Scheduling of Patients and Resources for Ambulatory Health Care

Stephen R. Smith, Bernard J. Schroer, PhD,
Robert E. Shannon, PhD

Engineering Analysis, Inc., University of Alabama in Huntsville

Abstract

This paper presents the results of a study of the scheduling of patients and resources in one of the Family Practice Units of The University of Alabama in Huntsville's Ambulatory Care Center. The objective of the study was three-fold:

1) Determine the maximum number of patients the physician staff can see in a 3-hour session,

2) Find the patient appointment schedule for this load that minimizes patient waiting without impacting through-put, and

3) Determine the number of nurses and examining rooms needed to process this patient load in the most efficient, cost-effective manner.

A detailed computer model of the Unit was written using SLAM (Simulation Language for Alternative Modeling) which incorporated parameters for theoretical distributions based on nearly a month's actual data taken in the Unit. Using the computer model a number of patient loads, appointment schedules, and resource allocations were examined. A general cost/benefit equation was formulated for evaluating the alternatives, and recommendations were made on how to optimize the Unit's operation.

INTRODUCTION

In recent years the demand for ambulatory care has placed increasing workloads on outpatient clinics. This coupled with the escalating costs of health care has served to intensify research in the area of resource scheduling and utilization. Attention is being focused primarily on achieving high utilization rates for the clinic's physician staff. The costs associated with these high rates are of two general types: the number and types of support personnel and facilities and the length of time patients have to wait for service.

The first type of cost is more directly measurable. The cost of hiring an additional nurse or of adding an examination room is known. The question is, how many nurses or examining rooms are needed to insure that the physicians are not forced to wait frequently or for long periods? In addition, will the cost of adding support personnel or facilities be offset by the increase in physician utilization? Answering these questions without actually trying different combinations of support personnel and facilities is difficult. Some type of mathematical model of the system that takes into account a variety of contributing factors must be employed.

The second type of cost, patient waiting, is not so easily measurable. However, it is receiving increasing attention as an important factor in the operating strategy of the outpatient clinic. Large backlogs of waiting patients not only necessitate larger reception or waiting areas, but they also have an adverse psychological (and sometimes medical) impact on patients and staff alike. There is the cost to the patient or to his employer of lost time in the doctor's office. In addition, clinic sessions may often run past the scheduled closing time, inconveniencing everyone.

A number of studies on outpatient scheduling have focused on balancing the patient's waiting time with the physician's idle time. As noted by Steidley and Vanloh (6), several factors complicate the search for the optimum appointment schedule: the incidence of no-shows (patients who break appointments); the number of walk-ins (patients appearing without an appointment); frequent late arrival of the physicians to start the clinic session. In addition there are other factors, including patients arriving late, variation in the types and length of patient services, the time for physicians to update patient records, physician consultations, and emergency interruptions.

Due to the complex nature of the above problems, deterministic models have limited application. Digital computer simulation then becomes a primary tool in evaluating various strategies.

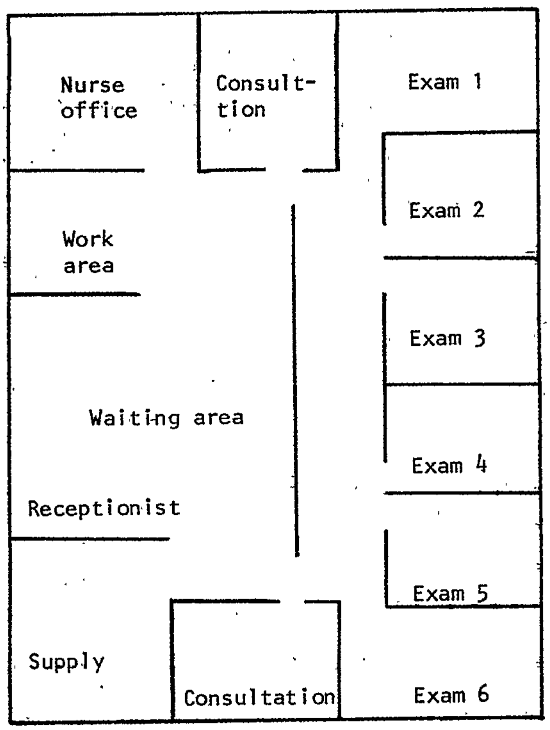
CH1437-3/79/0553-0561$00.75 © 1979 IEEE
1979 Winter Simulation Conference

PHYSICAL SYSTEM

This study uses the Monte Carlo, discrete event simulation technique to analyze the outpatient clinic at the Ambulatory Care Center (ACC) of the University of Alabama in Huntsville's School of Primary Medical Care. This clinic has four modules each staffed by a faculty physician; first, second, and third year residents; nurses; and other support personnel. Each module is designed to model most features of a small private clinic. The layout of a module is given in Figure 1.

