Proceedings of the 1990 Winter Simulation Conference
Osman Balci, Randall P. Sadowski, Richard E. Nance (eds.)

USING DISCRETE EVENT SIMULATION TO EVALUATE HOUSESTAFF WORK SCHEDULES

Robert W. Klein
Mark A. Dame

Regenstrief Institute for Health Care
1001 West Tenth Street
Indianapolis, Indiana 46202

Robert S. Dittus

Regenstrief Institute for Health Care
Indiana University School of Medicine
Department of Medicine
Indianapolis, Indiana 46202

David J. DeBrota

Eli Lilly and Company
Lilly Corporate Center
Indianapolis, Indiana 46285

ABSTRACT

The ability to examine alternatives without great expense or disruption is a major reason why simulation has become the primary tool for analyzing staffing and scheduling problems in industry. Similar techniques could be more widely applied to health care, since a hospital is analogous to a job shop. Data necessary for model construction has generally been unavailable, however, and gathering data has been prohibitively onerous or expensive. Mobile resources such as physicians with complex job descriptions and patients with "time-varying" arrival processes further complicate the modeling task. A doctor's work typically results from a complex combination of scheduled activities and random processes. We have found the INSIGHT simulation language and VISIFIT distribution fitting package to offer the flexibility and power needed for modeling health care activity. We use these tools to model and evaluate alternative housestaff work schedules in a setting where real data is scant, but "expert opinions" abound. Our model accurately predicts the effects on the sleep and activity profile of interns when their schedules are changed such that they are on call every fourth night instead of every third night.

1. INTRODUCTION

Following completion of medical school, physicians devote three to seven years in postgraduate clinical training, referred to as clinical residency training. These years of training have traditionally required physicians to work 80-100 hours per week within a hospital. Such intense work loads have raised concerns over physicians' fatigue and their capacity to deliver high quality medical care in the face of such fatigue. Lawsuits have claimed that tired or fatigued doctors provided inadequate care [McCall 1989]. In New York, the issue became so political that a commission was formed to address residency training [Bell 1988] and its guidelines were enacted into law [New York 1988] with little regard for the financial or educational consequences [Reiner 1989].

The true impact of housestaff working long hours and getting scant or interrupted sleep is not known. Hard evidence that long residency work hours significantly affect the quality of patient care is scarce. Nevertheless, residency directors are concerned that long hours may adversely affect the mental and physical health of the housestaff. In addition, hospital administrators are troubled about the risk of medical liability and the rising liability insurance costs. Also, hospitals risk being unable to recruit the top medical students to their training programs if the training program burdens the housestaff with too little sleep and long working hours. On the other hand, long hours maximize continuity of care and permit housestaff to better learn by observing the serial short-term course of disease processes. Moreover, if reduced hours by housestaff mean additional personnel are needed, the cost of health care could increase substantially.

Even when change has been mandated, many reasons for avoiding schedule revisions can be cited. First, in a hospital, unanticipated side effects of the change may have very serious consequences. In particular, the quality of care may suffer if the time until service is increased, if continuity of care is decreased, or if "over-tired" or "over-busy" physicians perform more of the work. Second, changing schedules can be disruptive to ancillary and support staff since they must always be able to identify the physicians responsible for the care of a given patient. Third, any change instituted may have a large administrative cost. Thus, the need to evaluate alternative schedules before actual implementation is apparent. Computer simulation can satisfy this need and additionally provide flexibility in examining a host of "what-if" questions.

2. SETTING

Wishard Memorial Hospital is an urban midwestern 450-bed public teaching hospital. At any time six medicine "ward teams" provide care to inpatients. Each ward team consists of a second or third year resident, two interns (first year residents), and usually a senior medical student (performing intern duties) collectively referred to as "housestaff". Most direct patient care is provided by interns or senior students. The residents perform initial patient "workups" (history and physical examination, acute stabilization, and admission orders), emergency care, and assist senior students with difficult procedures and order writing. During daily "staffing sessions", a faculty physician meets with each ward team to teach and review patient care. Junior medical students assist with record keeping and technician-type work but do not substitute for physicians, and by requiring instruction, may in some instances add to the housestaff workload.

