A SIMULATION MODEL OF A UNIVERSITY HEALTH SERVICE

OUTPATIENT CLINIC

by

Robert Baron
Research Assistant, Department of Industrial Engineering and Operations Research
University of Massachusetts
Amherst, Massachusetts

and

Edward J. Rising
Professor of Industrial Engineering and Operations Research
University of Massachusetts
Amherst, Massachusetts

October, 1972

Sixth Annual Joint Simulation Conference
San Francisco, California (January 17-19, 1973)

Abstract

This paper focuses on the development of a simulation model of a University Health Service Outpatient Clinic the implementation of which has resulted in significant improvements to system performance. The details of these improvements are published elsewhere; they amouted to savings in excess of fifty thousand dollars the first year the model was used, improved physician morale, and acceptance on the part of the Health Service staff of the simulation model as a tool for decision-making.

The Health Service provides complete outpatient medical care and limited inpatient care for about 19,000 people. The resulting outpatient load of over 400 persons a day requires the services of 12 full-time physicians. The simulation model for which appointment and walk-in patients are generated separately, was developed over a two-year period and takes the general form of a multiple stage, parallel queueing system with a variable number of servers. Validation problems are discussed, and data is presented.

Introduction

The literature in the field of health administration reflects the increasing importance of and burgeoning national interest in the delivery of health care through outpatient facilities. This paper reports on the details of a simulation model that was used as a portion of an overall systems analysis made of the delivery of outpatient care at the University of Massachusetts Health Service. The manner in which this systems analysis was carried out and the results it achieved were described in the Journal of the American College Health Association in a three article series in the June 1972 issue (1, 2, 3). These improvements were achieved through the reduction of physician idle time. In a concurrent study made by a team of sociologists who did a before and after set of interviews with the physicians, it was concluded that physician morale increased because of the work done.

The increase in throughput that was possible, together with the increased time spent with patients, and the fewer physician hours actually scheduled for patient contact, meant that the systems analysis was responsible for providing the students with services that would have required approximately 2.2 additional physicians operating under the old system. This meant a saving in excess of $50,000 in the first year in physicians' salaries alone, and if one also includes the support services that these two physicians would have required, this figure for savings would increase substantially.

The complete systems analysis used a simulation model to examine the effect of various strategies for scheduling the appointments of patients and for examining the effect of different working schedules for the physicians. Runs were compared on the basis of patient waiting time and physician idle time, the two most sensitive measures of effectiveness. Based on these criteria, the medical staff reviewed the results and decided on a scheduling pattern for themselves and their patients for the following academic year.

The key to the success of the enterprise was twofold: first, the entire system was analyzed by a team including sociologists, physicians and administrators as well as engineers; and second, the analysts were cast in the role of supplying staff support to both the clinic administrators and the medical staff. The clinic administrators formulated alternative scheduling patterns based on the questions that the medical staff raised, then the simulation analysts ran the model to replicate this situation. The results of the various simulation runs were examined by the medical staff and the clinic management under the guidance of the simulation analysts. The resulting decision represented the needs of the medical staff and the preferences of the clinic administration, and it took advantage of the technological expertise of the simulation analysts. The details of how the model was developed, the manner in which data were taken and the kinds of results that were obtained from the model are presented in the remainder of this article, but it must be remembered that this simulation model was only one facet of the enterprise.

Description of Facilities

In the fall of 1970, the University of Massachusetts Health Service delivered primary health care to approximately 19,000 students on a compulsory prepaid basis. There were, in addition to the outpatient department that is of interest here, approximately 70 inpatient beds, a laboratory and x-ray facilities, an emergency room, a pharmacy, and a mental health clinic that is separately housed. The University Health Service also operates a health education program and an environmental health and safety program.

The outpatient department usually treats between 400 and 500 patients per day. About half of these patients see a physician on either an appointment or a walk-in basis. The remaining patients visit clinics such as the nurse-practitioner clinic, where four nurses deliver primary care under the direct supervision of a physician, or special purpose clinics operated by nurses for things such as immunizations, TB tests, allergies, warts, obesity, etc. During the fall semester of 1970, the Health Service had twelve full-time physicians on its medical staff. Because of duties relating to administration, the inpatient area, the nurse-practitioner clinic, "on-call" periods during the evenings and weekends, and other tasks, only 260 physician hours per week were made available in the outpatient department during regular clinic hours. The rotating schedule meant that no more than seven physicians could be available at one time.

The outpatient department of the Health Service at the University of Massachusetts has many problems in common with other outpatient medical care delivery systems. The rapid growth experienced over the past several years has resulted in conditions common to most overcrowded health care facilities. The alleviation of the following conditions was identified as the immediate target of the study:

1. There was a long waiting time for patients.
2. The professional staff felt overworked and harassed.
3. There was much confusion and crowding in the waiting rooms at predictable times (on Monday, Tuesday, and Friday afternoon).
4. The physicians were still seeing patients as long as an hour past closing time.
5. During the day, physicians were sometimes idle because patients did not always keep appointments scheduled several weeks in advance.

6. The current building was (and is) overcrowded as it was designed for a student body of 10,000 and is currently serving a student body of over 19,000 students.

Procedure

Analysis of the targets of the work revealed that the basic problem was congestion in a complex queueing system. The procedure developed to solve this problem was based on the assumption that improved management of demand, through an expanded appointment system, better resource management, and more efficient physician scheduling, would make the system function more effectively.

The first step was to estimate the "demand" on the system. Specifically, the demand was defined as the number of physician visits per week that would occur during the regular clinic hours in the 1970 academic year. The estimated demand was divided into two components which were termed "controllable" and "uncontrollable". The contollable componentof demand was defined as those patients who made (or could be induced to make) an advance appointment for their physician care. The uncontrollable component of demand, or "walk-in" patients, was defined as those patients who arrived without notice. This latter category would include both "emergencies" and those patients whose need for medical care possibly could be postponed, but was not.

It was then necessary to estimate the number of patients who could be induced to make an appointment for their physician visits by estimating the size of the controllable component of the demand. It would then be possible to distribute the various appointment periods throughout the week in such a way that they would "complement" the walk-in demand. By scheduling more appointment periods during the periods of low walk-in demand, the appointment patients would "smooth" the load of physicians and facilities. Naturally, this distribution of appointments would have to take into account the pattern of arriving walk-in patients, which was known to be different on the various days of the week and which also changed hour by hour during the day.

The attempt to smooth the demand for physicians' services during regular clinic hours proceeded in two steps. The first step was to attempt to smooth the demand by day of the week. It was to be judged successful if there were a uniform number of patients arriving each day of the week. The second step was to smooth the demand across the hours of the day. The criteria of success of this step were the measures of effectiveness of the whole design procedure, the patient waiting time and physician utilization.

These two steps were performed separately. The first step, smoothing the demand over the days of the week, was performed by straightforward analysis of historical trends to produce estimates for the future. The second step of smoothing the demand over the hours of the day required a sophisticated Monte Carlo simulation model.

Alhtough this procedure has intuitive appeal it must still be recognized as a piecemeal attack that omits consideration of all ancillary services (except as their effectiveness may be enhanced by a "smoothed" demand). Considering the present state of the art, this piecemeal approach is the best that can be managed for this type of problem.

Development of the Simulation Model

The present model, schematized in Figure 1 simulates the operation of both the Physician and Nurse Practitioner Clinic.

The model was developed to date in five distinct stages. The first stage was the limited scope model named "ASIS" that simulated a single stage parallel queue (for a variable number of channels) to describe the Physician Clinic. This model had two levels of patient priority. The second model was an extension to a three level priority system for patients with the capability of experimenting with different priority rules and scheduling tactics. The third model was built to represent the Nurse Practitioner Clinic as a two stage, two priority, parallel queue system with a variable number of channels. The fourth model combined the second and third models into one where patients were allowed to cross over from the Nurse Clinic to the Physician Clinic. The fifth and present model is an extension of the fourth. In this model, the servers are allowed to switch functions at predetermined times to limit their patient load to any predetermined class (or classes) of patients. The program is streamlined with the use of modular subroutines to enable ease of modification.

