constraint or minimize overflows with constraints on utilization.	The Markovian decision model provides flexibility in service and arrival distributions; LP model results in an optimal or "good" admissions decision rule.
Kwak, Kuzdrall, & Schmitz [64]	Simulation	Determine which of five rules for dispatching surgical patients to ORs maximizes facility utilization.	Simulation used to test various dispatching rules as a simple way to develop daily surgery schedules.
Offensend [82]	Stochastic process	Determine admission policy, either census-based or work-load-based control policy, which minimizes nursing costs.	First study to show that work-load-based admissions control results in smoother work loads and lower nurse staffing costs.
Robinson, Wing, & Davis [88]	Simulation	Determine scheduling method which minimizes total cost of empty beds, patient overflows, and admission refusals.	Simulation showed accuracy of length of stay estimates to be more important than sophisticated scheduling rules.
Swain, Kilpatrick, & Marsh [105]	Heuristic	Determine number of additional elective admissions for each day that will maximize expected occupancy over the planning horizon subject to an overflow probability constraint.	This stochastic model predicts census, thus allowing census control through daily admission decision rules.

variations in the utilization of other resources. Offensend [82] was the first author to develop a hospital admission system based on nursing work-load requirements. In his research, a simulation model was used to compare occupancy-related admission policies with systems that considered the relationship between admissions and nursing requirements. Offensend found that work-load variance was reduced when patients were scheduled by a work-load-based model. In a similar study, Shukla [95] presented a scheduling system based on nursing work load and concluded that patient service mix had to be considered in the scheduling process in order to minimize changes in nursing staff assignment patterns.

Surgical-Suite Scheduling

Hospital resources also can be more effectively managed by considering the interaction between schedules admitting elective patients and those allocating surgical-suite facilities. One set of admission scheduling models [7] [20] [88] used a daily quota on the number of surgical patients admitted per day without explicitly considering available operating room capacity. Magerlein and Martin's [71] review of surgical-suite scheduling discussed procedures for scheduling patients in advance of their surgical dates and techniques for the assignment of patients to operating rooms at specific times of day. The literature on advance scheduling differentiates between methods which use a first-come, first-served decision rule and those based on the practice of blocked booking to reserve operating room (OR) time for individual surgeons or surgical specialties [65] [76] [77] [123]. Procedures for assigning patients to specific ORS at particular times are discussed by Barnoon and Wolfe [8], Charnetski [13], Esogbue [28], Goldman, Knappenberger, and Moore [39], and Kwak, Kuzdrall, and Schmitz [64]. Other work in the area of surgical-suite problems suggests using requirements planning systems for the management of surgical inventories [102].

Ambulatory Care Scheduling

While hospitals must contend with the uncertain arrivals of emergency patients, a similar problem exists in outpatient clinics where patients arrive without appointments. Outpatient clinics provide health care managers with a unique set of environmental conditions that must be considered in scheduling appointments including high no-show rates, early and late arrivals of patients, confusion and crowding in waiting areas with limited space, late arrivals of physicians, and inadequate equipment for the treatment of patients. Fetter and Thompson [32] studied the complexities of an outpatient clinic using a simulation model to investigate the performance of different levels of load factors (the percent of available appointments filled), the time between appointments, and office hours on physician and patient waiting times. In a simulation model of a university health service outpatient clinic, Rising, Baron, and Averill [87] evaluated a system which scheduled more appointments during periods of low walk-in demand in order to balance the daily demand for services. Using an empirical approach, Johnson and Rosenfeld [53] studied factors affecting waiting times in eight ambulatory care facilities and found significant differences in appointment systems and corresponding patient waiting times.

WORK-FORCE ACQUISITION DECISIONS

Work-force capacity is a function of the number of personnel hours available per unit of time and the composition of the work force in terms of the mix of employee skills. Most health care organizations determine the number of full- and part-time employees of various skill levels through the annual budgeting process. A summary of the literature on determining the work-force size in health care organizations is presented in Table 4. As shown, work-force acquisition decisions must consider such factors as (1) the stochastic nature of demand, (2) the difficulties in measuring the productivity of health care providers, (3) the flexibility facilitated by the substitution of different employee types, (4) the use of part-time employees for lowering operating costs and improving schedule flexibility, and (5) the use of overtime and temporary employees to provide additional work-force capacity.

Hospital Staffing

The majority of work-force size literature has addressed hospital nursing staffs. This emphasis is understandable; the nursing work force constitutes the single largest operating cost in most hospitals. In its simplest form, the inpatient work-force size problem is to determine the number of full-time equivalent nurses necessary to meet projected nursing hour requirements without allowing for substitution of tasks between nursing skill levels. Lowerre [68] provided an easy-to-use, general purpose staffing algorithm to derive the number of full-time equivalent employees needed to meet staff coverage standards when regular rest policies (such as two days off per week) and holiday and sick days are accounted for. Trivedi [111] described a goal programming model that derived staffing levels for both full- and part-time employees. The model also accounted for substitution between nursing skill levels, rest policies, and days off. The amount of substitution between registered nurses (RNs), licensed practical nurses (LPNs), and nurse's aides (NAs) in Trivedi's model was limited to a specific percentage of total demand where staffing coefficients specified the extent to which one nursing skill level could substitute for another.

Many authors have considered the use of "float nurses" to reduce the size of the full-time work force. Wolfe and Young [120] first suggested the concept of "controlled variable staffing" where nursing units were staffed at minimum levels and additional workers were temporarily assigned to units experiencing periods of high demand. Kao and Tung [59] described a linear programming model that derived the size of a float-nurse pool subject to the size of the permanent staff, the amount of overtime hours, and the requirements for temporary agency employees. Hershey, Abernathy, and Baloff [48] used a simulation model to illustrate that a variable staffing policy could reduce the size of the work force when nurses were placed in high demand units as needed instead of being permanently assigned to a single medical service.

Several authors have described more comprehensive approaches to the nurse staffing problem. A somewhat dated but detailed discussion of planning, scheduling, and controlling the nursing work force is presented in Abernathy, Baloff, and Hershey [1]. Based on this work, the authors later developed an integrated model for staffing and scheduling nursing units using stochastic programming with chance constraints for demand uncertainty and service level objectives [2]. The model evaluated fixed and variable staffing policies by deriving the size of each nursing unit, the number of employees in the float pool, and any personnel actions such as transfers, hires, and discharges. Perhaps the most complete investigation of nurse staffing is the recent work of Kao and Queyranne [57]. In this study, the authors used eight mathematical models to investigate the impact of demand uncertainty, the aggregation of work-force skill types into one equivalent skill level, and the

Table 4: Literature on work-force acquisition decisions.