During a medicine service rotation, prior to June 1989, a typical intern averaged about 97 hours per week at the hospital. Of this time, approximately 68 hours were spent in duties related to inpatients, about four hours at an outpatient clinic, and about four hours in educational conferences. The 21 "idle" hours included meals and approximately nine hours of sleep spread over the 2.3 on-call nights.

Two on-call teams were designated on any given day, one of which was composed of an entire medicine ward team (its resident, both interns, and senior medical student). The other on-call team was made up of both interns from another ward medicine team together with a resident not from a ward medicine team. Two of the six ward teams were thus represented on call (by the presence of their interns) each night, and each intern was on call every third night. Residents and seniors were on call every sixth night.

On-call periods on Monday through Friday begin at 3:00 pm, and on Saturday and Sunday begin at 7:00 am; on-call periods end at 7:00 am the following day. All housestaff not on call on a particular day arrive at the hospital by 7:00 am and leave the hospital as soon after "changeover" (3:30 pm Mon-Fri, 9:30 am Sat-Sun) as they have finished their day's work, typically around 5:00 pm on weekdays and 11:30 am on weekends.

During on-call periods, new patients admitted to the hospital are assigned on a rotating basis among the four interns on call. While on call, interns maintain responsibility for the care of their own patients as well as for approximately one fourth of the patients of interns not on call. On days after interns are on call, they and the rest of their ward team are exempt from being assigned new admissions, but perform all other patient care duties as usual.

During non-call periods (7:00-3:00 pm Mon-Fri), each new patient admitted to the hospital is allocated to one of the four teams not on call the previous night. Each new admission is assigned to the team with the lowest cumulative number of new admissions to date.

In late 1988, it was decided that beginning in mid-1989, the service would change to a system in which interns average call no more frequently than every fourth night. Since the specifics of the new system had not been determined but a change had been mandated, we had the opportunity both to aid in the design of a new schedule and to demonstrate the validity of a simulation model.

3. METHODOLOGY

When creating a simulation model, the first step must be to define its purpose [Pritsker 1989]. Then a suitable simulation language can be chosen and definition of the entities used in that language can be undertaken. Building a preliminary model is generally recommended before a data collection effort is mounted. In our case, the long-term goal of the model is to use it for ranking, selection and analysis of alternative schedules. Thus it is helpful to have a one-to-one correspondence between model elements and system elements. Specifically, the interns, residents and seniors are defined as resources serving a variety of transactions including patient workups, patients' care requests, and other demands on the resources' time such as conferences, staffing sessions, and changeover meetings. The INSIGHT simulation language [Roberts 1983] is used because of its power in modeling flexible decisions by resources in determining which activities to pursue, including the ability of one activity to interrupt another. Its statistical integrity and its strength in providing input process distributions [Roberts and Klein 1989] substantiate the choice of INSIGHT.

To minimize skepticism surrounding the use of simulation modeling in health care, we also attempt the more difficult task of using the model to estimate absolute performance measures. In particular, we want it to predict total time at the hospital, and the times interns and residents spend in various activities, most importantly on-call sleep and patient care. Patient-related statistics such as reasonable lengths of stay and times to receive service are deemed crucial to the model's credibility and are used for verification and validation. Information about continuity and quality of care are also desired.

3.1 Model Formulation

Categorizing patient care requests is fundamental to the model. Care requests by inpatients are classified as one of emergency, urgent, ancillary, non-urgent, or verbal. Each care request is assumed to arrive as a "page" which notifies the physician of the care demand. Verbal care requests can be satisfied almost immediately (e.g. over the telephone); all others require additional physician time. Demands from intensive care (ICU) patients typically have higher priorities than those from ward patients. Additionally, at any time, every task is defined to be either unstarted, overdue, started, or incomplete. An unstarted request may become overdue and thus increase in priority if not serviced within a certain time. When enough of a started task is done, the remainder assumes a lower priority, and is thereafter referred to as incomplete (e.g. when an intern has examined a patient but not yet recorded the exam findings in the chart).