Figure 1

TYPICAL CLINIC MODULE LAYOUT

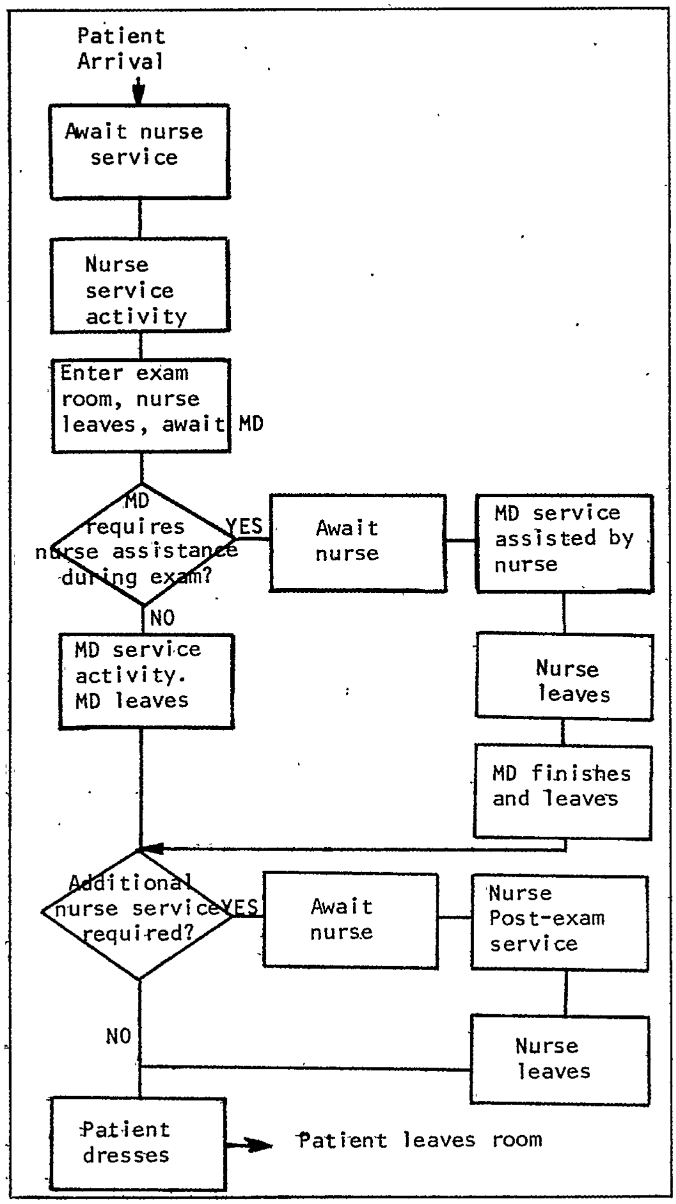


Each module operates two, three-hour sessions per day and is staffed by one to four physicians, each physician having his own patient appointments. The particular physician staff for a session which most closely resembles the private clinic setting consists of the faculty physician and two second/third year residents.

For analysis purposes, the significant events in a patient's flow through the module are depicted in Figure 2. Events such as medical records retrieval, receptionist sign-in activities, lab and x-ray services, billing, follow-up appointments, physician consultation activities and interruptions due to emergencies are ignored. Instead the study focuses primarily on nurse/physician/patient interactions under normal conditions.

Figure 2

PATIENT FLOW THROUGH MODULE



SIMULATION MODEL

Since a patient's movement through the system is dependent upon a number of interactive, random processes, a tailored high-level simulation language was needed. SLAM was chosen because it provides a convenient framework for the discrete process orientation while at the same time allowing the modeler to code his own event-oriented subroutines (1).

The goal of the simulation was to evaluate different patient appointment strategies in conjunction with various configurations of nurse and exam room assignments. Based upon interviews with ACC staff members, observation of the clinic in operation, and research of pertinent literature, the variables in Table 1 were identified as significant. These were used to initialize the simulation.

TABLE 1

MODEL VARIABLES

Description
Number of physicians and whether staff or resident
Number of nurses
Number of exam rooms assigned to each physician
Patient appointment schedule for each physician
Nurse pre-exam service time parameters
Physician's service time parameters
Percentage of patients requiring nurse's assistance during the physician's exam
Mean time the nurse spends in the exam room with the physician
Percentage of no-shows
Percentage of walk-ins.
Parameters for patient arrival about appointment times
Parameters for physician arrival times in the clinic
Mean time per patient required for physician to update the patient's records (dictation)
Mean number of patients the physician sees before stopping to update patient records
Percentage of patients requiring nurse's service after the physician is finished
Parameters for patient dressing times.

Figure 3 depicts the basic operation of the model. From one to four physicians can be modeled with any number of nurses and any number of exam rooms.