	Methodology	Managerial Objective	Comments
Abernathy, Baloff, Hershey, & Wandel [2]	Chance-constrained programming model	Determine appropriate staff level so that probability of an overflow is less than established risk level for a given service policy.	Formulates control problem as stochastic programming model solved by an iterative technique with a random loss function.
Carlson, Hershey, & Kropp [11]	Recursive mathematical programming simulation	Determine number of physicians and PEs, task assignments, services to be offered, and facility size to minimize total annual costs subject to waiting-time restrictions.	Methodology allows optimal aggregate staff planning, while simulation disaggregates the aggregate plan to quantify the impact of stochastic demand on waiting times. Used to select the optimal PE skill level. Model can be used for staff sizing and task allocation.
Connell, Adam, & Moore [19]	Heuristic	Determine number of employees required per period to minimize total food-service production costs.	Demonstrates the benefits of well-known aggregate planning models in a service setting.
Gupta, Zoreda, & Kramer [42]	Queuing theory	Determine number of hospital messengers needed for a given service level.	Shows the relationship between staff size and call waiting time, a measure of service.
Hershey, Abernathy, & Baloff [48]	Simulation	Determine savings resulting from variable staffing (compared to fixed staffing) using three alternative criteria for setting the float staff size.	Demonstrates the economic benefit of using float nurses. Proposes alternative methods of implementing variable staffing.
Ittig [52]	Linear programming	Determine number of physicians of each type, patient group–physician assignments, and population acceptance rate to maximize services provided subject to operating budget and physician supply.	Excellent physician planning model for HMOs. Objective function uses weights to specify preferences for relative comprehensiveness of primary and specialist physician services; weights can be adjusted to generate physician staffing alternatives.
Kao & Queyranne [57]	Various methodologies (listed below)	Determine regular work force, overtime, and agency personnel required from each skill class to minimize total nursing costs.	Models represented by triplet: S=single period, M=multiple period; A=aggregate, D=disaggregate; D=deterministic demand, P=probabilistic demand.
MDP model	Two-stage stochastic programming with recourse		Very complete formulation of staff size problem but extremely difficult to solve.
MAP model	One variable convex minimization		Good approximation to MDP model.
SDP model	Stochastic programming		Extremely difficult to solve.
SAP model	Algorithm		Solved as a newsboy-type problem.
MDD model	Linear programming		Classic formulation with consideration of time-varying demand.
SDD model	Linear programming		Classic formulation of nurse work-force size problem.
MAD model	Piecewise linear programming approximation		Good approximation to MDD model.
SAD model	Decision rule		Simple dissaggregation of SAD solution gives same result as SDD linear programming model.
Keller & Laughman [61]	Queuing theory	Determine number of physicians to minimize total costs of patient waiting time and physician capacity.	Normative use of queuing theory to find optimal staff size.
Lowerre [68]	Heuristic	Determine minimum nursing staff size subject to nurse coverage needs and days-off requirements.	Easy-to-use algorithm applicable to a broad range of continuous staffing situations.

Schneider & Kilpatrick [91] Overall planning model	0-1 integer programming	Determine number of physicians and PEs of each type, task assignments, services provided through extraordinary means, and services not provided by the HMO which minimize operating costs subject to capital constraints, number of subscribers, and HMO capitation rate.	HMO planning model specifies the physician-PE team composition, service mix, and exam room requirements. Model can determine staff size and task allocation. Appropriate for HMO start-ups.
Subscriber maximization	0-1 integer programming	Determine number of PEs of each type, their task, assignments, services to be provided by extraordinary means, and initial capital required to maximize the number of subscribers subject to initial physician supply.	Model appropriate for existing physician practices. Model can determine staff size and task allocation.
Shuman, Young, & Naddor [97]	Linear programming	Determine number of physicians and PEs to hire, their assignments to tasks and sites, the level of technology, and the number of graduates of each type of PE to minimize the sum of patients' fees, clinic construction costs, community contributions, education costs, and care shortage costs.	Multisite planning model establishes the mix of PEs and their assignments to services and to clinic locations. Considers cost from the perspective of the community. Model can determine staff size and task allocation.
Smith, Over, Hansen, Golladay, & Davenport [98]	0-1 integer programming	Determine number of physicians and PEs of each type to minimize total wage costs subject to labor substitution possibilities and weekly patient demand.	Provides the optimal physician-PE team composition for a specified weekly demand. Employs "homogeneous encounter groupings" to classify service outputs and specifies alternative labor inputs for each output. Model can determine staff size and task allocation.
Trivedi [11]	Goal programming	Determine number of full- and part-time RNs, LPNs, and aides, subject to 5 weighted goals: (1) maximize budget surplus, (2) maximize number of full-timers, (3) minimize budget deficit, (4) minimize number of part-timers, and (5) minimize number of nursing hour shortages.	Very complete staffing model allows for multiple solutions by adjusting goal weights. Accounts for substitution among nursing skill levels, days-off provisions, and part-time, full-time, and overtime nurses.
Willemain & Moore [118]	Linear programming	Determine rate of visits to physicians and PEs to maximize total patient visits subject to labor capacity constraints.	Provides for optimal physician capacity utilization through physician-PE practice teams.
Wolfe & Young [120] [121]	Linear programming	Determine number of nurses of each skill level to meet service demands at minimum total nurse staff cost.	Specifies size and composition of the nursing staff through assignment of task complexes to nursing skill levels using a multiple assignment model with human judgment.

effects of time-varying demand on the work-force budgeting problem. Using data from a public hospital in Arizona, the findings supported the use of nursing staff models incorporating the stochastic nature of demand and aggregating nursing requirements across skill levels.

The staffing of inpatient support services has been investigated by a limited number of authors. Gupta, Zoreda, and Kramer [42] used queuing theory to determine the size of a messenger unit whose function was to transport patients and other objects in response to requests from all units in a hospital. Connell, Adam, and Moore [19] applied aggregate planning methodologies to a hospital food service to obtain production rates and overtime levels.

Ambulatory Care Staffing

The presence of alternative delivery patterns complicates the process of setting work-force sizes in ambulatory-care facilities. These alternative patterns are delineated by differences in the mix of health care personnel, the tasks allocated to each personnel class, and the technologies employed in delivering health care services. While different types of physician extenders (PEs) such as nurse practitioners and physician assistants are found in medical group practices, surgical centers, and emergency clinics, there are few commonly accepted standards regarding work-force composition and size decisions with respect to these resources. The literature on ambulatory care facilities has considered three types of substitution of tasks among health care providers. These include (1) vertical and (2) horizontal interchangeability between personnel types and (3) the substitutability of labor and technology in the delivery of health care services.

Vertical substitution is the dominant delivery alternative in the literature on work-force size. Several methodologies have been used to determine the optimal work-force size of each personnel group, given the opportunity to increase a physician's productivity by allowing PEs to perform a limited set of tasks previously assigned only to physicians. The work of Shuman, Young, and Naddor [97] determined the annual number of physicians and assistants required for a region of ambulatory care clinics using a linear programming model with constraints specifying the types of services assistants can perform and the maximum number of assistants each physician can supervise. Willemain and Moore [118] also presented a linear programming model to evaluate capacity in patient visits per day for a physician with a specified number of PEs. Schneider and Kilpatrick [91] developed two mixed integer programming models to investigate the formation of health care teams consisting of a combination of physicians, PEs, and/or nurses. The first model allowed liberal use of PEs and determined the work-force mix for an HMO that minimized the operating cost of providing comprehensive health services for a defined population. An alternate model was formulated to determine the size and composition of the supporting personnel that maximized the population that could be served by an existing staff of physicians. Both models also included the possibility of subcontracting for medical services not provided by the HMO. In a later study, Carlson, Hershey, and Kropp [11] used the first Schneider and Kilpatrick model with a recursive optimization-simulation approach to determine the work-force composition and size in the context of the stochastic nature of patient arrivals, task times, and patient needs for service. Smith, Over, Hansen, Golladay, and Davenport [98] studied the economics of a variety of health care teams where the effects of volume and patient mix on the average cost of delivering a specific service were considered in an integer programming model.

Horizontal substitution refers to a situation in which equally skilled practitioners can each perform a subset of the others' tasks. In his model of ambulatory health service, Ittig [52] illustrated the use of horizontal substitution between primary care specialists in pediatrics, internal medicine, and obstetrics/gynecology where each of these practitioners could treat a fraction of the others' patients. Using a linear programming model, Ittig also demonstrated vertical substitution with family practice physicians who could treat patients of all other primary care specialists. The linear programming model maximized the services available to a patient population such as those found in an HMO, subject to financial and physician availability constraints.

A generally accepted principle is that when technology is substituted for personnel a higher level of technology will increase the productivity of personnel and thereby decrease the size of the work force. Shuman et al. [97] determined the composition of the work force in a regional planning system in conjunction with various technology options. They assumed a monotonically increasing relationship between higher levels of technology and the number of patient visits delivered by an ambulatory care provider. Using a broader definition of technology, Ittig [52] evaluated the substitution effect of ambulatory care for more costly hospitalization services. He used a linear function to describe the relationship between hospital bed-days saved through the use of ambulatory care physician visits. Experimentation with this model in an HMO organization indicated that increased flexibility in terms of substitution of ambulatory services for hospital care could change the composition of the work force to provide more comprehensive ambulatory services.