Other demands upon a physician's time include out-patient clinics, admission workups of ICU and ward patients, morning rounds, educational conferences, afternoon "card" rounds, staffing, changeover, background work, teaching junior students, and meals. Table 1 displays the ordered list of preferences an intern uses in choosing his next activity upon completion of some activity. Another list, shown in Table 2, ranks activities by preemption levels. Activities with lower levels can be interrupted due to the advent of a new request with a higher level. Activities of equal level, neither interrupt nor can be interrupted by each other. For instance, a physician can immediately leave a conference (level 4) in response to an emergency (level 10) or leave incomplete non-urgent care (level 2) to begin a conference.

Table 1. Decision Priority List

1. service a page
2. resume servicing started emergency care demand
3. service unstarted emergency care demand
4. go to clinic (while in clinic, physicians are protected from pages and the emergency care demands that pages may spawn)
5. resume servicing started urgent care demand
6. service overdue urgent care demand
7. resume servicing incomplete emergency care demand
8. resume/go to changeover (if ON-CALL and in progress)
9. resume servicing started workup of ICU patient
10. service overdue workup of ICU patient
11. service unstarted urgent care demand
12. service unstarted workup of ICU patient
13. resume/go to staffing (if in progress)
14. resume/go to work rounds (if in progress)
15. resume/go to card rounds (if in progress and resident available)
16. resume servicing incomplete urgent care demand
17. resume servicing patient moving from ICU to ward
18. service patient waiting to move from ICU to ward
19. resume servicing incomplete workup of ICU patient
20. resume servicing started workup of ward patient
21. service overdue workup of ward patient
22. resume/go to conference (resuming conference is guided by additional rules such that one does not resume a conference with only a few minutes remaining)
23. resume/go to changeover (if in progress and NOT on call)
24. resume servicing started ancillary care demand
25. service overdue ancillary care demand
26. resume servicing started non-urgent care demand
27. service overdue non-urgent care demand
28. eat (weekday lunches are available between 11 am and 2 pm for everyone who doesn't go to a noon conference; dinners 5-9, breakfasts 6:30-6:50, and weekend lunches are only for those on-call)
29. service unstarted workup of ward patient
30. resume servicing incomplete workup of ward patient
31. service unstarted ancillary care demand
32. service unstarted non-urgent care demand
33. resume servicing incomplete ancillary care demand
34. resume servicing incomplete non-urgent care demand
35. do background daily chores
36. go home (if after 3:30 and not on-call)
37. sleep (if between 10:30 pm and 6:30 am)
38. go idle

As specified in the INSIGHT simulation language, the model consists of many parallel networks with the resources (physicians) free to move among activities in most of them. Each physician activity not directly related to patient care has its own network. The largest and central network models the patients who arrive, are worked up, stay a period of time on the ICU and/or ward, and eventually are discharged or die. At random times after patients have been worked up they may require attention from a physician. In addition, the activity times for routine care activities in other networks, such as staffing and background work, may depend on the number of a given physician's patients (transactions) in the central patient network. There are a total of 24 queue-activity node pairs where transactions may wait and then capture resources. Eight other activity nodes help to regulate transaction movement. Including attributes needed to collect specific statistics into the 64 tables, 112 transaction and network attributes are defined and maintained. Fifty decision nodes are used to model the process defined in Table 1 and the similar processes for residents and senior students.

The arrivals and departures of patients, their transitions between the ward and the ICU, and the occurrences and urgencies of their care requests are all determined randomly according to probability distributions. The transactions in many of the other networks initiate scheduled activities such as clinics or conferences, so their arrival times may be deterministic. Thus, it is helpful that INSIGHT can read and process scheduled events from an external file and intersperse them with the stochastic events which arise during the simulation.

A pilot model was built using triangular service and exponential patient arrival time distributions, based only on some readily available data, such as an average patient census of 90 (13% ICU), 91 admissions/week, an average length of stay of 6.5 days, and "expert" opinion. A total of 41 distributions were specified. Preliminary analysis revealed the areas requiring more extensive data collection or better estimates.