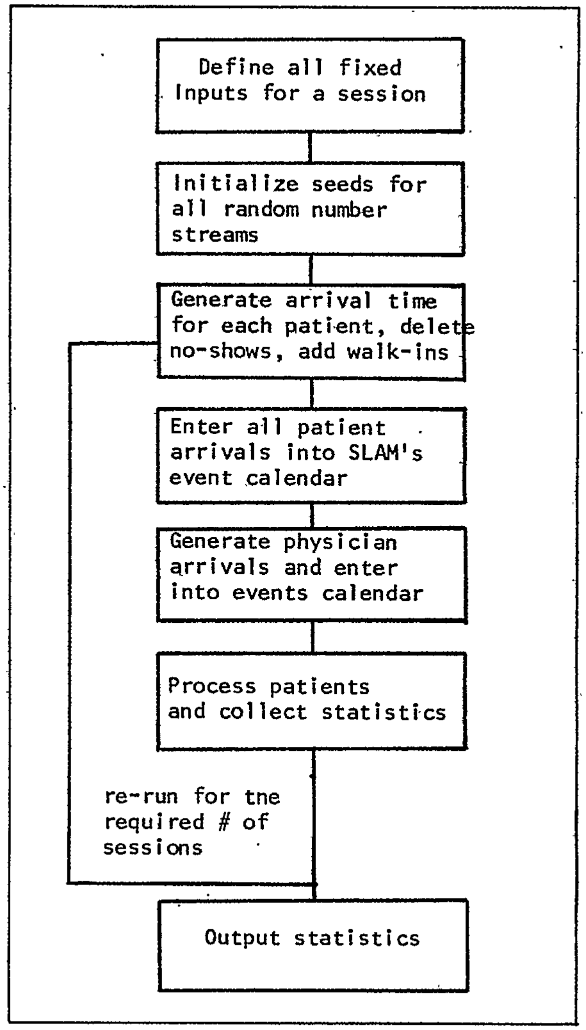
Since the nurses perform essentially three different activities, priorities are assigned to determine which activity is done when there are conflicting demands. The order of precedence for these activities is:

1) Provide assistance for a physician seeing a patient in an exam room (primarily when the patient is female or a child)

2) Perform required post-exam service for a patient before he leaves (primarily injections, catheterizations, etc.)

Figure 3

MODEL OPERATION



3) Perform pre-exam service for a patient from the waiting room and take him to an exam room to await his physician (primarily includes taking patient's weight, blood pressure, and temperature)

Although each nurse works for all the physicians, the physicians are ranked on the basis of seniority. Given two physicians requiring the same type of assistance from the nurse, the nurse will always work for the physician of higher rank.

Primary outputs from a run included the following:

1) Mean session length

2) Mean time in the system for a patient (for all patients and by physician)

3) Mean waiting time in the waiting room (all patients and by physician)

4) Mean time waiting in an exam room (all patients and by physician)

5) Mean time in the system for patients entering during the first hour, the second hour, and the third hour

6) Mean utilization rate of each physician, the nurses, and the exam rooms

7) Mean time each physician finished with his last patient.

EMPIRICAL DATA

The data collection was done by the employees working in the module. Using these employees rather than an impartial observer left questions as to objectivity. Thus an outside observer was also used to watch the operation for significant activities that had not been identified and to spot-check the collected data. A rubber stamp was developed to stamp each patient charge ticket. The appropriate employee then entered the appropriate data on the charge ticket. The following data were entered: patient number and sex, physician number, time of appointment, time of arrival, time nurse called the patient, time patient entered the examining room, time physician entered the room, time physician left the examining room, and time patient left the examining room.

The receptionist entered patient number, sex, physician, time of appointment and time of arrival (which she took from the general patient sign-in sheet). The nurse entered the time she called the patient and the time she left him in the exam room. The physician recorded the times he entered and left the exam room.

Data were collected for a three-week period on 412 patients. The data were then edited to eliminate erroneous entries and outliers. Statistics were computed for each parameter and hypotheses formulated to test the empirical distributions. Goodness of fit tests using the Kolmogorov-Smirnov and Chi-Square methods were performed with Table 2 giving the final results. The data revealed that second and third year residents worked about equally fast, thus the two categories were combined. In addition, the nurses were assumed equivalent in speed of service.

Values for the other parameters were either extracted from the data or estimated by the medical staff. The values of each of these parameters are given in Table 3.

TABLE 2

SERVICE TIMES

Parameter	Mean	Standard Deviation	Empirical Distribution
Patient arrival about appointment times	-8.25	14.03	Normal
Nurse pre-exam service	5.18	5.29	Exponential
Faculty physician service	11.16	9.14	Log normal
2nd/3rd year residence service	14.51	11.14	Log normal
1st year residence	24.78	19.17	Log normal
Patient dressing time	6.44	7.78	Exponential

TABLE 3

OTHER INPUT PARAMETERS

Parameter	Value Assigned
Patients requiring nurse assistance during physician exam	50%
Time required for nurse assistance during physician exam	(μ,σ)= 6.,2. (Normal distribution)
Number of patients physician sees before updating records	3
Time required for physician to update records	5 min. (constant)
No-show rate	10%
Walk-in rate	10% (Poisson distr. Lambda = 1.0)
Walk-in arrival times	Uniform distr. (0,180)
Patients requiring nurse post-exam service	75%
Nurse post-exam service time	4 min. (constant)
Physician arrival times	Uniform distr. (0,5)

VALIDATION AND EXPERIMENTAL DESIGN

The simulation model was run with four variable inputs:

1) Number of physicians scheduled (1 to 4) and the rank of each

2) Patient appointment schedule for each physician (given in minutes from session start-up)

3) Number of exam rooms assigned to each physician (generally each physician has exclusive use of the exam rooms assigned to him)

4) Number of nurses.

The model was validated using the actual values of the four input variables that were used in a clinic session and using the SLAM's trace capability which shows the step by step transition of patients through the system. The results compared closely with the collected data as indicated in Table 4.