Other methodologies for determining the work-force size in ambulatory care facilities do not consider substitution between personnel types. Keller and Laughman [61] used a queuing-theory approach to determine the number of physicians needed for the morning and afternoon shifts at a clinic. Rather than using actual cost parameters, they obtained the optimal number of physicians by specifying the ratio of the cost of physician idle time to the cost of patient waiting time. Vargas, Hottenstein, and Aggarwal [113] developed utility functions using a common unit of output measurement to express an HMO's multiple performance attributes, such as average patient waiting time and staff idleness. A search algorithm was used to maximize total system utility and determine the mix and quantity of labor and nonlabor inputs subject to budget, technology, and facility capital constraints.

WORK-FORCE ALLOCATION DECISIONS

As in the work-force acquisition literature, the vast majority of research on work-force scheduling deals with inpatient nursing resources. This is demonstrated in Table 5, which summarizes past research on work-force allocation decisions. The nurse scheduling problem in the most general form is to allocate a budgeted number of nurses to a specific medical unit based on the forecasted occupancy level which then is used to determine the type of nursing skills required for patients typically admitted to the unit. Kaplan [60] described a feedback control procedure that compared the scheduled number of nursing hours with the actual number to determine whether the judgments of the administrator making scheduling decisions succeeded in using nursing personnel efficiently. In a very complete formulation of the hospital-wide scheduling problem, Warner and Prawda [116] allowed limited substitution of tasks between nursing skill levels in a mixed integer quadratic programming model that maximized the quality of care by minimizing nursing care

Table 5: Literature on work-force allocation decisions.

	Methodology	Managerial Objective	Comments
Arthur & Ravindran [5]	Goal programming	Determine set of individual nurse day-shift assignments which minimizes deviations from four goals: (1) minimum staff level, (2) desired staff level, (3) schedule preferences, and (4) special requests.	Schedule preferences quantified by nurse ratings of certain schedule attributes. Assumes a single skill level. Model assigns nurses to days, manual assignment to shifts.
Cavaiola & Young [12]	0-1 integer programming	Determine task assignments for each skill level which maximize total appropriateness of care score for all assignments.	The Collaborative Patient Assessment Instrument is used to classify patients by required nursing activities. Patients then are clustered by required nursing activities for input into the staff assignment model.
Liebman, Young, & Bellmore [67]	0-1 integer programming	Determine task assignments for each skill level which maximizes total perceived effectiveness of all nursing task assignments.	Long-term care setting. Model can be used for task assignment or to evaluate effectiveness of alternate staffing configurations.
Lyons & Young [70]	0-1 integer programming	Determine task assignments for each skill level which maximize total appropriateness of care score for all assignments.	Residential mental health care setting.
Miller, Pierskalla, & Rath [74]	Cyclic coordinate descent algorithm	Determine set of individual nurse day-shift assignments which minimizes the costs of not staffing at desired levels and the costs of not meeting nurse work schedule preferences.	Schedule preference costs determined by individual schedule aversion weights and the extent to which an individual's past schedules have met desires. Assumes a single skill level.
Trivedi & Warner [112]	0-1 integer programming	Determine allocation of available float nurses to units which minimizes total severity of need for nursing personnel and maintains equity of need between units at the start of each shift.	Multivariate regression model developed to quantify how severely a unit needs additional nurses given staff level and unit census.
Warner [115]	Multiple choice programming	Determine set of individual nurse day-shift assignments which maximizes total schedule preference weights of all nurses subject to minimum nursing coverage constraints.	Nurses score certain attributes of schedules according to their aversion to working them and head nurses score special requests to derive schedule preference weights. Incorporates substitution by defining coverage requirements as minimum RN, minimum RN+LPN, and minimum RN+LPN+aide needed each shift.
Warner & Prawda [116]	Mixed integer quadratic programming	Determine number of RNs, LPNs, and aides to schedule in each unit during each shift which minimizes the cost of nursing care shortages subject to a budgeted number of nurses available in each skill class.	Considers (1) cost of nursing care shortage as a convex decreasing function, where cost equals zero at the desired staff level, and (2) limited substitution among skill levels.

shortage costs. One of the unique features of this model was the inclusion of a shortage cost for a particular skill class in a specific ward when the schedule could not accommodate the nursing-hour requirements across the planning horizon. A nonlinear cost function for nursing shortages was used to reflect that hospitals may tolerate small shortages, but the cost of large shortages increases exponentially with the magnitude of the shortage. Trivedi and Warner [112] offered a methodology for allocation float nurses to medical units at the start of each shift based on predicted severity of need. The head nurse's judgments were used to develop regression equations for predicting the severity of need for various nursing skill levels from control variables, such as the number and classification of patients and the size of the medical unit's current nursing staff. The severity-of-need predictions were used in conjunction with a branch-and-bound procedure to evaluate the set of possible float nurse allocations that minimized the total severity index of the hospital and maintained staffing equity between medical units.

In response to the desire for the equitable allocation of high-quality schedules among employees, many nurse scheduling models incorporated the individual preferences of each nurse for a particular shift assignment and for days off. Warner [115] implemented an integer programming model that minimized the sum of penalty costs—including the individual's preferences for certain schedules—and penalties for deviations from the desired staffing coverage. Miller, Pierskalla, and Rath [74] also used mathematical programming, but they considered hospital staffing policies and employee scheduling constraints as two classes of constraints. Feasible nurse schedules were included in a set of binding constraints while desired hospital staffing policies were treated as nonbinding constraints whose violation incurred a penalty cost in the objective function. A goal programming approach was offered by Arthur and Ravindran [5] as a methodology for incorporating both individual preferences and hospital objectives.

Several work-force allocation models were concerned with the assignment of specific daily work tasks to employees of different skill classes. Wolfe and Young [121] also used their linear programming model to assign nurses of six personnel classes to sixteen different tasks. Mixed integer programming models were used for task assignment in long-term care organizations [12] [67] and mental health facilities [70]. Q-sort techniques (see [81]) were used by Liebman, Young, and Bellmore [67] and Lyons and Young [70] to quantify subjective judgments about the appropriateness of various personnel assignments and to evaluate them based on the effective utilization of personnel. Cavaiola and Young [12] presented a multiple regression-based technique for classifying long-term care patients according to their required nursing care activities.

FUTURE RESEARCH DIRECTIONS

The remaining sections of this paper are devoted to a discussion of those topics in health care capacity management that we believe offer the highest potential return given the changes currently taking place in the industry. Despite the extensive topic coverage in the existing literature, changes in the health care environment have significantly reduced the usefulness of the past research for several reasons. First, researchers have shown the most interest in problems that exist in inpatient facilities, but the present trend is toward increased use of ambulatory care facilities as a substitute for hospitalization. Second, capacity management has been studied under the assumption that health care managers primarily were concerned with the objectives of (1) minimizing operating and investment costs, (2) maximizing facility and workforce utilization, (3) minimizing the distance traveled by the patient, (4) minimizing patient waiting time, and (5) minimizing the number of patients receiving inappropriate care. As will be shown below, managers and researchers now must consider an expanded range of objectives. Third, demand typically has been represented as a function of the demographic characteristics of the population in the geographic region served by the health care facility. However, a health organization's market share is constrained not only by distance or travel boundaries but by factors such as membership in an integrated health care system and the competitor's location and service mix. Finally, there is little integration in the literature between work force and facilities resources, which are interconnected and interdependent components of a health care delivery system.