3.2 Data Collection

Efforts were undertaken to characterize the frequency and duration of types of work and the dynamics of the medicine patient population. A variety of sources (log books, admission records, departmental summaries) were scrutinized to delineate the nature of the patient arrival and departure processes and the transitions of patients between ICU and ward units. Patient arrival rates were estimated from available records of admission times during representative time periods.

Table 2. Preemption Levels

```
------- 20 -------
remain at clinic
remain at home
service a page
------- 10 -------
continue servicing started emergency care demand
service unstarted emergency care demand
move patient from ICU to ward when ICU is full
------- 9 -------
go to clinic
------- 8 -------
continue servicing started urgent care demand
service overdue urgent care demand
remain at changeover (if ON CALL)
------- 7 -------
continue servicing incomplete emergency care demand
go to changeover (if ON CALL)
continue servicing started workup of ICU patient
service overdue workup of ICU patient
service unstarted urgent care demand
service unstarted workup of ICU patient
------- 6 -------
go to / remain at staffing
go to / remain at work rounds
go to / remain at card rounds
------- 5 -------
continue servicing incomplete urgent care demand
service patient to be moved from ICU to ward
continue servicing incomplete workup of ICU patient
continue servicing started workup of ward patient
service overdue workup of ward patient
------- 4 -------
go to / remain at conference (resuming conference is
   guided by separate rules such that one does not
   resume a conference with only a few minutes
   remaining)
go to / remain at changeover if in progress (if NOT on
   call)
continue servicing started ancillary care demand
service overdue ancillary care demand
continue servicing started non-urgent care demand
service overdue non-urgent care demand
------- 3 -------
eat a meal
service unstarted workup of ward patient
------- 2 -------
continue servicing incomplete workup of ward patient
service unstarted non-urgent or ancillary care demand
continue servicing incomplete non-urgent or ancillary
care demand
------- 1 -------
do background daily chores
go home (if NOT on call)
sleep
```

Except for scheduled activities such as conferences and workrounds, no measures of physician activity times were readily available. Consequently, a linear opinion pool [Berger 1980] Delphi-assisted poll of six faculty familiar with the medicine service, and a survey of the 18 housestaff on the medicine service rotation were used to elicit minimum, maximum, and most likely times for all activities modeled.

3.3 Input Modeling

These activity time estimates were then used in conjunction with VISIFIT [DeBrota et al. 1989a, b] to fit Johnson S_B distributions which generate realistic service times. The Johnson system of distributions [Johnson 1949] is a four-parameter family useful for modeling distributions containing any skewness and kurtosis. Figure 1 shows two Johnson S_B distributions overlaid with a triangular, all with the same minimum, maximum and mode. For right skewed distributions, typical of service times, a triangular distribution usually overestimates the true mean [Klein and Baris 1990]. We concur with Wilson et al. [1982] that using a "beta-pert" estimate of the standard deviation as one-sixth of the mean (a procedure readily available in VISIFIT), produces an intuitively appealing and apparently valid representation of most such distributions. Therefore, this procedure was used to parameterize most distributions in the model. The more peaked Johnson S_B in Figure 1 depicts such a distribution, while the other Johnson S_B has the same variance as the triangular.

The survey also gathered information on the daily number of care requests in each category per patient, as well as information used for verifying and fine-tuning the model, such as the times housestaff go home and the periods of sleep obtained on call nights. Piecewise-linear time-varying non-homogeneous Poisson process generators [Klein and Roberts 1984] were parameterized with the historical average arrival rates of ICU and ward patients for various relevant times of day on both weekdays and weekends.