The language used for modeling was a local version of GASP II (4). Since at the outset of the project the University of Massachusetts Research Computer Library contained no debugged and documented simulation language, GASP II was chosen for two reasons: (1) the author's familiarity with Fortran and his understanding of GASP principles; and (2) ease of implemention for the available CDC 3600 and simple modification and expansion in the future.

Since then, several additions and modifications have been made to the GASP II simulation package to satisfy our needs, and this extended version has been found very useful for other projects.

The GASP II (discrete, next event) simulation language contains two groups of Fortran subroutines. One group takes care of the filing and retrieving of simulation events while the other group serves the statistical functions of the simulation, i.e., sampling of distributions, collecting of statistics, etc.

The model builder has only to construct Fortran subroutines using GASP II conventions to model the system, and the execution of the

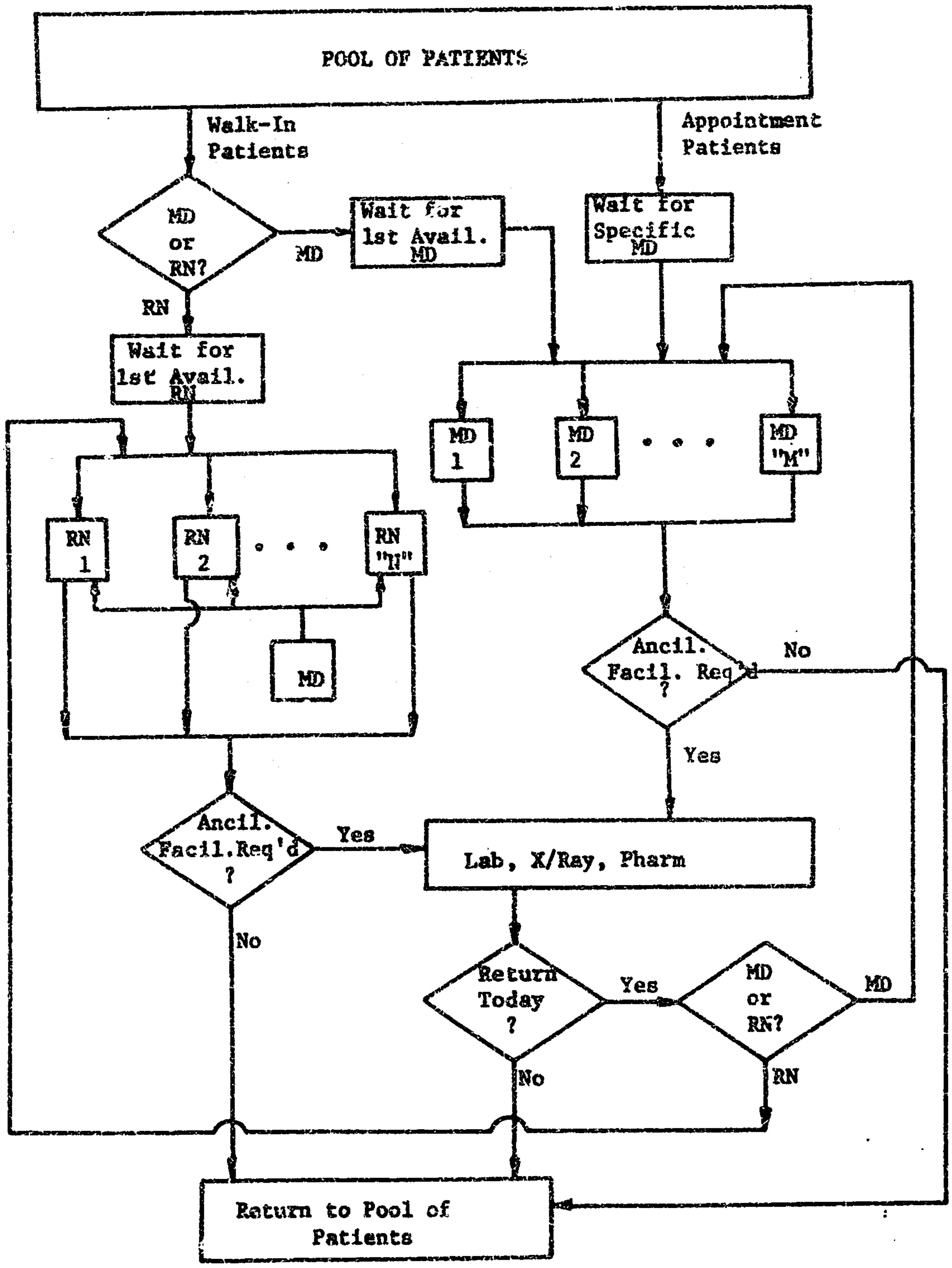


Figure 1. Schematic Patient Flow Diagram Indicating Some of the Logic of the Simulation Program

simulation is then taken over by the GASP II executive package.

Data

Because the system was conceptualized as a complex queueing system, information was needed on the arrival patterns of patients and the way they spent their time in the system. This latter information was broken down into waiting time, the routing of patients through the system, and the amount of time required to serve their needs at each of the places in the system where they received service.

The data used to determine arrival patterns were taken from the encounter form that all patients fill out prior to any service they receive from the Student Health Service. After it is filled out by the arriving student, it is stamped with the date and time and placed with the medical record.

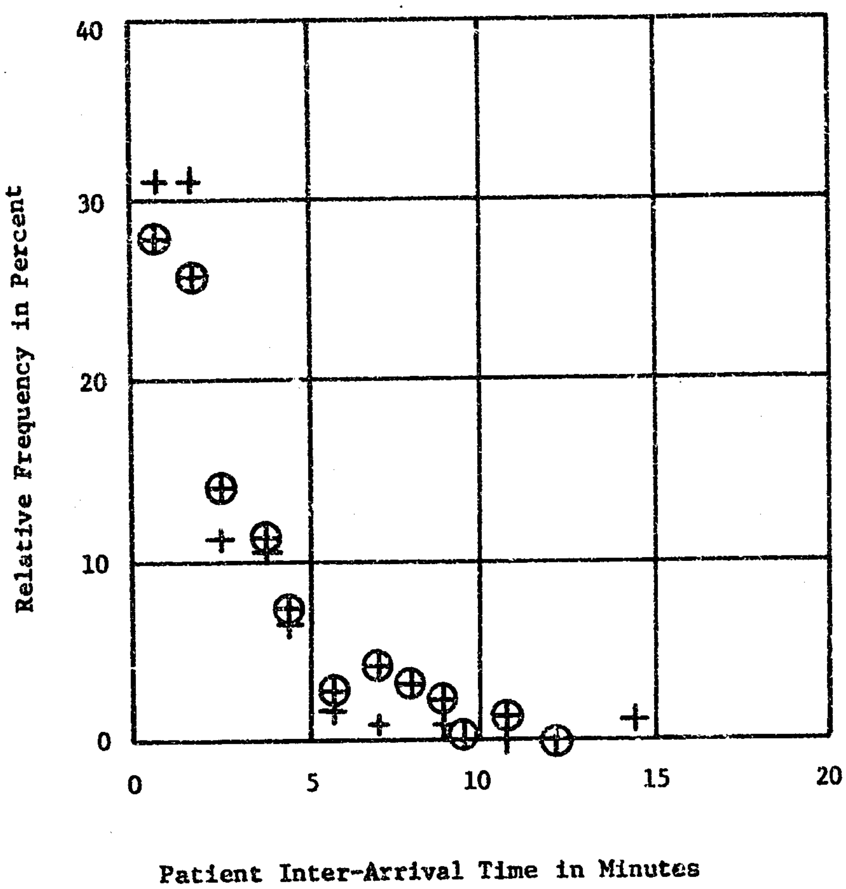
Data on the time physicians spent consulting with patients and the time patients spent in the laboratory, x-ray, etc. were taken during three separate weeks. Clerks were stationed near the entrance to physicians' examining rooms and other facilities and were furnished with date-time stamping clocks. Special record sheets provided for the purpose were stamped as each patient entered and left each service. These records were also time stamped and collected when the patient left the building. The information stamped on these forms gave an accurate account of the services the patient used and of the time necessary to provide the service in question as well as all waiting time involved. During two data-taking periods, the number of these special records collected agreed with the medical encounter forms within about 7 per cent; also, less than 5 per cent of the special forms provided to collect service time and routing data were unusable because of a missing arrival or departure stamp.