A major challenge for researchers is to incorporate organizational goals in capacity decisions in order to evaluate alternative strategic plans and their impact on performance. This requires the development of measures of customer service, resource flexibility, cost, and market share, and the quantification of their interrelationships with capacity decisions. One of the most significant changes in the health care industry is the increasing number of patients who purchase health care services through a single health plan rather than through piecemeal fee-for-service purchases. To remain competitive in this environment, health care organizations must engage in strategic planning to determine the organization's allocation of resources and desired marketplace position. Therefore, during the strategic planning process, health care providers must decide what services they will provide for specific markets. This requires that providers shift from a process orientation towards product-line management, thus creating an emphasis on the profitability of the service mix.

Future work in capacity management also must consider the increased interest of the health care community in quality-of-care issues. Consumers and some health care providers are concerned that quality of care will significantly decrease as a result of new payment mechanisms that provide a financial incentive to decrease the amount of care provided. This concern is based on the assumption that more medical services lead to higher quality care and that under financial constraints health care providers will use less costly, and possibly lower quality, services and/or limit the quantity of care. Donabedian, Wheeler, and Wyszewianski [23] offered a model that describes the relationship between cost and quality. According to Donabedian et al., as costs increase so do health benefits until the point is reached at which additional expenditures do not lead to an increase in health status. Beyond this quality-neutral region, extra medical services become harmful to the health of the patient. Empirical studies have shown that high quality medical services can be delivered at a lower cost in specific care settings. Various researchers have reported that specialized care units such as cardiac surgery improve quality and significantly lower costs when they are able to increase their volume of patients [26]. Extensive research also indicates that quality can be maintained when performing services at a lower cost using less expensive providers such as physicians' assistants, nurse practitioners, and midwives as substitutes for more expensive physician care [15] [41]. Furthermore, it has been found that nonphysician health care providers are rated higher on quality measures of the interpersonal relationship between the health care provider and the patient [86]. In research attempting to measure the relationships between cost, quantity, and quality of health care, researchers have used unique definitions of the quality of the care provided, primarily because there is no universally accepted definition of health care quality. Thus, the research on capacity management will need to draw on the continually developing definitions and measures of quality of care as they are developed and refined.

While there are opportunities for capacity management research in a variety of disciplines, it is our objective in the following sections to concentrate on research agendas for the operations management community. Our intent is not to present an exhaustive set of research questions; rather, the purpose of this discussion is to motivate future research on this topic. The research areas developed in this paper are based on capacity decisions that have been identified by health care managers as important managerial problems for the next decade.

Vertical Integration

Vertical integration is an attempt to control the sources of demand and/or to enhance control of the inputs through ownership of supply and distribution channels. A recent trend is the use of forward integration by health care organizations to reduce the uncertainty of demand volume. Hospitals are expanding into the field of ambulatory care to provide a source of inpatients and to increase the utilization of existing facilities and work-force resources. Health care organizations also are developing insurance and HMO plans that guarantee a patient population for a given period (usually one year). It has been projected that vertical integration in the next decade will lead to a situation where health care services will be controlled by 20 to 40 companies called the "SuperMeds" [25].

The desired result of vertical integration is a balanced health care organization that matches demand with most appropriate source of care within the vertically integrated system. Past research has been limited to the work of Dokmeci [22] and Parker and Srinivasan [83] who examined problems related to the design of these systems but did not fully capture the trade-offs inherent in vertical integration. The primary focus of research in this area should be to determine the system configuration that will assure patients of high-quality, cost-effective health care services. Research questions to be addressed include the following:

1. What degree of vertical integration is most appropriate for a health care organization's operations strategy and competitive position? In particular, how many distinct levels of care or different types of facilities are needed, and how many facilities of each type should be offered?
2. Should the vertically integrated system be formed by contractual affiliations with other health care organizations or through direct ownership?
3. What services should be delivered at each level in the system? Specifically, how much duplication of services at different levels of care should be offered to satisfy consumer preferences, utilize resources effectively, and guarantee a high-quality service?
4. How much capacity should be available at each care level? (The progressive patient care literature provides a basis for research on balancing capacity between different levels of care in a fully integrated delivery system.)
5. What is the relationship between system configuration and future market share?
6. What are the advantages and disadvantages of sharing resources between different levels of care in terms of both acquisition and allocation decisions? For example, is the cost of cross-training a nurse practitioner in both inpatient and outpatient services offset by the benefits of increased staffing flexibility and greater opportunities for professional growth?

7. Should we study capacity decisions in these systems with an integrated model or should we use a disaggregation approach similar to the materials requirements planning framework?
8. Is a centralized approach to the management of capacity resources a more effective strategy than decentralization?

Multihospital Systems

The response of many independent hospitals to an increase in competition has been the formation of multi-institutional hospital systems [40]. Several factors—including increasing capital requirements, new technology, and constrained capital markets—have forced independent hospitals to join systems to improve their ability to compete for financial resources [124]. Multihospital systems also present the potential for significant operational savings and downsizing opportunities through the regionalization of services [69]. Past research on regional facility acquisition decisions indicates the need for an integrative model for facility size, location, and service mix across a multiple-site system. Such a model should incorporate the following factors:

1. Economies of scale within specific service types.
2. Facilitywide economies of scale.
3. The potential for improved quality when services are delivered at a limited number of facilities, thus increasing patient volume at each delivery site.
4. The impact of facility decisions on market share potential.

Multihospital systems must account for increased opportunities to share common facility and work-force resources. Research issues to be investigated include the following:

1. How does the optimal size of an inpatient care unit change when a hospital is part of a system and has the ability both to transfer patients to a nearby affiliated hospital and to receive patients from other hospitals?
2. What are the advantages and disadvantages of centralized admission scheduling for a multihospital system? To what extent will a centralized admission scheduling approach minimize fluctuations in daily occupancy and resource requirements at each hospital?
3. Should work-force scheduling models schedule workers across facilities, thus allowing the possibility of assigning an employee to more than one facility? Can individual preferences for specific facilities be incorporated into allocation models?
4. What are the benefits and costs of a centrally managed pool of work-force resources that float across affiliated hospitals?
5. How does the optimal float-pool size change when models are expanded to include interhospital transfers and the associated costs of cross-training, productivity losses, wage differentials, and scheduling and labor control systems?
6. Should the interhospital float pool be limited to specific categories of work-force resources?
7. How can Trivedi and Warner's [112] model be used for allocating a centralized float pool in a multihospital system?

Hospital Downsizing

Downsizing has been identified as a key strategic alternative for survival in the turbulent hospital industry [16] [108] [114]. The process of downsizing involves reducing the number of hospital beds and related components in the delivery system. However, there is a concern that many health care providers are overlooking the importance of reducing capacity in order to minimize cost. According to Egdahl, "Most of these cost-management approaches will fail in their long-term goal of overall communitywide (cost) containment because they do not decrease the capacity of the health care system, especially the number of hospital beds" [24, p. 126].

Research on capacity reduction in inpatient services must simultaneously consider facility and work-force resources. This requires that demand be represented in terms of products rather than only in terms of requirements for a bed and associated nursing services within a specific medical unit. A product-line approach defining demand as a package of services consumed by a patient is necessary for obtaining a well-balanced inpatient delivery system. Other factors that should be included in inpatient capacity sizing models are

1. Revenue and delivery costs for each product
2. Marginal cost of eliminating or reallocating a capacity resource
3. Decreased aggregate demand uncertainty as HMOs and other insurance plans limit the patient's and physician's choice of hospitals
4. An increase in the proportion of demand represented by emergency patients as ambulatory care is substituted for hospitalization
5. Reduction in the variability of patient length of stay as new reimbursement methods limit the variation in physician practice patterns
6. Restrictions on sizing and product mix decisions imposed by a hospital's competitive strategy; for example, a hospital may choose to deliver a full line of services at the lowest cost possible or it may limit the number of services offered based on the distinctive competence that gave the hospital its strong market position.