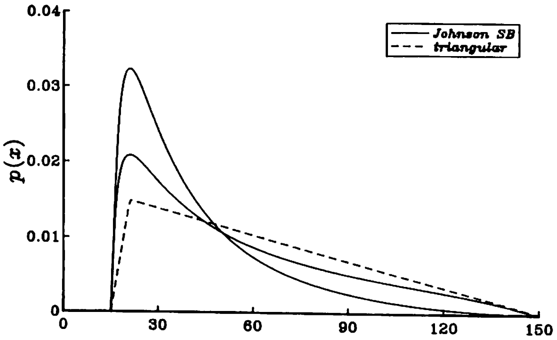

Figure 1. A Triangular and Two Johnson S_B Distributions with same Mode and Range

Figure 2 shows the rate function used to model ICU patient arrivals. The resulting time-varying Poisson distribution avoids the extensive data processing and the unnatural discontinuities inherent in the more common method of using a different Poisson process for a large number of time periods (e.g. every hour). Note also that the zero rates depicted are real, because arrivals near the end of shifts remain in the emergency room rather than being immediately assigned to a medicine team.

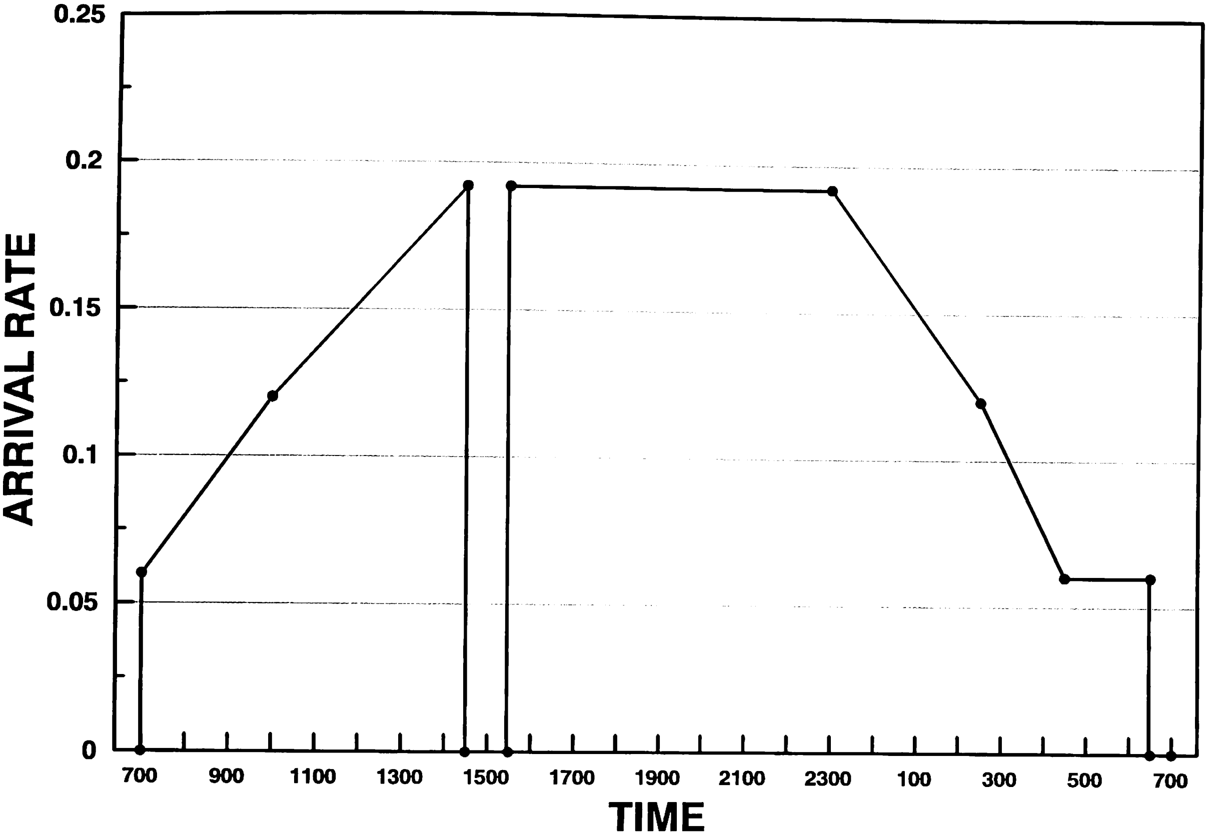

Figure 2. Time-Varying ICU Patient Arrival Rate Function

3.4 The Experimental Frame

Multiple six-week periods of simulation are used to collect statistics on individual and aggregated housestaff and on patients. Each six-week period during which statistics are gathered is preceded by a two-week simulation period without statistics collection in order to mitigate any initialization bias. Confidence intervals produced from five such runs are used to verify that measures not used in initial model construction are consistent with reality. Patient length of stay and times housestaff left the hospital are then used in fine-tuning patient transition probabilities and housestaff priority rules. Once the model adequately represents the medicine service, we can begin to consider the proposed changes.