Simulation Runs

A word is in order concerning the interpretation of the results obtained from the simulation model. It quickly became clear that Monday, the day with the largest number of walk-in patients, and Thursday, the day with the largest number of appointment patients, were the days that were most sensitive to any scheduling tactics. Therefore, the bulk of the simulation studies were limited to situations found on Monday and Thursday.

The gross patient arrivals over the day (for one Monday and one Thursday) are shown in Figure 2. Figure 2 illustrates that the patient arrivals over the entire day are distributed in a negative exponential form. From this it was assumed that this form of distribution could be used throughout the day to generate walk-in arrivals even though the mean value of the distribution was changed hour by hour to correspond to the observed values.

The arrival rates for each hour of each day were available from the arrival date-time



+ Monday, April 6, 1970; $\bar{x}$ = 2.167, s=2.402, n = 237

+ Thursday, April 9, 1970; $\bar{x}$ = 2.626, s = 2.838, n = 202

Figure 2. Frequency Distribution of Patient Inter-Arrival Time in Minutes

stamp made on all medical encounter forms. This arrival pattern is illustrated in Figure 3 which shows the average number of all patients entering the Health Service to receive care for each hour of the day for Mondays and Thursdays during the fall semesters of 1969 and 1970. The similarity of the pattern between Monday and Thursday data demonstrates that there was little biasing effect of class hours, which tend to be scheduled at the same hours on Monday, Wednesday, and Friday, or at the same hour on Tuesday and Thursday. Although the 1970 data were not available when the analysis was performed, it is presented here to show that stability of the pattern.

The arrival pattern from 1969 was used to generate the walk-in patients for the Monte Carlo simulation model. Operationally, the arrival pattern was incorporated into the simulation model as inter-arrival times, and the parameters of this distribution were changed during each hour of simulated time. By this process, we were reasonably assured that the arrival pattern of the walk-in patients would replicate the pattern of walk-in patients which would actually occur.

The consultation times (service times) that physicians spent with patients were measured in three separate categories. Those categories were for appointment patients, walk-in patients, and the time required for "second service". These distributions are shown by the histograms in Figure 4A, 4B, and 4C.* The sample mean and sample standard deviation for appointment service times were found to be 12.74 minutes and 9.56 minutes and for walk-in service times were found to be 9.61 minutes and 7.48 minutes respectively. The values actually used in the Monte Carlo simulation were generated from a log-normal form and resulted in a distribution whose mean and standard deviation were 12.35 minutes and 9.05 minutes for appointment patients and 9.57 and 8.22 minutes for walk-in patients. These values correspond very closely to the values of 12.6 and 9.8 minutes respectively which were reported for two Air Force Ambulatory care facilities, and which were used by Fetter and Thompson in a portion of their study (5). The Nuffield study of ambulatory care facilities in England also reports similar figures (6); the average consultation time for new patients visiting physicians whose specialties roughly correspond to a practice of Internal Medicine is 11.8 minutes.

The examination of ten days of data showed that approximately fifteen per cent of the patients who see physicians are sent elsewhere in the clinic (e.g., laboratory, x-ray, etc.) and return to see the same physician again on the same day. A log-normal distribution with a mean and variance of 15.54 and 11.09 respectively was found to be an appropriate model for

* Service times in the medical practice are generally best described by either a Gamma or Lognormal distribution. Since generation of variates from the Gamme distribution is time consuming, the Lognormal form was used here.

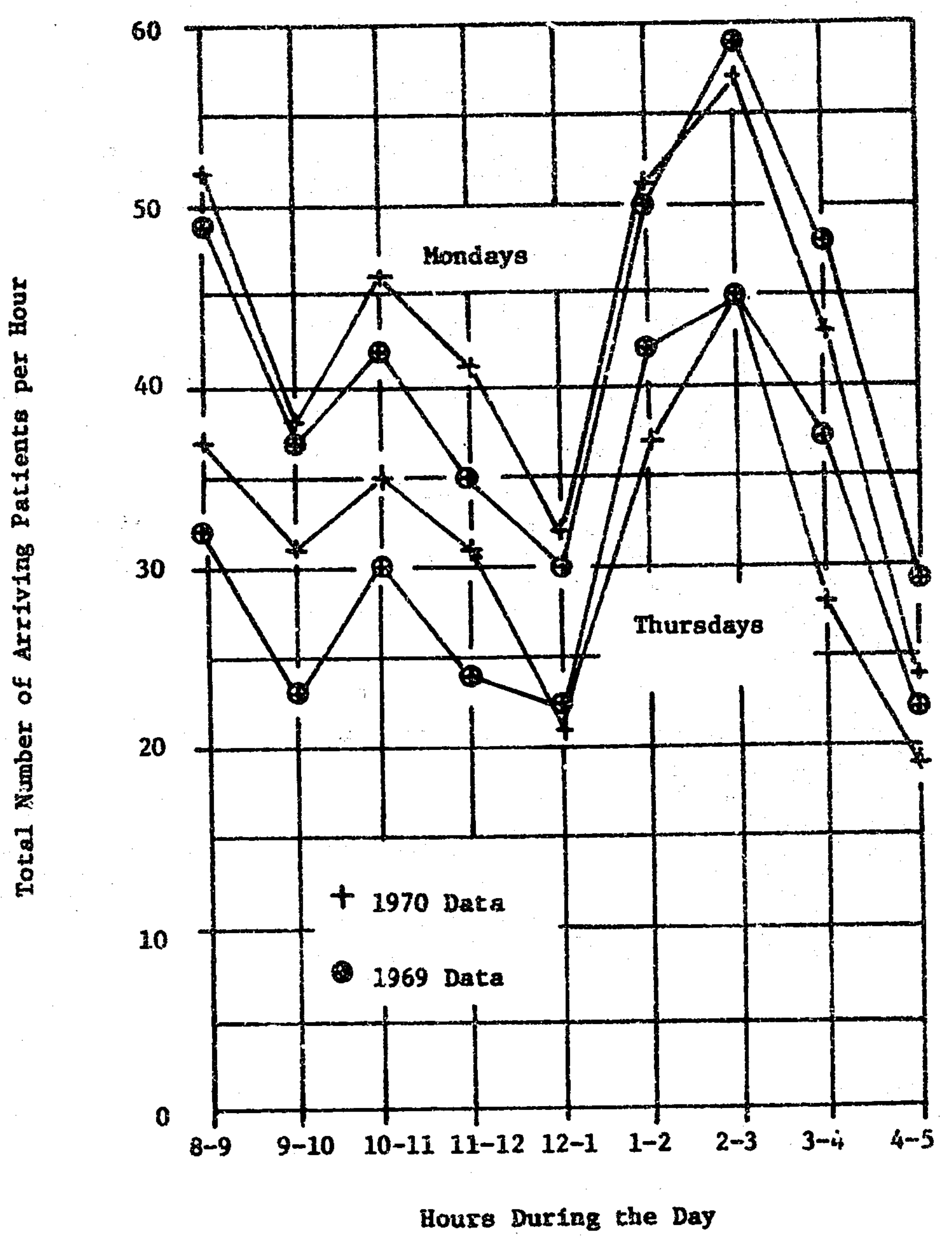


Figure 3. Hourly Arrivals at Student Health Service (Monday and Thursday Averages; Fall Semester 1969 and 1970)