Subcontracting Services

The subcontracting of health care services is a recent innovation that involves contractual agreements between health care provider organizations and HMOs, PPOs, and other insurance plans. For example, an HMO and a hospital might enter into a contractual affiliation which specifies the amount and types of reimbursement to be provided when HMO subscribers use inpatient services [63]. In general, the motivation for this type of affiliation is (1) to lower costs for the purchaser in exchange for a referral system for the supplier, (2) to offer a more comprehensive range of services to HMO subscribers, and (3) to gain a marketing advantage through affiliation with a well-established provider. Although the nature of subcontracting relationships has been described in the literature [78], the effects of subcontracting on capacity management have not been investigated.

Research questions that need to be addressed from the perspective of the purchaser include the following:

1. How much capacity should be purchased from supplier organizations and for what levels of care, given the characteristics of the membership population?
2. Should contractual agreements be made for aggregate capacity, such as patient days, or for specific types of services, such as intensive care patient days?
3. What are the advantages of guaranteeing a minimum demand level to the supplier? How much of a price discount should the purchaser require if it is to assume some of the risk concerning actual demand rates?

4. What factors should be considered in the selection of a supplier? In particular, how will the quality of services be measured? Should quality levels be guaranteed in the contract?
5. Should the purchaser use single or multiple sources of services? What are the costs and benefits of offering patients and physicians flexibility in their choice of providers?

There also is a need for research that considers the supplier's side of subcontracting relationships:

1. What are the benefits of these agreements to the supplier?
2. How much capacity should be allocated to each purchaser group? Should capacity be reserved for fee-for-service patients?
3. How large a price discount should be given to guarantee a minimum level of demand? Should the price discount also consider the potential for increased market share?
4. How much risk, if any, should the supplier assume if the contracted demand level is not realized? That is, should the supplier demand a higher price if it accepts some of the risk for actual patient volume?
5. If contractual agreements lead to increased future market share, how will this impact current and future facilities and work-force acquisitions?
6. If a supplier provides services to multiple purchasers, how would priorities be established and maintained between patients in each group? For example, would the supplier distinguish between customers from different organizations in the admissions and surgical scheduling processes?

Freestanding Ambulatory Care Clinics

There has been tremendous growth in freestanding ambulatory care facilities, such as urgent care clinics and surgicenters, that offer a limited range of services at a low cost and at a convenient location for the patient. These facilities are open for more hours than the typical physician's office; as a result, many consumers are using freestanding clinics for both urgent and primary care needs. Prepaid health organizations are developing agreements with freestanding facilities in order to offer after-hours services to their members and thus eliminate unnecessary visits to hospital emergency rooms. While past research has examined acquisition decisions for ambulatory care facilities, these studies have focused on minimizing the distance traveled, minimizing the costs of delivering health care services, and holding physician and patient waiting times to reasonable limits. Since freestanding centers are in competition with hospital-based emergency rooms and physician's offices, future research efforts should concentrate on location and size decisions in a competitive marketplace with a profit-based objective function. Retail location models [36] [51] incorporate the impact of a variety of factors on demand and, as a result, offer an insight for studying freestanding ambulatory clinics. The following features should be considered in future research:

1. The location, size, and service mix of both competitor's facilities and affiliated clinics
2. The interaction between floor-space requirements and work-force acquisition decisions as described in [91] and [113]
3. Patient travel time and distance
4. Comparison of the cost of support services—such as radiology, laboratories, and pharmacies—offered within a facility to the cost of subcontracting support services at another site

5. Distance to the nearest inpatient care facility for cases that cannot be handled in a freestanding clinic
6. The role of freestanding clinics in generating demand for other facilities in a vertically integrated health care delivery system.

HMOs

HMOs have become a dominant alternative to traditional fee-for-service medicine. The acceptance of HMOs is part of a trend toward capitation to encourage efficiency by limiting payments to a fixed fee per person per year regardless of the quantity of the services delivered. This financial structure requires the HMO to assume the roles of both a comprehensive health care provider and a risk management organization [100]. As a result, HMOs need flexible delivery systems to accommodate the total health needs of subscribers at the lowest possible cost. While the planning models of Ittig [52] and Schneider and Kilpatrick [91] capture the ability to substitute between types of health care providers, neither model features the full range of delivery alternatives that can be provided by multiple site organizations. Flexibility in HMO delivery systems can be gained by moving specialized health care providers between clinics and by hiring a more general-purpose work force that can substitute for other personnel. Research on multiple location systems should investigate the effects of the following factors on work force composition and size decisions:

1. The sharing of personnel between clinics to allow providers to deliver services on a regular basis at more than one location
2. The substitution of technology for labor by taking advantage of the economies of scale available when demand is concentrated at a limited number of facilities
3. The increase in quality of service that occurs when only a limited number of clinics deliver a specialized service, thus increasing volume of patients treated by specialized health care providers
4. Varying the degree of labor substitution between clinics according to the preferences of the subscribers served at a specific clinic.

The growth in both the number and the size of HMOs also has generated a need for research on the design of HMO delivery systems. Research should develop an integrated model for determining the number of clinics, the location and size of each clinic, and the services offered at each location. As a multiple-site health care organization, an HMO faces the same size and location issues as those previously cited in the multihospital systems section. However, as a result of capitation-based payments, the following factors also are relevant in the design of HMO delivery systems:

1. Forecasted resource usage for different demographic cohorts within the HMO population
2. The relationship between pricing structure in the form of copayment and deductible levels and the demand for services; for example, how much will charging a fixed price per visit limit a subscriber's demand for services?
3. The financial impact of providing capacity inadequate relative to demand: in the short term, restricting the availability of services will increase waiting times causing some individuals to purchase services outside the HMO; the long-run implications of this capacity strategy may include lost future market share, additional health problems as a result of the lack of timely care, and the substitution of inappropriate care

4. A marketing policy which addresses whether a customer must limit visits to a single clinic location within the HMO system
5. Variation in resource utilization patterns across providers in the HMO.

Since the majority of HMO subscribers can change their membership in health plans during an annual open-enrollment period, it is imperative that HMO capacity acquisition research feature the relationship between current capacity decisions and their effects on future market share and competitive behavior. Both work-force and facility-acquisition models must be expanded to a multiple-period planning horizon in order to incorporate potential changes in the demand for services and the corresponding demographic characteristics of the changing membership population. Another perspective of HMO acquisition research is to emulate the work of Schneider and Kilpatrick [91] and formulate acquisition models that maximize the number of HMO subscribers served by an existing staff of health care providers. This type of approach would allow for determining the market segment that can be best served by an HMO's current resources.

Diagnosis Related Groups (DRGs)

Medicare's 1983 decision to use a fixed-price, prospective payment reimbursement system (PPS) for hospital care based on diagnosis-related groups has provided additional impetus for the movement from process to product-line management. DRGs define medically meaningful groups that are predictive of hospital resource consumption (see [4] for a thorough description of DRGs). According to Crawford and Fottler, ''DRGs have the potential for changing the manner in which hospitals are managed and the way in which hospital services are planned and administered. Although PPS/DRG is currently confined to those hospital patients who are reimbursed by Medicare, it is likely that this program, or a similar one, will be expanded to all third-party insurance systems in the future'' [21, p. 74]. In light of this shift to predetermined, fixed-price reimbursement for hospital services, in this section we discuss the impact of capitated payment systems on inpatient facility and work-force allocation decisions.

Although admissions and nurse scheduling are well-established research topics, there are numerous opportunities for research on scheduling in a product-line environment. As discussed above, most admission scheduling models in the literature maximize bed occupancy levels without considering the effect of an admission on the utilization levels of other resources. Similarly, existing nurse scheduling models use forecasted bed occupancy and historical service-mix factors to determine nursing skill requirements. In contrast, a product-line-based scheduling procedure is needed to treat demand for hospital services in terms of all of the resource requirements associated with a given product with the objective of balancing the utilization of work-force and facilities resources. Future research efforts also should consider differences in physician practice patterns and variations between patient resource requirements within specific products. Such an integrative scheduling approach would implicitly consider quality standards by appropriately matching available capacity with patient requirements.