TABLE 4

MODEL VALIDATION RESULTS

Variable	Actual (min)	Model (min)
Mean patient time in the system	56.16	54.70
Mean patient wait time	29.32	25.00
Mean session length (last patient out)	180.00	211.10
Mean time last physician finishes	180.00	183.58

Using the sample mean and variance for patient time in the system as calculated in the validation run, it was determined that 50 replications were required to insure 85% confidence of being within 5 minutes of the mean time in the system (4).

As noted earlier, the physician staff that most closely resembles the private clinic setting consists of three physicians, the faculty physician and one second and one third year resident. This configuration was assumed in all model runs. Thus, for design purposes, the experiment could be described by a three-dimensional grid, patient appointment schedule, number of nurses, and number of exam rooms. The number of nurses was varied from two to four. The number of exam rooms was varied from six to nine with the assignments for each physician given in Table 5.

TABLE 5

ASSIGNMENT OF EXAM ROOM

Number of exam Rooms	Assigned to faculty phy	Assigned to Resident	Assigned to Resident
6	2	2	2
7	3	2	2
8	3	3	2
9	3	3	3

Considerable attention was focused on the selection of various patient appointment schedules. A review of pertinent literature on the subject revealed that the modified-wave approach has the greatest potential for maximizing resource utilization and patient throughout while minimizing patient waiting. (2,5) Rather than having the patients scheduled uniformly (e.g., two every fifteen minutes), the modified-wave schedules more patients toward the beginning of each hour and less toward the end of the hour. This allows the physician to "catch up" and be back on schedule at the beginning of each hour, hopefully eliminating the long patient waits caused by carry-over from the previous hour. An initial study at the ACC where this approach was implemented, indicated improved performance over the old scheduling method. (4)

EXPERIMENTAL RESULTS

The experiment was conducted in three phases. In the first phase the objective was to determine the maximum patient load for the three physicians that would insure high physician utilization and low patient waiting times while holding the session to three hours. In the second phase the object was to process this patient load in the most efficient, cost-effective manner by varying patient scheduling, number of nurses, and number of exam rooms. The objective of the third phase was to optimize the resources including nurses, physicians, and exam rooms.

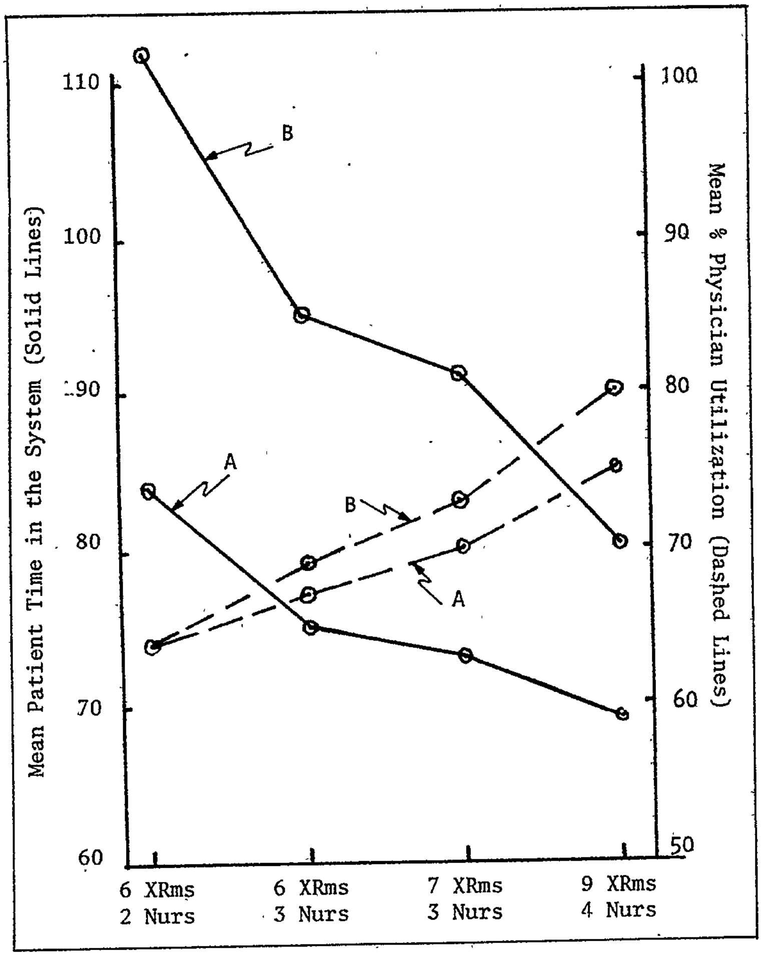
MAXIMUM PATIENT LOAD

The first set of simulation runs was made with 38 patients. Given 100% utilization, the faculty physician can see at most 14 patients in a three-hour session and the other two physicians at most 12 patients each. Two sets of 12 runs each were made using this patient load with different appointment schedules.