Among the outcome measures available from INSIGHT tables are: (1) total and maximum uninterrupted sleep obtained by housestaff while on call, (2) percentages of times housestaff spend in various activities, (3) attendance by housestaff at educational conferences, (4) time spent by residents teaching junior students, (5) total time housestaff spend at the hospital, (6) the times until various types of patient care delivery commence, and (7) measures of patients' continuity of care.

3.5 Validation

Since all output measures gathered for the every third night call system were used in the verification of the baseline model, the validity of its predictions remained to be established. Before the process of validation could be completed, however, an administrative decree mandated that an every fourth night call system be adopted immediately unless significant adverse outcomes could be foreseen. By changing primarily the model inputs relating to the frequency of call, a modification of the baseline model in which interns took call every fourth night was created. Although results indicated a slight increase in the percentage of care provided by "tired" interns and an average decrease of about one hour of sleep per intern night on call, the new system was judged acceptable and was adopted in June, 1989.

The implementation of the new schedule created a research platform for examining the validity of our modified model. Thus, we conducted a work measurement study of interns and residents. Senior medicine students were hired and trained to follow a physician and record each time an activity was begun or ended. In a pilot study, three students followed one intern for twelve hours to assess interobserver variability. Two hundred random minutes were sampled, and the categorizations of intern activity on the three record sheets were identical on 186 (93%). After a clarification of the definition of what constituted a workup, agreement exceeded 96%. A total of twelve intern days, six resident days, four intern nights, and two resident nights were subsequently observed. For each physician day, a simulation matching number and type of patients, patient arrivals, clinic and conference schedules was built and run for 50 replications. From the simulations, means and variances for the percentage of time in each category of activity were calculated. Using these statistics, confidence intervals were constructed and treated as control limits for the actual observations. A similar procedure was followed for the aggregate of all physician days. When aggregated to represent single days, all observed data fit within 95% confidence intervals generated by the simulation.

We conclude that the model can be used to predict the effects of future schedule changes on the housestaff's activity profile. See [Carson 1986, 1989; and Banks 1989] for further discussions of validation.

4. RESULTS

Table 3 combines output from the baseline (every third night call system) model with output from the model of the new system. This displays the differences in the activity profiles of interns under the two systems. It is apparent that while average time at the hospital is reduced by 9.5 hours per week, total time working is only reduced by about 3 hours. Since no new resources are added, any change in work done must be attained by shifting work from interns to residents or seniors, or by missing desirable activities such as educational conferences. Similar statistics for residents and seniors allow the specifics of work shifting to be predicted.

Table 3. Intern Activity Profile

ACTIVITY	Old hr./week	New hr./week	Change
Workups	12.4	11.6	-0.8
Ward patient care	12.7	11.7	-1.0
ICU patient care	9.7	9.5	-0.2
Background work	11.2	10.3	-0.9
Answering pages	1.6	1.5	-0.1
Workrounds	6.7	6.4	-0.3
Card rounds	1.8	1.8	0.0
Staffing	10.0	10.2	+0.2
Changeover	1.7	1.5	-0.2
Clinic	4.0	4.0	0.0
Conferences	3.9	3.9	0.0

Total working	75.5	72.4	-3.1
Meals at hospital	2.7	2.2	-0.5
Idle at hospital	19.0	13.0	-6.0

Total at hospital	97.1	87.6	-9.5

As would be expected, the every fourth night call schedule results in decreased average nightly sleep when on call, since one less intern is available to perform each night's work. Another vital statistic is the amount of "important" care provided by "tired" physicians. The simulation model produces enough data to evaluate systems using a variety of complex criteria to define "tired". For instance, we are able to find the percentage of urgent or emergency care requests begun by a physician who has been at the hospital at least 15 hours, with less than four hours of total sleep and a maximum uninterrupted sleep period of less than two hours since arrival. The change from every third to every fourth night call increases the number of such services by approximately 50%.