The use of a product-line concept of scheduling implies that dependent demand approaches such as materials requirements planning (MRP) may be appropriate in health care settings. For a hospital, the resources necessary for treatment of a patient diagnosed as belonging to a specific DRG or product could be determined from prescribed medical procedures so that all patients would utilize approximately the same quantities and types of resources during their stays. While it may appear that dependent demand procedures would be directly applicable in this environment, there are several factors that need to be considered in the development of product-line scheduling procedures. First, the representation of a product in the form of a bill of material does not accurately describe the dependent demand relationship in a health care environment. The bill of material used in MRP systems is an arborescent network in that no part or activity in the bill has more than one part or predecessor; most health care products can be more appropriately described by a project network. For instance, a number of simultaneously occurring testing and preparatory activities precede surgery and other concurrent activities succeed the surgical procedure. Second, a product-based scheduling procedure should explicitly consider the availability of all resources in determining admissions and work-force schedules.

One possible method for the application of dependent demand procedures in hospitals may be a modification of Smith-Daniels and Aquilano's [99] CPM-MRP model. This technique enables the modeling of a product with a project network structure, determines a schedule of activities for each patient based on existing resource limitations, and allows for the minimization of fluctuations across facility and work-force resources. Furthermore, CPM-MRP can deal with the complexities of simultaneously scheduling surgical suites and other hospital resources.

CONCLUSION

In summary, this paper reviews the literature dealing with capacity management decisions in health care and develops several research agendas based on the results of past research efforts and the structural changes occurring in the health care industry. In the past, research efforts often failed during the implementation phase because they were unable to balance the conflicting objectives of physicians, nurses, hospital administrators, boards of directors, and other health care professionals. However, with the growth of large health care organizations such as the SuperMeds, it has been suggested by numerous authors that decisions pertaining to capacity resources now will be more well-defined because of the ease in establishing the objectives of these health care organizations. The future challenge will not be to balance diverse objectives; rather, it will be to plan and integrate capacity units. As a result, the next decade offers numerous research opportunities as health care providers position themselves in a highly competitive marketplace. [Received: August 20, 1986. Accepted: October 19, 1987.]

REFERENCES

[1] Abernathy, W. J., Baloff, N., & Hershey, J. C. The nurse staffing problem: Issues and prospects. *Sloan Management Review*, 1971, *13*(1), 87-99.

[2] Abernathy, W. J., Baloff, N., Hershey, J. C., & Wandel, S. A three-stage manpower planning and scheduling model—A service-sector example. *Operations Research*, 1973, *21*, 693-711.

[3] Abernathy, W. J., & Hershey, J. C. A spatial-allocation model for regional health-services planning. *Operations Research*, 1972, *20*, 629-642.

[4] Arbitman, D. B. A primer on patient classification systems and their relevance to ambulatory care. *Journal of Ambulatory Care Management*, 1986, *9*, 58-81.

[5] Arthur, J. L., & Ravindran, A. A multiple objective nurse scheduling model. *AIIE Transactions*, 1981, *13*, 55-60.

[6] Banker, R. D., Conrad, R. F., & Strauss, R. F. A comparative application of data envelopment analysis and translog methods: An illustrative study of hospital productivity. *Management Science*, 1986, *32*, 30-44.

[7] Barber, R. W. A unified model for scheduling elective admissions. *Health Services Research*, 1977, *12*, 407-415.

[8] Barnoon, S., & Wolfe, H. Scheduling a multiple operating system: A simulation approach. *Health Services Research*, 1968, *3*, 272-285.

[9] Berry, R. E. Returns to scale in the production of hospital services. *Health Services Research*, 1967, *2*, 122-139.

[10] Bithell, J. F. A class of discrete-time models for the study of hospital admission systems. *Operations Research*, 1969, *17*, 48-69.

[11] Carlson, R. C., Hershey, J. C., & Kropp, D. H. Use of optimization and simulation models to analyze outpatient health care settings. *Decision Sciences*, 1979, *10*, 412-433.

[12] Cavaiola, L. J., & Young, J. P. An integrated system for patient assessment and classification and nurse staff allocation for long term care facilities. *Health Services Research*, 1980, *15*, 281-306.

[13] Charnetski, J. R. Scheduling operating room surgical procedures with early and late completion penalty costs. *Journal of Operations Management*, 1984, *5*, 91-102.

[14] Chase, R. B., & Tansik, D. A. The customer contact model for organization design. *Management Science*, 1983, *29*, 1037-1050.

[15] Christianson, J. B., & McClure, W. Competition in the delivery of medical care. *New England Journal of Medicine*, 1979, *301*, 812-818.

[16] Coddington, D. C., Palmquist, L. E., & Trollinger, W. V. Strategies for survival in the hospital industry. *Harvard Business Review*, 1985, *63*(3), 129-138.

[17] Cohen, M. A., Hershey, J. C., & Weiss, E. N. Analysis of capacity decisions for progressive patient care hospital facilities. *Health Services Research*, 1980, *15*, 145-160.

[18] Cohen, M. A., & Lee, H. L. The determinants of spatial distribution of hospital utilization in a region. *Medical Care*, 1985, *23*, 27-38.

[19] Connell, B. C., Adam, E. E., & Moore, A. N. Aggregate planning in health care foodservice systems with varying technologies. *Journal of Operations Management*, 1985, *5*(1), 41-55.

[20] Connors, M. M. A stochastic elective admissions scheduling algorithm. *Health Services Research*, 1970, *5*, 308-319.

[21] Crawford, M., & Fottler, M. D. The impact of diagnosis related groups and prospective payment systems on health care management. *Health Care Management Review*, 1985, *10*(4), 73-84.

[22] Dokmeci, V. F. A quantitative model to plan regional health facility systems. *Management Science*, 1977, *24*, 411-419.

[23] Donabedian, A., Wheeler, J. R. C., & Wyszewianski, L. An integrative model of quality, cost, and health. *Medical Care*, 1982, *20*, 975-992.

[24] Egdahl, R. H. Should we shrink the health care system? *Harvard Business Review*, 1984, *62*(1), 125-132.

[25] Elwood, P. M., & Paul, B. A. *Here come the SuperMeds*. Excelsior, MN: Interstudy, 1984.

[26] Enthoven, A. C. *Health plan: The only practical solution to the soaring costs of medical care*. Reading, MA: Addison-Wesley, 1980.

[27] Erickson, G. M., & Finkler, S. A. Determinants of market share for a hospital's services. *Medical Care*, 1985, *23*, 1003-1018.

[28] Esogbue, A. M. O. *Experiments on scheduling disciplines in surgery: A simulated queuing approach*. Unpublished manuscript, Case Western Reserve University, 1969.

[29] Esogbue, A. O., & Singh, A. J. A stochastic model for an optimal bed distribution problem in a hospital ward. *Operations Research*, 1976, *24*, 884-898.

[30] Evans, R. G. "Behavioural" cost functions for hospitals. *Canadian Journal of Economics*, 1971, *4*, 198-215.

[31] Evans, R. G., & Walker, H. D. Information theory and the analysis of hospital cost structure. *Canadian Journal of Economics*, 1972, *5*, 398-418.

[32] Fetter, R. B., & Thompson, J. D. The simulation of hospital systems. *Operations Research*, 1965, *13*, 689-711.

[33] Fetter, R. B., & Thompson, J. D. A decision model for the design and operation of a progressive patient care hospital. *Medical Care*, 1969, *7*, 450-462.

[34] Folland, S. T. Predicting hospital market shares. *Inquiry*, 1983, *20*, 34-44.

[35] Geller, N. L., & Yochmowitz, M. G. Regional planning of maternity services. *Health Services Research*, 1975, *10*, 63-75.

[36] Ghosh, A., & Craig, S. Formulating retail location strategy in a changing environment. *Journal of Marketing*, 1983, *47*, 56-68.