Figures 4 and 5 show the results of four selected runs using Schedule A, the modified wave and Schedule B, two patients every 15 minutes. Physician utilization is high for both schedules. Note that the true utilizations are higher than those shown because the physicians are doing some activities not modeled. As expected, increasing the number of exam rooms and the number of nurses permits greater physician utilization. Mean time for a patient in the module is very high--especially for Schedule B. Since all the patients in Schedule B arrive within an hour and a half of session start-up, most patients wait for long periods. Patient time in the system is also unacceptably high for Schedule A when we consider that, on the average, only 22 and 25.4 minutes are needed to process a patient for the faculty physician and the resident physicians, respectively. Thus, the patients would spend from 50 minutes to over an hour and a half waiting for service.

Figure 4

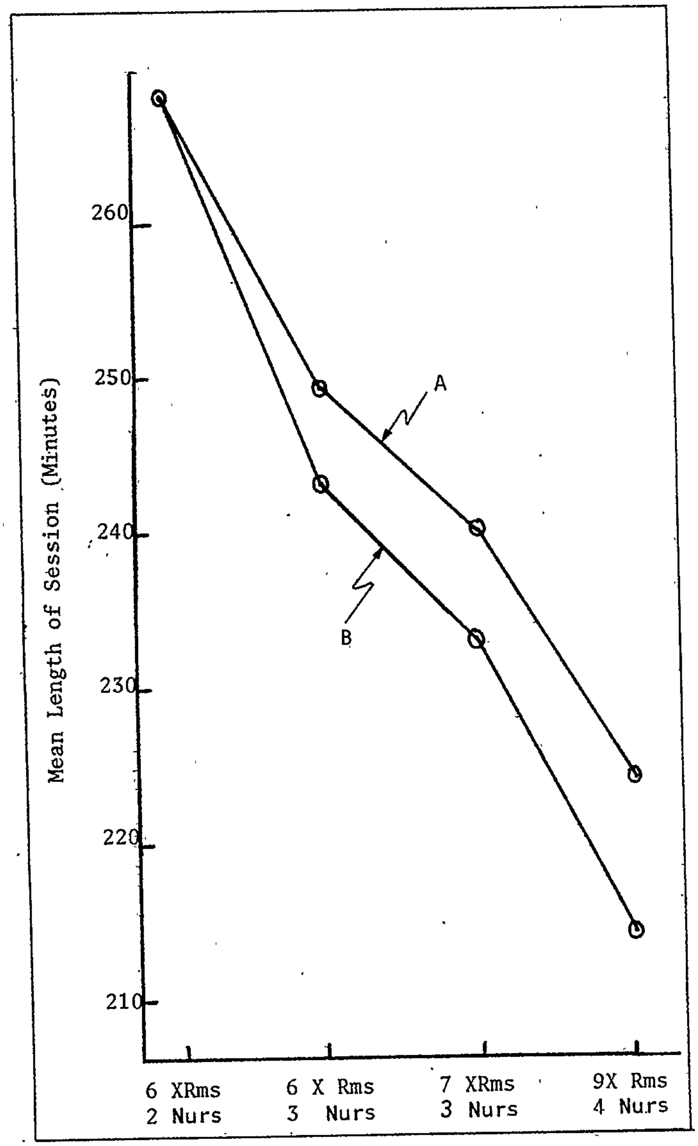
COMPARISON OF SCHEDULES A AND B



Session times for these two schedules would run from 30 minutes to an hour and 20 minutes past closing time, where Schedule B is slightly better due to higher physician utilization rates. Clearly a patient load of 38 is too high for these 3 physicians.

Figure 5

COMPARISON OF SCHEDULES A AND B



Three other patient loads were considered; 30, 29, and 27 patients, respectively for Schedules C, D, and E. Figures 6 and 7 give the results of selected runs with these schedules. Although physician utilizations are less than with the previous 38-patient cases, they are still reasonable and mean patient waiting time has fallen in most cases to the 35- to 45-minute range. Session lengths for Schedules C and D are still probably unacceptable in all but the last case. For Schedule E, session lengths are near the 3-hour mark, though generally on the high side. This correlates with the actual data which shows 412 patients in roughly 15 sessions for an average of 27.5 patients per session.

Thus, 27 patients (11, 8, and 8 for the faculty physician and residents, respectively) appears to be a reasonable patient load. All phase 2 runs were made using two different appointment schedules for these 27 patients (Schedules E and F).

Figures 8 and 9 indicate that for all practical purposes Schedules E and F are equivalent. In fact, there are only a few ways 11, 8, and 8 patients can be logically scheduled using the modified-wave approach.