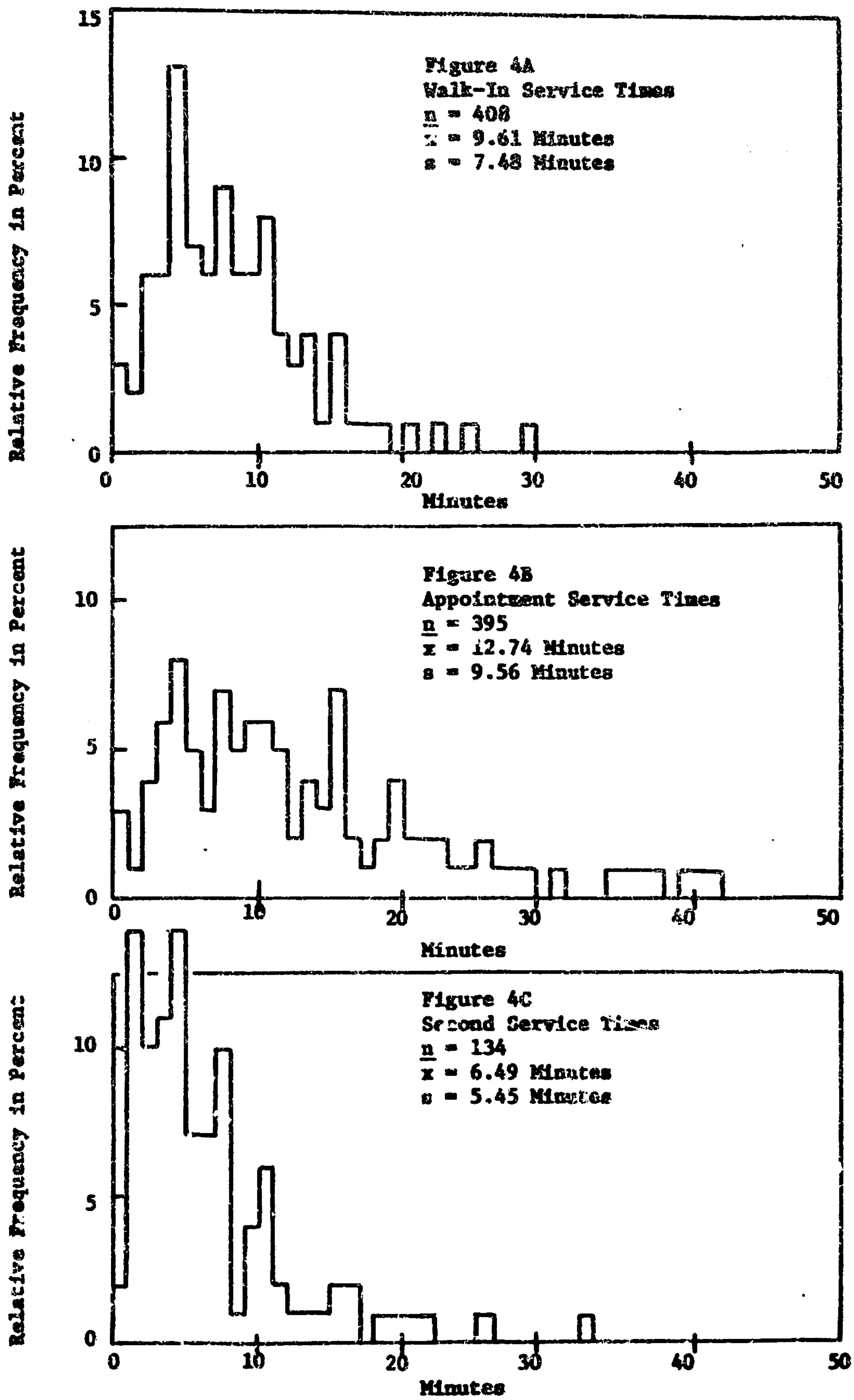


Figure 4. Histograms of Service Times

the elapsed period of time. These patients returning to see the physician were observed to interrupt the flow of new patients. A return visit to a physician seen earlier in the same day was termed 'second service"; the sample mean and standard deviation were 6.49 minutes and 5.45 minutes. The values used in the Monte Carlo simulation were generated from a log-normal form and resulted in a distribution whose mean and standard deviation were 6.41 minutes and 4.91 minutes. No published data were found for comparison purposes.

In the actual operation of the clinic, the physicians see patients in a sequence governed by three priority considerations. First priority is given to emergency patients entering the system and patients who are returning from a visit to the laboratory, x-ray, etc. to see the same physician they have already seen earlier on the same day. Second priority is given to patients who have made an advance appointment with a specific physician. Last priority is given to walk-in patients who are then seen on a first-come, first-served basis. The walk-in patients are seen by any physician as soon as he becomes free of higher priority work. Most physicians use two examining rooms; a patient is being seen in one room while the next patient to see the physician is being prepared in the other. When a physician finishes with one patient, the priority system is used to select a patient for the examining room just vacated. This priority system is administered by a nurse who controls the flow of patients through release of medical records to individual physicians in the proper order.

Within the Monte Carlo simulation model, the priority rules are replicated by the use of two "files" for each physician and one "file" common to all physicians. There is a "priority file" and an "appointment patient file": for each physician, and the "walk-in patient file" is held in common. Each time a physician completes a service to a patient, the files are searched to locate a patient to fill the examining room just used, while the physician sees the patient already waiting in the second examining room. The files are searched in the following order: priority file, appointment file, and walk-in file.

During the operation of the simulation, information was collected on waiting time for appointment patients, walk-in patients, and all patient who undergo "second service". In addition, the simulation collected information on physician utilization and the amount of time beyond the end of regular clinic hours that was required to service all the patients who have arrived during regular clinic hours. These results were displayed in histograms with the mean value, standard deviation, maximum and minimum values.

The effects of two decisions were examined on the basis of the results obtained from the simulation. The first decision was the hours during the day it was best to schedule the physicians, and the second decision involved which of the hours scheduled were to be set aside for appointment patients. The selection of the best schedule of physician hours and the best time for appointment periods proceeded in three steps.

First, an intuitively attractive set of appointment periods was selected that approximately complemeneted the known hourly arrivals of walk-in patients. In the second step, this appointment pattern was held constant and the number of physicians was changed across various hours of the day within daily resource constraints (52 physician-hours, and a maximum of seven physicians at one time). The third step was to hold constant the best physician schedule found in step two, and then to go back and rearrange the appointment slots in an attempt to improve the solution. The second and third steps were repeated in an attempt to secure additional improvements, but none was obtained.

In general, the best physicians' schedule found was that seven physicians should work during the last six hours of the normal eight hour day. In actual practice, this pattern had minor deviations occasioned by the need to stagger the schedules of the physicians to accommodate lunch hours, coffee breaks, and a period for "rounds" in the inpatient area. The clinic was kept open nine hours per day to accommodate the daily eight hour working schedules of the physicians.

Table 1 presents the arrangement of appointment periods by hour of the day and by day of the week that produced the best simulation results when used with the above physician schedules.

The waiting times of appointment patients are relatively insensitive to how the appointment slots are arranged throughout the day because of the priority these patients are given. The waiting time of walk-in patients is highly sensitive to the arrangement of appointment slots through the day and therefore the waiting time of walk-in patients was the most useful criterion to use to make the decisions. It was also found that the number of minutes the clinic runs overtime is sensitive to the pattern of arrivals.

In general, it was found that provided a queue is built up early in the day by either scheduling fewer physicians at the beginning of clinic hours or by scheduling a group of early appointments, the physicians' idle time is relatively insensitive to the way the appointment periods are arranged throughout the day. Proof of this finding was first reported by Welch and Bailey (7) in 1952.