[37] Gitlow, M. S. A methodology for determining the optimal design of a free standing abortion clinic. *Management Science*, 1976, *22*, 1289-1298.

[38] Goldman, J., Knappenberger, H. A., & Eller, J. C. Evaluating bed allocation policy with computer simulation. *Health Services Research*, 1968, *3*, 119-129.
[39] Goldman, J., Knappenberger, H. A., & Moore, E. W. An evaluation of operation room scheduling policies. *Hospital Management*, 1969, *107*, 40-51.
[40] Goldsmith, J. C. Outlook for hospitals: Systems are the solution. *Harvard Business Review*, 1981, *59*(5), 130-141.
[41] Greenfield, S. Efficiency and cost of primary care by nurses and physician assistants. *New England Journal of Medicine*, 1978, *298*, 305-309.
[42] Gupta, I., Zoreda, J., & Kramer, N. Hospital manpower planning by use of queuing theory. *Health Services Research*, 1971, *6*, 76-82.
[43] Hancock, W. M., Magerlein, D. B., Storer, R. H., & Martin, J. B. Parameters affecting hospital occupancy and implications for faculty sizing. *Health Services Research*, 1978, *13*, 276-289.
[44] Hancock, W. M., Martin, J. B., & Storer, R. H. Simulation-based occupancy recommendations for adult medical/surgical units using admissions scheduling systems. *Inquiry*, 1978, *15*, 25-32.
[45] Hancock, W. M., Warner, D. M., Heda, S., & Fuchs, P. Admissions scheduling and control systems. In J. R. Griffiths, W. M. Hancock, & F. C. Munson (Eds.), *Cost containment in hospitals.* Ann Arbor, MI: Health Administration Press, 1976.
[46] Handyside, A. J., & Morris, D. Simulation of emergency bed occupancy. *Health Services Research*, 1967, *2*, 287-297.
[47] Hefty, T. R. Returns to scale in hospitals: A critical review of recent research. *Health Services Research*, 1969, *4*, 267-280.
[48] Hershey, J. C., Abernathy, W. J., & Baloff, N. Comparison of nurse allocation policies—A Monte Carlo model. *Decision Sciences*, 1974, *5*, 58-72.
[49] Hershey, J. C., Weiss, E. N., & Cohen, M. A. A stochastic service network with application to hospital facilities. *Operations Research*, 1981, *29*, 1-22.
[50] Hornbrook, M. C., & Monheit, A. C. The contribution of case-mix severity to the hospital cost-output relation. *Inquiry*, 1985, *22*, 259-271.
[51] Huff, D. L. A programmed solution for approximating an optimum retail location. *Land Economics*, 1966, *42*, 293-303.
[52] Ittig, P. T. A model for planning ambulatory health services. *Management Science*, 1978, *24*, 1001-1010.
[53] Johnson, W. L., & Rosenfeld, L. S. Factors affecting waiting time in ambulatory care services. *Health Services Research*, 1968, *3*, 286-295.
[54] Kao, E. P. C. A semi-Markov model to predict recovery progress of coronary patients. *Health Services Research*, 1972, *7*, 191-200.
[55] Kao, E. P. C. A semi-Markovian population model with application to hospital planning. *IEEE Transactions on Systems, Man, and Cybernetics*, 1973, *SMC-3*, 327-335.
[56] Kao, E. P. C. Modeling the movement of coronary patients within a hospital by semi-Markov processes. *Operations Research*, 1974, *22*, 683-699.
[57] Kao, E. P. C., & Queyranne, M. Budgeting costs of nursing in a hospital. *Management Science*, 1985, *31*, 608-621.
[58] Kao, E. P. C., & Tung, G. C. Bed allocation in a public health delivery system. *Management Science*, 1981, *27*, 507-520.
[59] Kao, E. P. C., & Tung, G. C. Aggregate nursing requirements planning a public health care delivery system. *Socio-Economics Planning Science*, 1981, *15*, 119-127.
[60] Kaplan, R. S. Analysis and control of nurse staffing. *Health Services Research*, 1975, *10*, 278-296.
[61] Keller, T. F., & Laughman, D. J. An application of queuing theory to a congestion problem in an outpatient clinic. *Decision Sciences*, 1973, *4*, 379-394.
[62] Kolesar, P. A Markovian model for hospital admissions scheduling. *Management Sciences*, 1970, *16*, 384-396.
[63] Kralewski, J., Countryman, D., & Shatin, D. Patterns of interorganizational relationships between hospitals and HMOs. *Inquiry*, 1982, *19*, 357-362.
[64] Kwak, N. K., Kuzdrall, P. J., & Schmitz, H. H. The GPSS simulation of scheduling policies for surgical patients. *Management Science*, 1976, *22*, 982-989.
[65] Lamont, C. A. An O. R. block booking system—A review. *Hospital Administration in Canada*, 1974, *16*, 26-33.
[66] Lave, J. R., Lave, L. B., & Silverman, L. P. Hospital cost estimation controlling for case-mix. *Applied Economics*, 1972, *4*, 165-180.
[67] Liebman, J. S., Young, J. P., & Bellmore, M. Allocation of nursing personnel in an extended care facility. *Health Services Research*, 1972, *7*, 209-220.

[68] Lowerre, J. M. On personnel budgeting for continuous operations (with emphasis on hospitals). *Decision Sciences*, 1979, *10*, 126-135.

[69] Luft, H. S., & Crane S. Regionalization of services within a multihospital health maintenance organization. *Health Services Research*, 1980, *15*, 231-248.

[70] Lyons, J. P., & Young, J. P. A staff allocation model for mental health facilities. *Health Services Research*, 1976, *11*, 53-68.

[71] Magerlein, J. M., & Martin, J. B. Surgical demand scheduling: A review. *Health Services Research*, 1978, *13*, 418-433.

[72] McClain, J. O. A model for regional obstetric bed planning. *Health Services Research*, 1978, *13*, 378-394.

[73] McGuirk, M. A., & Porell, F. W. Spatial patterns of hospital utilization: The impact of distance and time. *Inquiry*, 1984, *21*, 84-95.

[74] Miller, H. E., Pierskalla, E. P., & Rath, G. J. Nurse scheduling using mathematical programming. *Operations Research*, 1976, *24*, 857-870.

[75] Milsum, J. H., Turban E., & Vertinsky, I. Hospital admission systems: Their evaluation and management. *Management Science*, 1973, *19*, 646-666.

[76] Morgan, D. M. Block booking in the O.R.—A solution to a multitude of sins. *Canadian Hospital*, 1972, *49*, 54-56.

[77] Morgan, D. M. Improved scheduling through block booking. *Canadian Hospital*, 1973, *50*, 41-45.

[78] Morrisey, M. A. The nature of hospital-HMO affiliations. *Health Care Management Review*, 1984, *9*(2), 51-60.

[79] Navarro, V. A systems approach to health planning. *Health Services Research*, 1969, *4*, 96-111.

[80] Navarro, V., Parker, R., & White, K. A stochastic and deterministic model of medical care utilization. *Health Services Research*, 1970, *5*, 342-357.

[81] Nunnally, J. C. *Psychometric Theory*. New York: McGraw-Hill, 1967.

[82] Offensend, F. L. A hospital admissions system based on nursing work load. *Management Science*, 1972, *19*, 132-138.

[83] Parker, B. R., & Srinivasan, V. A consumer preference approach to planning the rural primary health-care facility. *Operations Research*, 1976, *24*, 991-1025.

[84] Parker, R. D. Variation of the occupancy of two medical units with the amount of sharing between the units. *Health Services Research*, 1968, *3*, 214-223.

[85] Prastacos, G. P. Blood inventory management. *Management Science*, 1984, *30*, 777-800.

[86] Ramsey, J. A., McKenzie, J. K., & Fish, D. G. Physicians and nurse practitioners: Do they provide equivalent health care? *American Journal of Public Health*, 1982, *72*, 55-57.