Figure 6

COMPARISON OF SCHEDULES C, D, AND E

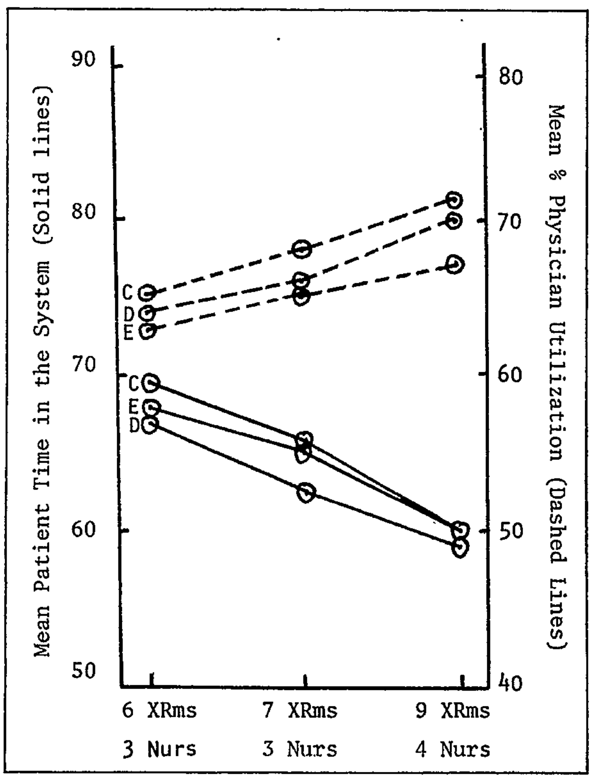


Figure 7

COMPARISON OF SCHEDULES C, D, AND E

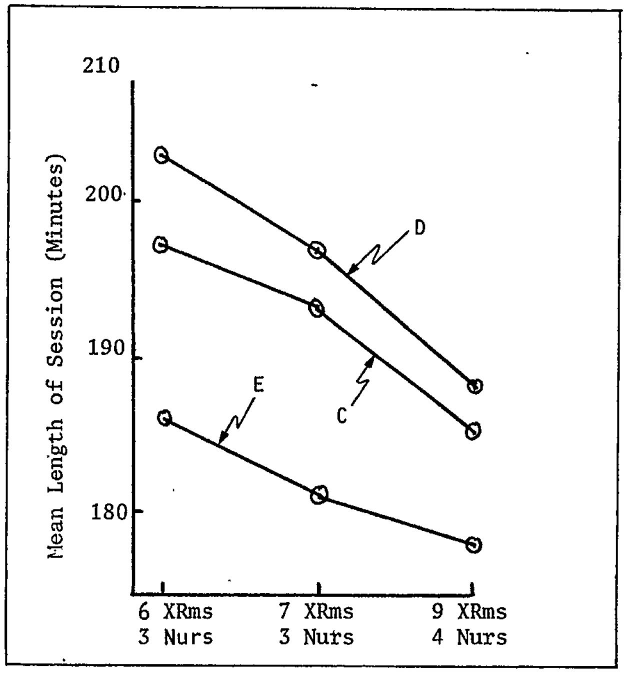


Figure 8

COMPARISON OF SCHEDULES E AND F

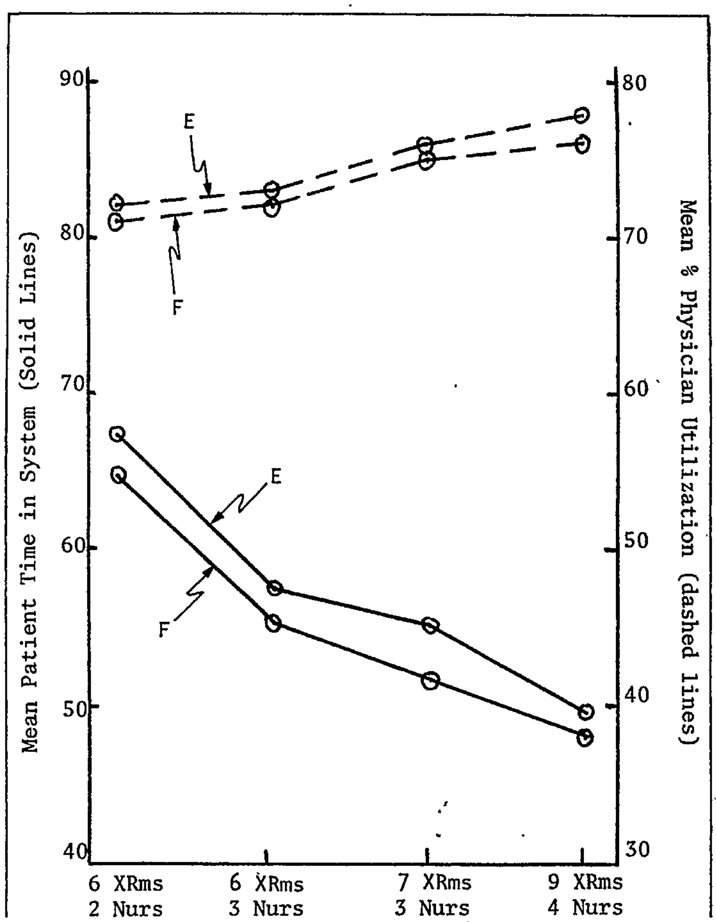
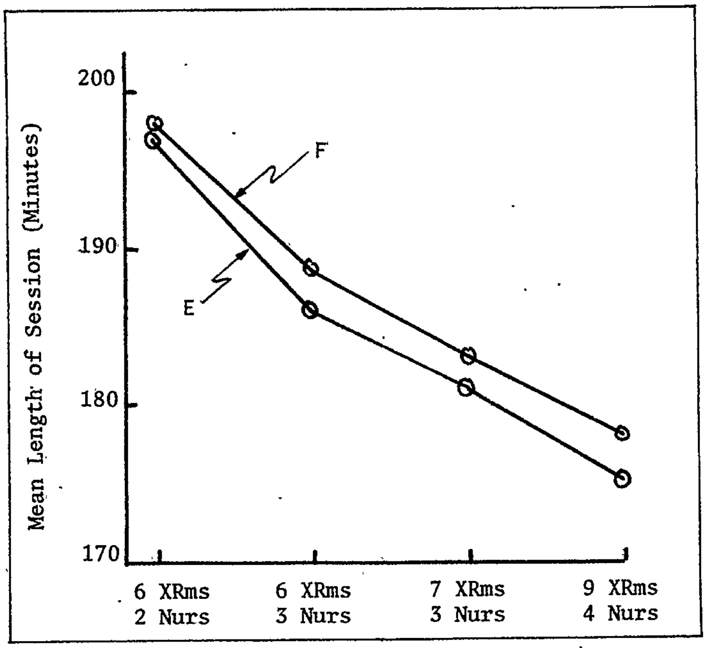