Comparisons between Predictions and Performance

A week or two after the best simulation results were implemented, minor adjustments of some physicians' schedules and appointment patterns were made. After a two month period of operation, data were then taken on service

TABLE 1

Hourly Schedule of Appointment Periods
Available During The Week

Hours	Mon	Tues	Wed	Thurs	Fri	Hourly Totals
8-9	9	7	7	9	7	39
9-10	21	21	21	21	21	105
10-11	13	14	14	15	14	70
11-12	7	7	12	13	12	51
12-1	0	0	0	0	0	0
1-2	10	15	15	17	13	70
2-3	12	12	17	18	18	77
3-4	17	17	13	17	17	87
4-5	7	7	7	6	7	34
Daily Total	96	100	112	116	109	533

times, patient routing, and patient waiting times. The routine management information system based on the encounter form operated constantly, and it yielded arrival data by hour of the day separately for each day.

The real test of the accuracy of the methodology lay in the comparison of predicted and actual outcome measures. However, since the predictions were based on service time data from the previous year, a follow-up study was made to determine if these data were stable from year to year. In this case, examination of service time distributions and patient routing provided this assurance; the follow-up data verified the base data. Furthermore, the evaluation of the control of patient arrivals over the days of the week has been given earlier and rests on the data presented for "smoothed" daily arrivals for the days of the week shown in Figure 5 and on the stability of the pattern of hourly arrivals between 1969 and 1970 shown earlier in Figure 3. Since this evidence was judged satisfactory, the way was cleared to examine the comparison of predicted and actual values of waiting time as a measure of success of the overall methodology.

The two main measures of the model's validity were the patient waiting times and the amount of time necessary for the clinic to remain open to serve the remaining patients after closing time. Initially, the model provided results which were much better than those of the real system. In checking our input data and assumptions, we discovered two discrepancies:

(1) the sum total of tasks measured did not add up to the length of the workday; and

(2) we discovered the system acting somewhat differently when the data takers were visible as compared to a normal day.

The first discrepancy was attributed to the fact that as in all labor-intensive work, allowances have to be made for fatigue and factors beyond operator control and the second was attributed to the classical "Hawthorne" effect which means that data collected with the subject's knowledge tends to show better performance than is the case under normal circumstances.

The allowance made to correct the above was to simulate equally lengthened coffee and lunch breaks for the staff. This allowance produced simulated results where waiting time is compared for five sets of values. The sets of simulation results shown are: first, the case where all physicians arrive at the clinic on time and leave and return promptly from coffee breaks and lunch. In this case, there is no time "lost" from treating patients. The second case is where all physicians lost twenty minutes per day from scheduled clinic duties (an aggregate loss of 140 physician minutes per day); the third case is similar to the second case except the physicians lose 40 minutes per day each from scheduled clinic

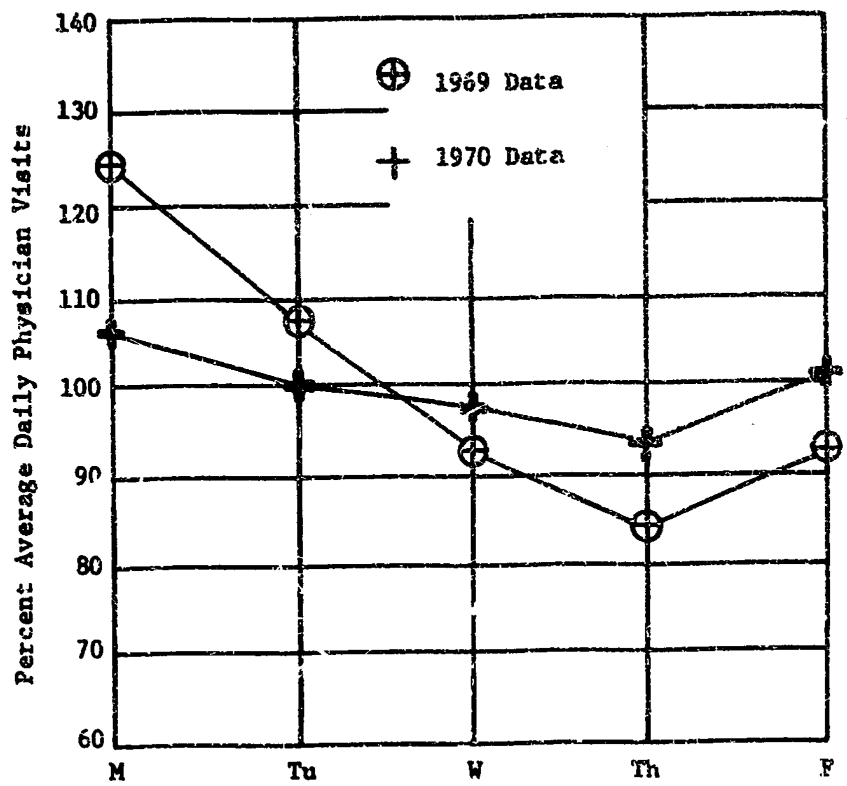


Figure 5. Effect of "smoothing" physician Visits by Day of the Week

duties (an aggregate loss of 280 minutes per day). The fourth and fifth sets of results further increase the loss of physician time to 60 and 80 minutes per physician, which aggregated to 420 and 560 minutes respectively.

In Figure 6, the data from Table 2 have been ploted for Monday and Thursday. This figure shows how the waiting time for appointment patients, walk-in patients and second service changes with the amount of physician "lost" time when the patient load on the clinic remains constant. The data for Monday show that walk-in waiting time is larger than appointment waiting time, and both these values increase for increased values of "lost" time. The Thursday data, the day with the greatest number of appointments, show the same general trends except that walk-in waiting time is less than appointment waiting time for low values of "lost" physician time. This phenomenon appears when the system is not congested and is the effect of physicians using two examining rooms. Use of two examing rooms makes an appointment patient's priority one of "second" in line for a particular physician while a walk-in is taken by any physician as soon as he is free.

Figure 6 shows that the parameters of the distributions of waiting times produced by the model are a function of the physician time that is "lost" to the clinic. The simulated results shown for "no lost time" are an idealized situation that can be used to set lower bounds on waiting times. The simulated values produced for the various amounts of waiting time provide estimates of what is likely to happen to waiting time under these various conditions of "lost" time. It should be pointed out that a "corridor consultation" between two physicians, or an emergency phone call would result in "lost" time to the clinic and therefore would have the same effect on the results obtained from the model as would tardiness and extended coffee breaks.

The entire distribution of simulated values whose mean values agreed best with the actual mean values are shown on the same graph in Figure 7 A, B, and C. Examination of this figure reveals that the simulation model produces distributions of waiting times very close to the actual values. Of particular interest is the fact that all three sets of predicted and actual distributions conform closely across their entire range of values.

The agreement between the form of the predicted and actual values shown in Figure 7, A, B, and C leads to the conclusion that the model behaves in much the same manner as the real world. Efforts are continuing to refine the model to obtain even better predictions and to give deeper insights into the operation of the real system.