[87] Rising, E. J., Baron, R., & Averill, B. A systems analysis of a university-health-service outpatient clinic. *Operations Research*, 1973, *21*, 1030-1047.

[88] Robinson, G. H., Wing, P., & Davis, L. E. Computer simulation of hospital patient scheduling. *Health Services Research*, 1968, *3*, 130-141.

[89] Ruth, R. J. A mixed integer programming model for regional planning of a hospital inpatient service. *Management Science*, 1981, *22*, 521-533.

[90] Schmitz, H. H., & Kwak, N. K. Monte Carlo simulation of operating room and recovery room usage. *Operations Research*, 1972, *20*, 1171-1180.

[91] Schneider, D. P., & Kilpatrick, K. E. An optional manpower utilization model for health maintenance organizations. *Operations Research*, 1975, *23*, 869-889.

[92] Shannon, G. W., Bashur, R. L., & Metzner, C. A. The concept of distance as a factor in accessibility and utilization of health care. *Medical Care Review*, 1969, *26*, 143-161.

[93] Shonick, W. A stochastic model for occupancy-related random variables in general-acute hospitals. *Journal of the American Statistical Association*, 1970, *65*, 1474-1500.

[94] Shonick, W., & Jackson, J. R. An improved stochastic model for occupancy-related random variables in general-acute hospitals. *Operations Research*, 1973, *21*, 952-965.

[95] Shulka, R. K. Admissions monitoring and scheduling to improve work flow in hospitals. *Inquiry*, 1985, *22*, 92-101.

[96] Shuman, L. J., Hardwick, C. P., & Huber, G. A. Location of ambulatory care centers in a metropolitan area. *Health Services Research*, 1973, *8*, 121-138.

[97] Shuman, L. J., Young, J. P., & Naddor, E. Manpower mix for health services: A prescriptive regional planning model. *Health Services Research*, 1971, *6*, 103-119.

[98] Smith, K. R., Over, A. M., Jr., Hansen, M. F., Golladay, F. L., & Davenport, E. J. Analytic framework and measurement strategy for investigating optimal staffing in medical practice. *Operations Research*, 1976, *24*, 815-841.

[99] Smith-Daniels, D. E., & Aquilano, N. J. Constrained resource project scheduling with CPM-MRP. *Journal of Operations Management*, 1984, *4*(4), 369-387.

[100] Sorkin, A. L. Health care and the changing economic environment. Lexington, MA: Lexington Press, 1986.

[101] Staff, P. J., & Vagholkar, M. K. Stationary distribution of open Markov processes in discrete time with applications to hospital planning. *Journal of Applied Probability*, 1971, *8*, 668-679.

[102] Steinberg, E., Khumawala, B., & Scamell, R. Requirements planning systems in the health care environment. *Journal of Operations Management*, 1982, *2*(4), 251-259.

[103] Studnicki, J. The geographic fallacy: Hospital planning and spatial behavior. *Hospital Administration*, 1975, *20*, 10-21.

[104] Sumner, A. J., & Hsieh, R. K. C. Long range predictions of examining room requirements. *Health Service Research*, 1972, *7*, 221-230.

[105] Swain, R. W., Kilpatrick, K. E., & Marsh, J. J. Implementation of a model for census prediction and control. *Health Services Research*, 1977, *12*, 380-395.

[106] Swartzman, G. The patient arrival process in hospitals: Statistical analysis. *Health Services Research*, 1970, *5*, 320-329.

[107] Thomas, D. E., & Stokes, H. H. How many beds should a hospital department serve? *Health Services Research*, 1976, *11*, 241-251.

[108] Thomas, W. H. A model for predicting recovery progress of coronary patients. *Health Services Research*, 1968, *3*, 185-213.

[109] Thompson, J. D., & Fetter, R. B. Economics of occupancy with varying mixes of private and other patient accommodations: A simulation. *Health Services Research*, 1969, *4*, 42-52.

[110] Toregas, C., Swain, R., ReVelle, C., & Bergman, L. The location of emergency service facilities. *Operations Research*, 1971, *19*, 1363-1373.

[111] Trivedi, V. M. A mixed integer programming model for nursing service budgeting. *Operations Research*, 1981, *29*, 1019-1034.

[112] Trivedi, V. M., & Warner, D. M. A branch and bound algorithm for optimum allocation of float nurses. *Management Science*, 1976, *22*, 972-981.

[113] Vargas, G., Hottenstein, M., & Aggarwal, S. Using utility functions for aggregate scheduling in health maintenance organizations (HMO's). *AIIE Transactions*, 1979, *11*, 327-335.

[114] Vraciu, R. A. Hospital strategies for the eighties: A mid-decade look. *Health Care Management Review*, 1985, *10*(4), 9-19.

[115] Warner, D. M. Scheduling nursing personnel according to nursing preference: A mathematical programming approach. *Operations Research*, 1976, *24*, 842-870.

[116] Warner, D. M., & Prawda, J. A mathematical programming model for scheduling nursing personnel in a hospital. *Management Science*, 1972, *19*, 411-422.

[117] Weiss, G. H., & Zelen, M. A semi-Markov model for clinical trials. *Journal of Applied Probability*, 1965, *2*, 269-278.

[118] Willemain, T. R., & Moore, G. T. Planning a medical practice using paramedical personnel. *Health Services Research*, 1974, *9*, 53-61.

[119] Wind, Y., & Spitz, L. Analytical approach to marketing decisions in health-care organizations. *Operations Research*, 1976, *24*, 973-990.

[120] Wolfe, H., & Young, J. P. Staffing the nursing unit, Part I: Controlled variable staffing. *Nursing Research*, 1965, *14*, 236-243.

[121] Wolfe, H., & Young, J. P. Staffing the nursing unit, Part II: The multiple assignment technique. *Nursing Research*, 1965, *14*, 299-303.

[122] Young, J. P. *A queuing theory approach to the control of hospital inpatient census*. Unpublished doctoral dissertation, Industrial Engineering Department, Johns Hopkins University, 1962.

[123] Zimmerman, O. C. Utilizing operating room time. *Hospital Topics*, 1963, *41*, 99-102.

[124] Zuckerman, H. S. Industrial rationalization of a contact industry: Multi-institutional hospital systems. In S. E. Berki (Ed.), *Annals of the American Academy of Political and Social Science: Health Care Policy in America*, 1983, *468*, 216-230.

Vicki L. Smith-Daniels is Assistant Professor of Operations Management in the Department of Purchasing, Transportation and Operations at Arizona State University. She received her Ph.D. from The Ohio State University and a B.B.A. from the University of San Diego. Her current research interests are in the areas of process industries scheduling, multistage lot sizing, sequence-dependent setup-time sequencing, lot sizing in a project scheduling environment, and health care operations. Her articles have appeared in *IIE Transactions*, *Journal of Operations Management*, *International Journal of Production Research*, and *Production and Inventory Management*. She is a member of the Decision Sciences Institute, TIMS, ORSA, and APICS.

Sharon Bergman Schweikhart is a Ph.D. candidate in operations management at the University of Minnesota. She received an M.B.A. from the University of Minnesota and a B.S. in industrial engineering from Wayne State University. Her current research interests are in the areas of health care operations management, facility location, and quality management. She is a member of the Decision Sciences Institute, IIE, and TIMS.

Dwight E. Smith-Daniels is Assistant Professor of Operations Management in the Department of Purchasing, Transportation and Operations at Arizona State University. He received his Ph.D. from the University of Arizona and a B.B.A. from the University of Michigan. His current research interests are in the areas of scheduling and materials management in project networks, use of the net present value criterion in operations scheduling, and health care management. His publications have appeared in *Decision Sciences*, *IIE Transactions*, *Journal of Operations Management*, and *Production and Inventory Management*. He is a member of the Decision Sciences Institute, TIMS, PMI, IIE, and APICS.