Figure 9

COMPARISON OF SCHEDULES E AND F



OPTIMAL ALLOCATION OF RESOURCES

Patient Schedule E was used in the third phase of the experiment. Figures 10 through 12 show some of the results where the number written beside each curve corresponds to the total number of exam rooms. The most significant improvement in the physician utilization, patient time in the system, and session length occurs by increasing the number of nurses from two to three and the number of exams rooms from six to seven. However, as the number of nurses increase, the mean nurse utilization decreases sharply, regardless of the number of exam rooms. Also, mean exam room utilization decreases slightly by increasing the number of both nurses and exam rooms.

Figure 10

SCHEDULE E VARYING NURSES AND EXAM ROOMS

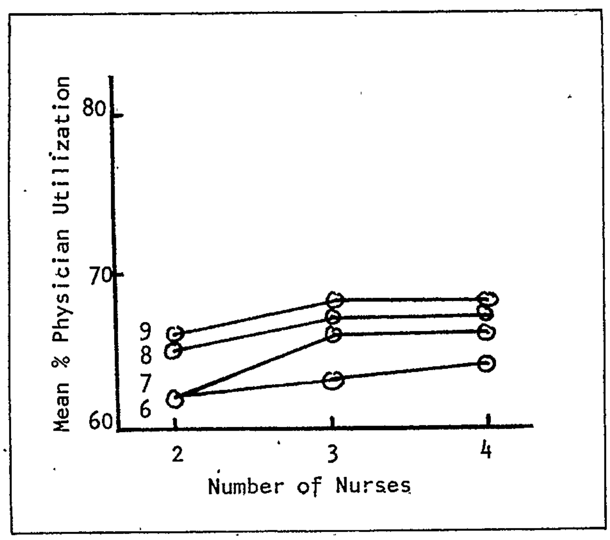


Figure 11

SCHEDULE E VARYING NURSES AND EXAM ROOMS

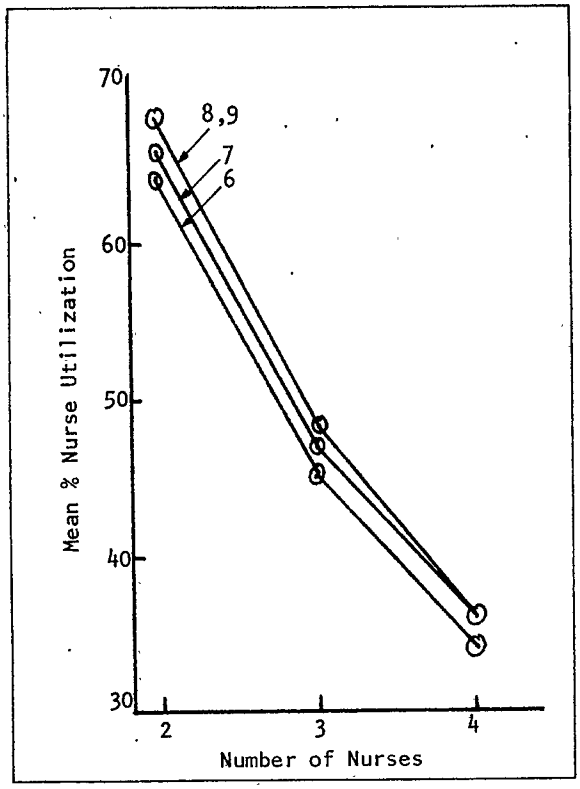
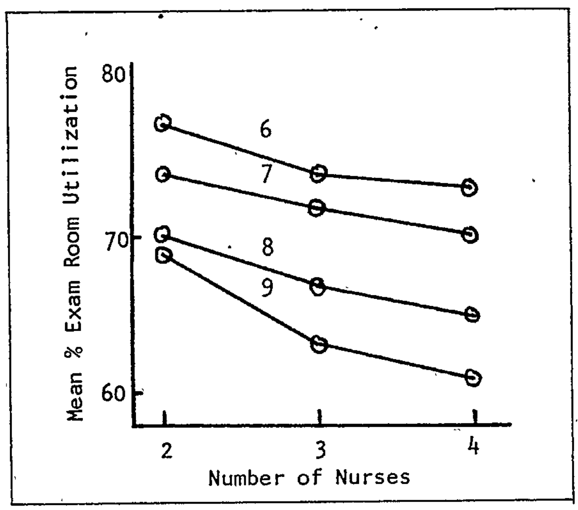


Figure 12

SCHEDULE E VARYING NURSES AND EXAM ROOMS



Improvement is not necessarily achieved by keeping the nurses busy and the exam rooms occupied, but by keeping the physicians busy and the patient backlog low. Although it is very desirable to have high utilization of nurses and exam rooms, it is most important to have a nurse and an exam room available when needed. Thus, intuitively it appears that three nurses and seven exam rooms are the most cost-effective arrangement for these physicians under this patient load. A careful assessment of the cost/benefit ratios associated with these options is necessary before a final selection can be made.