It is felt that the next step is to examine the "state sensitivity" of arrival rates

TABLE 2

Simulated Waiting Times for Patients For
Various Amounts of "Lost" Physician Time

		Amount of Lost Time Per Day Per Physician				
		0	20	40	60	80
Walk-In Patients	Mon	13	18	28	34	60
	Thurs	12	12	22	30	51
Appointment Patients	Mon	13	17	20	25	29
	Thurs	12	14	21	26	32
Second Service Patients	Mon	10	10	13	17	20
	Thurs	10	10	15	16	19
Aggregate Lost Time (Min./Day)		0	140	280	420	560

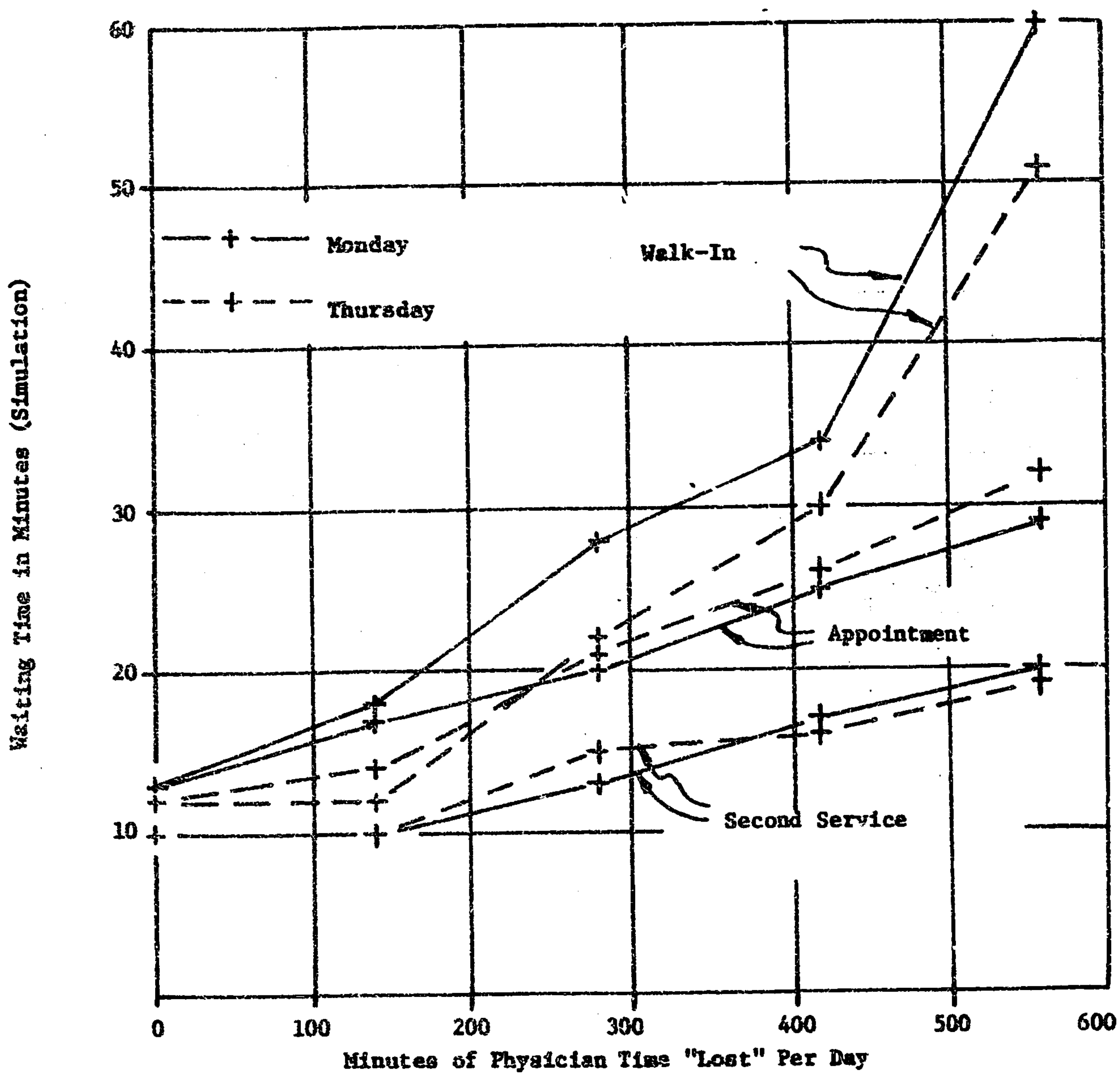


Figure 6. Simulated Waiting Time as a Function of the number of Minutes of Physician Time "Lost" Per Day

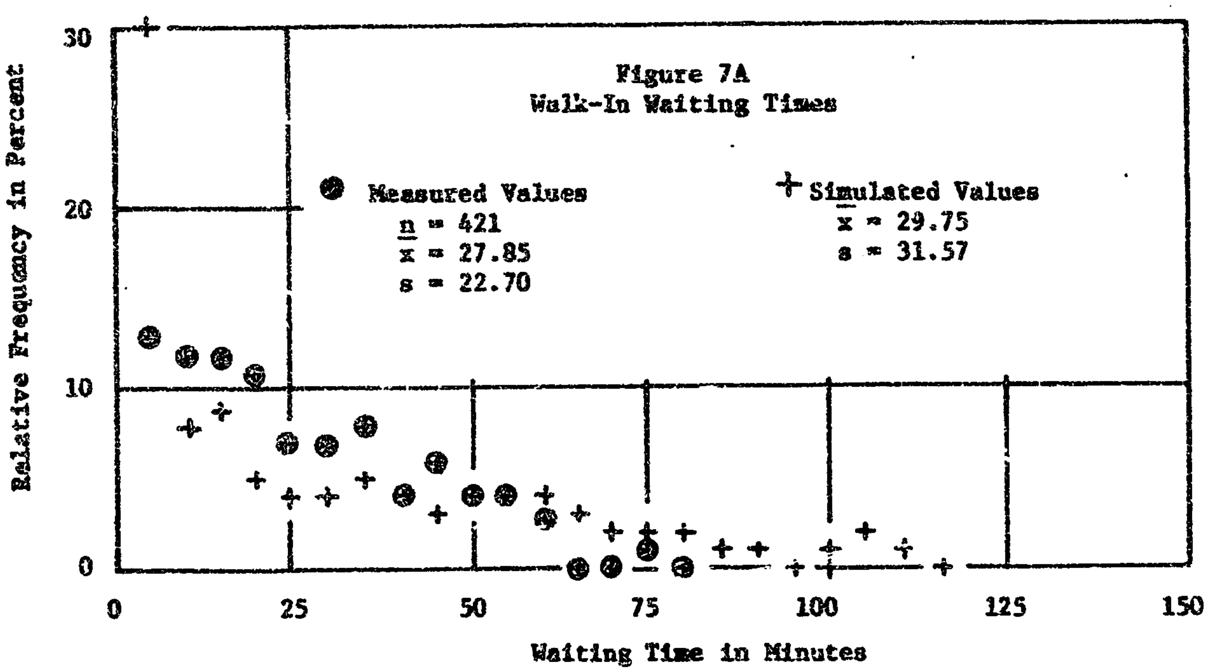


Figure 7. Comparison of Simulated and Actual Waiting Times

7A. Walk-In Waiting Times
7B. Appointment Waiting Times
7C. Second Service Waiting Times

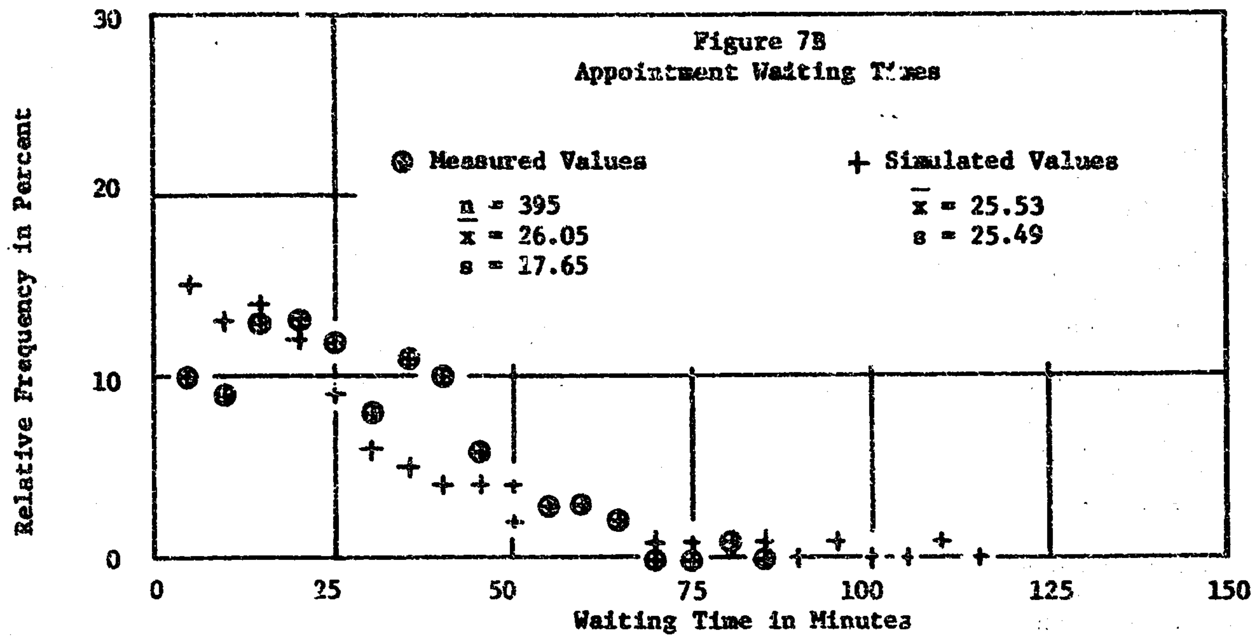


Figure 7B
Appointment Waiting Times
Measured Values
n = 395
x = 26.05
s = 17.65
Simulated Values
x = 25.53
s = 25.49
Relative Frequency in Percent
Waiting Time in Minutes
0
10
20
30
0
25
50
75
100
125
150

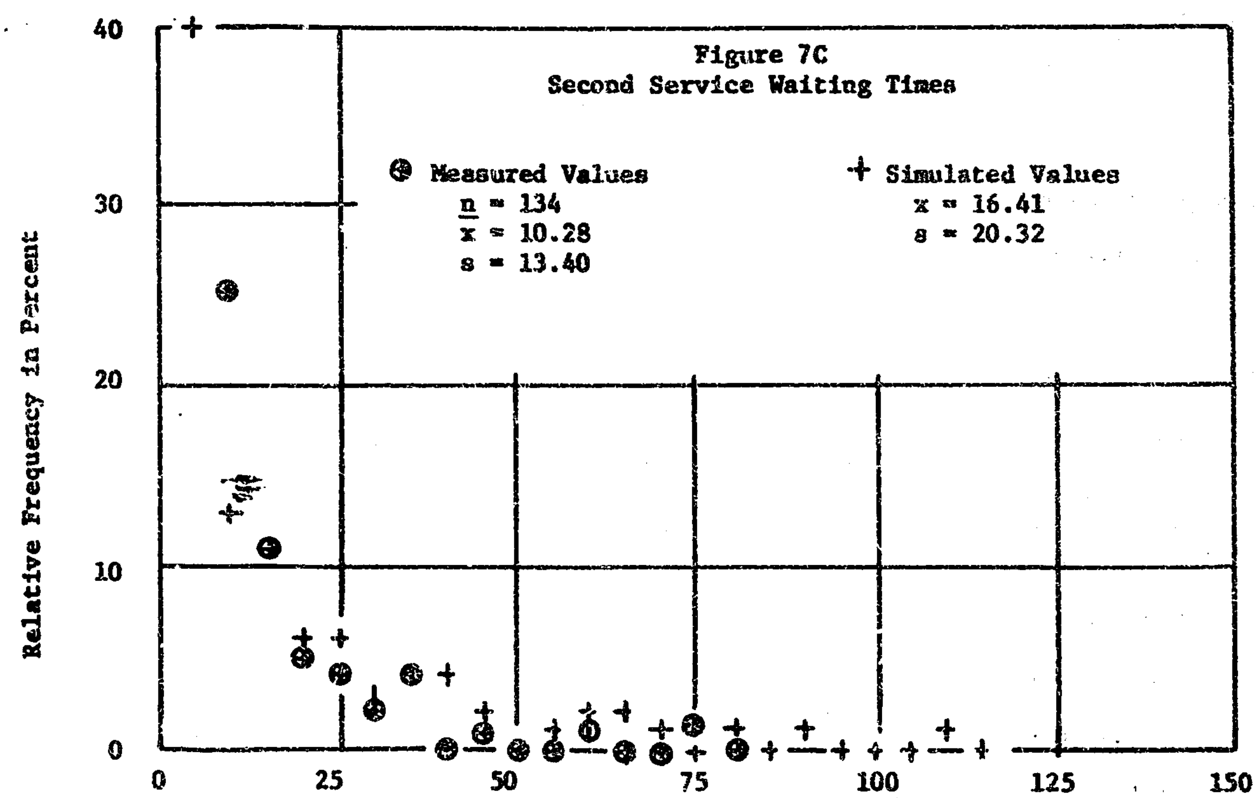


Figure 7C
Second Service Waiting Times
Measured Values
n = 134
x = 10.28
s = 13.40
Simulated Values
x = 16.41
s = 20.32
Relative Frequency in Percent
0
10
20
30
40
0
25
50
75
100
125
150

and service times. It has been observed that when the system is congested some of the arriving patients tend to look around the waiting room and then depart without entering the system. It has also been observed that in the last hour of the day, or when the physicians perceive the waiting room is crowded, they tend to reduce the length of their service time. Both of these cases illustrate that the current "state" of the system appears to have an effect on system parameters. It will be of interest to determine simple and straightforward techniques to measure these effects, apply them in the model, and then determine their effect on the predictive capability of the model.

Results

The data generated by the simulation model were thoroughly analyzed, and it was decided to implement the system changes that were studied. After the changes were accepted and considered routine by the operating staff, data was collected on the actual operation of the system. It was found that the number of patients that were seen by physicians increased by 13.4 per cent with a corresponding decrease of 5.1 per cent in the number of physician hours that were allocated to the walk-in an appointment clinics. In addition, interviews with the physicians confirmed that with the new operating policies, less overtime was required to finish treating those patients who remained in the system after clinic hours were closed. Data was not available to document the physicians' comments due to "student disorders" that were common during the end of the school year.

Another effect of the changes that were implemented was an increase of 5.0 per cent in the overall average time that patients spent with physicians. This increase was due to the increase in the number of patients that were seen by appointment. For both 1969 and 1970, the mean consultation time was 12.7 minutes and 9.6 minutes respectively for appointment and walk-in patients. The increase in throughput of 13.4 per cent and the average increase in service time of 5 per cent were based on a staff level of twelve full-time physicians; and this makes these increases of substantial value.

If these increases had had to be provided under the old system, it would require approximately 2.2 additional physicians. The fact is that the increased service was provided by the same staff and actually used by the student population which increases in size each year. At an average wage of approximately $25,000 for each physician in 1970, this results in a saving in excess of $50,000 the first year in physicians' salaries alone. Substantially more savings can be attributed to this analysis if the salaries of support personnel and equipment charges, hiring costs, fringe benefits, etc. are included.

The waiting time for both walk-in and appointment patients changed from 1969-70 to 1970-71, but the changes occurred in such a way that the overall average waiting time remained the same. The mean waiting time for walk-in patients decreased from approximately 38 minutes to 28 minutes, and the mean waiting time for appointment patients increased from approximately 12 minutes to 26 minutes. Taking into account the increase in the proportion of appointment patients in 1970-71, the weighted average is approximately 27 minutes, which is the same as the weighted average for 1969-70.

In a concurrent study performed by two sociologists, in which the physicians were interviewed both before and after the changes described took place, it was concluded that the physicians' morale improved.