The most easily measured cost/benefit parameter is the cost of running a session vs the revenue generated during the session. The net cost for a session can be calculated by:

cost =

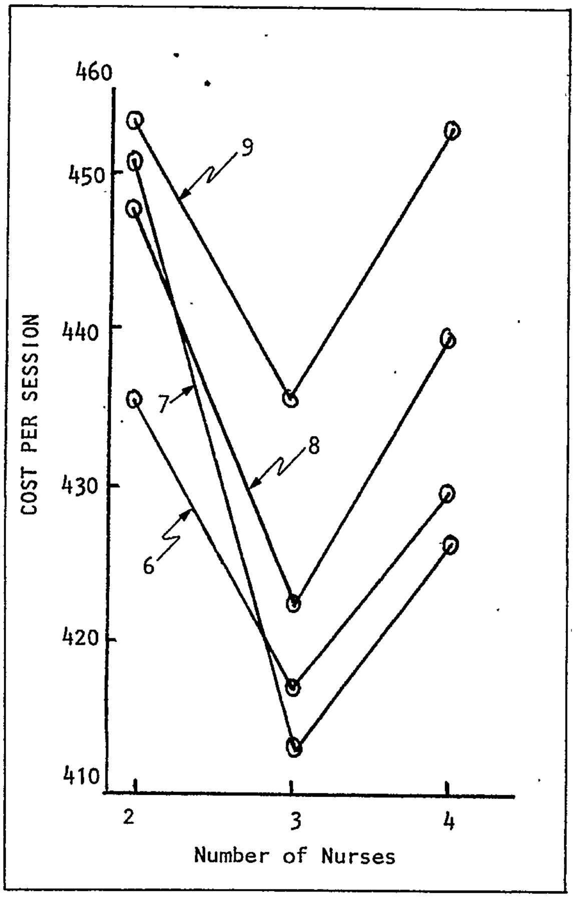
(# of physicians) * (physician cost per session)
+(# of nurses) * (nurse cost per session)
+(# of physicians) * (overtime cost for late sessions)
+(# of nurses) * (overtime cost for late sessions)
+(mean patient wait) * (#of patients)
* (cost of waiting)
+(# of exam rooms) * (cost per exam room per session)
+(cost of the waiting area per session).

Using some hypothetical values for these cost factors, the above formula was applied to the Schedule E outputs. The results are given in Figure 13.

In all cases three nurses were optimal and seven exam rooms were slightly less costly than six with three nurses. Nine exam rooms were the least desirable and four nurses slightly better than two. Of course, these plots can be totally different depending on how the costs are estimated. Patient waiting cost is a value judgment and depends on who is making the estimate (physician or patient).

Figure 13

COMPARISON OF OPERATING COSTS



CONCLUSIONS

In summary, the following general observations may be made:

- Given this particular physician staff, the modified wave scheduling method with 27 patients is optimal.
- Additional patients raise resource utilization only slightly and force the session into overtime.
- Nurse and exam room utilizations are not good indicators of clinic efficiency.
- In determining resource allocation and patient loading, careful cost/benefit analysis should be done and costs assigned to such items as patient waiting, late-running sessions, and crowded reception areas.
- For this particular physician staff the clinic operates most efficiently and cost-effectively with three nurses and seven exam rooms.

As the mix of the resident staff at the ACC changes, and as total patient load fluctuates, decisions must be made concerning the number of nurses and exam.rooms. This study provides a good estimate of the effects of these changes in the system. In addition, other alternatives, such as increasing the hours of operation and varying the number of physicians, may be easily evaluated. One model enhancement that might be particularly useful would be to consider different types of patient visits which require unusual service or service times (such as physicals).

REFERENCES

1. Pritsker, A. Alan B.; Pegden, Claude Dennis, Introduction to Simulation and SLAM, John Wiley and Sons, New York, 1979, pp. 64-73.

2. Rockhart, J., et.al., "A Predictive Model for Ambulatory Patient Service Time," Medical Care, 12:512-9, June 1974.

3. Schroer, Bernard J.; Smith; Herbert T., "Effective Patient Scheduling," The Journal of Family Practice, Vol. 5, No. 3, 1977, pp. 407-411.

4. Shannon, Robert E., Systems Simulation: The Art and Science, Prentice Hall, 1975, pp. 186-192.

5. Silver, Michael, "Scheduling: Least Developed Art," Family Practice News, Vol. 5, No. 23, Dec. 1, 1975.

6. Steidley, K. David; Vanloh, Sidney, "A Monte Carlo Computer Model to Investigate Patient Scheduling," Radiology, 122:359-364, Feb. 1977.