Concluding Remarks

It is widely accepted that simulation modeling is as much an art as it is a science. The builder of a model must combine all the basic modeling elements in such a way that the finished product performs as much like a real world system as possible. In practice, this seldom happens in a direct and straightforward way. Ordinarily, a crude model is constructed which then goes through a series of refinements until the resulting model resembles certain aspects of the real world closely enough to be useful for decision making. Then the model is frequently used to play the "what if" game. The analyst uses the model to investigate what would happen if certain parameters in the real world were deliberately changed or happened to change. For example: How would the system respond if the demand doubled? What would happen if a physician were sick? And so on.

As more and more questions are investigated with the simulation model, adaptations to the basic model have to be made. The analyst finds himself in the position of not having a single simulation model but rather has an entire family of models, many of which may have been patched together in a hurry to be used only once and discarded; others are used over and over. When an analyst finds that one of these adaptations is used several times and more usage is forseen, it may become desirable to spend some time reprogramming the model to add new features and use the opportunity to improve the elegance of the programming.

We have conceptualized this sort of model development on a diagram in Figure 8. As one proceeds horizontally on this diagram, one sees the basic changes that are incorporated into the model to make it resemble reality more closely or comprehend a larger portion of the system. As one moves in the vertical direction on this diagram one may see growth or adaptations of a basic model type, each adaptation being identified with a specific question being asked. It is the authors' experience that a

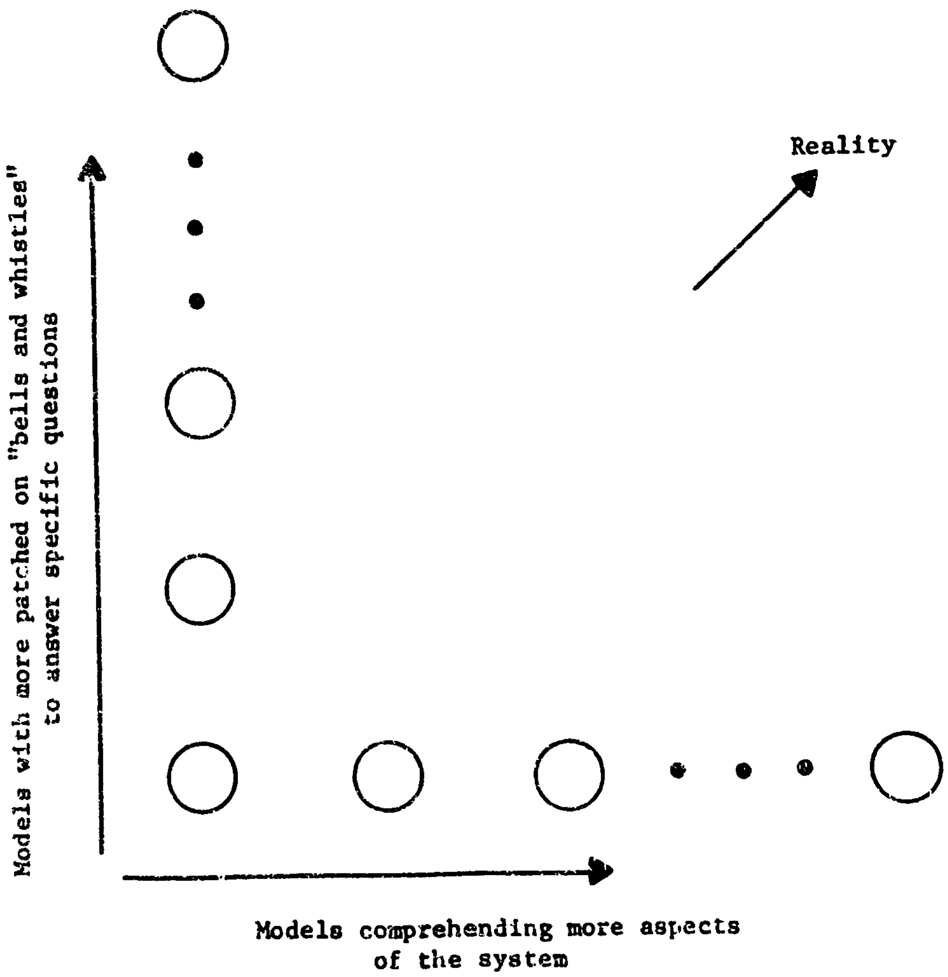


Figure 8. Model Development

basic model type can be expanded in the vertical direction through added on features until it becomes so cumbersome that it is difficult to use efficiently. This may be reflected in such things as unnecessarily elaborate data requirements, extreme running times, and exceeding the capabilities of the available computer configurations. In cases like this, it becomes necessary to invest some time in a programming reorganization to achieve more efficient execution. In this case, a new version of the program would be written which is represented in Figure 8 by a step horizontally.

In practice, when a new question is to be investigated, there is always a conflict between the deisre to accommodate this question with a quick modification or to make a basic change in the model. The quick addition of a new feature to the current version of a program secures the answer to one more new question. A basic change in the model allows this new feature to be included in the model in a way that improves the efficiency as well as the flexibility of the program. It is unnecessary to say that the quick modification route is the one we have usually taken.

The point to be made from this examination of the philosophy of modeling described in Figure 8 deals with the problem of model validation. This diagram shows that the model of any real world situation has many versions, and it is neither economic nor is it possible to validate all the versions by comparing model output to real world data.

Suppose, for example, as in our present case, a comparison of one version of our model with real world data produces results as shown in Figures 7, A, B, and C. What then is the status of the validation of various adaptations of that model? Our view agrees with Naylor (8) that the validation of simulation models is mostly a matter of confidence and faith - not just a matter of mathematics. We believe that mathematics plays its part; it is necessary to compare some basic form of the model with the real world in a careful and systematic way using whatever mathematical tools may be appropriate and available. Once this is done, we feel it is not very fruitful to continue to go over this ground each time the model is adapted to a new problem. Once a model has been validated with real world data, we feel that the intuition of the real world system managers is adequate validation for adaptations.

We feel that the results of this study indicate the methodology presented for demand smoothing and the scheduling of physicians and their appointment patients is successful in this application. The gains in efficiency that we documented were substantial, and they were all in the right direction. We believe additional work will produce further gains. This success, we believe, was due in large part to the predictive capability of the simulation model. The model was "tailor made" to fit this facility.

A generalization emerges from this work when it is considered together with a study of the literature and observation of other outpatient facilities. It is that the methodology described here could be used to good effect on other outpatient facilities. The main obstacle to the widespread use of this approach is the cost and effort necessary to construct realistic simulation models. We believe that since all the published models of outpatient clinics follow the same general form, that is, queueing models solved by Monte Carlo simulation, it would be possible to develop a generalized model that is sufficiently flexible to overcome this obstacle. We are working on this task.

REFERENCES

1. Averill, B.W., Rising, E.J., McBride, T.C., Gage, R. W., and Piedmont, E.B. "The Outpatient Care Delivery System - A New Approach. I. Developing a New System." Journal of the American College Health Association 20: (334-339), June 1972.

2. Rising, E.J., Averill, B.W., and Baron R. "The Outpatient Care Delivery System - A New Approach. II. The Effects of Operations Research." Journal of the American College Health Association 20: (June 1972), 339-344.

3. Baron, R., Rising, E.J., Averill, B.W., and Los, T.J. "The Outpatient Care Delivery System - A New Approach. III. A Computerized Appointment System." Journal of the American College Health Association 20: (June, 1972), 344-350.

4. Pritsker, A.A.B., and Kiviat, P.J. Simulation with GASP II. Prentice-Hall, Inc., 1969.

5. Fetter, R.B. and Thompson, J.D. "Patient Waiting Time and Doctors' Idle Time in the Outpatient Setting." Journal of Health Services Research, Summer 1966.

6. Nuffield Report. Waiting in Hospital Outpatient Departments. Oxford University Press, 1965.

7. Welch, J.D., and Bailey, N.T.J. "Appointment Systems in Hospital Outpatient Departments." Lancet, 1952.

8. Naylor, T.H., and Finger, J.M. "Verification of Computer Simulation Models." Management Science, Vol. 14, No. 2, 1967.