5. CONCLUSIONS

A discrete-event simulation model of work performed by hospital interns and residents was built despite limited data and many complicated modeling issues. The model incorporates such complex features as interruptions of one activity by another, involved decision processes by mobile resources, and highly variable work (patient) arrival processes. Where historical data were lacking and too expensive to collect, surveys of experts produced sufficient information to use VISIFIT to specify input processes. Validation efforts showed that the model accurately predicted the results of changes to the schedule.

With its validity established, the model has since been used to investigate the ramifications of alternative systems including the addition of a night-float team and more help from subspecialty fellows. The consequences of changes in patient load or decreased availability of senior students were also examined. The prospect for regulatory or legislative restrictions on housestaff hours makes the model a valuable tool for evaluating scheduling rules and plans.

ACKNOWLEDGEMENTS

The authors wish to thank Dr. John Fitzgerald for his help in specifying the initial model, and Dr. Stephen Roberts for his suggestions on the content and organization of this paper.

REFERENCES

Banks, J. (1989), "Testing, Understanding, and Validating Complex Simulation Models," In *Proceedings of the 1989 Winter Simulation Conference,* E.A. Macnair, K.J. Musselman, and P. Heidelberger Eds. IEEE, Piscataway, NJ, 549-551.

Bell, B.M. (1988), "The New Hospital Code and the Supervision of Residents," *NY State Journal of Medicine 88*, 617-619.

Berger, J.O. (1980), *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag, New York, NY, 273.

Carson, J.S. (1986), "Convincing Users of Model's Validity is Challenging Aspect of Modeler's Job," *Industrial Engineering 18*, 6, 74-85.

Carson, J.S. (1989), "Verification and Validation: A Consultant's Perspective," In *Proceedings of the 1989 Winter Simulation Conference,* E.A. Macnair, K.J. Musselman, and P. Heidelberger Eds. IEEE, Piscataway, NJ, 552-557.

DeBrota, D.J., R.S. Dittus, J.J. Swain, S.D. Roberts, J.R. Wilson, and S. Venkatramen (1989), "Modeling Input Processes with Johnson Distributions," In *Proceedings of the 1989 Winter Simulation Conference,* E.A. Macnair, K.J. Musselman, and P. Heidelberger Eds. IEEE, Piscataway, NJ, 308-317.

DeBrota, D.J., R.S. Dittus, S.D. Roberts, and J.R. Wilson, (1989), "Visual Interactive Fitting of bounded Johnson distributions," *Simulation 52*, 199-205.

Johnson, N.L. (1949), "Systems of frequency curves generated by methods of translation," *Biometrika 36*, 149-176.

Klein, R.W. and P.M. Baris (1990), "Selecting and Generating Variates for Modeling Service Times," *Computers and Industrial Engineering* (In press).

Klein, R.W. and S.D. Roberts (1984), "A time-varying Poisson process generator," *Simulation 43*, 193-195.

McCall, T.B. (1989), "No Turning Back: A Blueprint for Residency Reform," *Journal of the American Medical Association 261*, 909-910.

New York State Minimum Standards for Hospitals, new part 405 (1988), Attachment prepared for executive briefing session on house staff credentialing, privileging, and scheduling, Greater New York Hospital Association.

Pritsker, A.A.B. (1989), "Why simulation works," In *Proceedings of the 1989 Winter Simulation Conference,* E.A. Macnair, K.J. Musselman, and P. Heidelberger Eds. IEEE, Piscataway, NJ, 1-6.

Reiner, S.C. (1989), "The Impact of the New Code Regulations on Postgraduate Medical Education in New York State," *New York State Journal of Medicine 89*, 457-461.

Roberts, S.D. (1983), *Simulation Modeling and Analysis with INSIGHT*, Regenstrief Institute, Indianapolis, IN.

Roberts, S.D. and R.W. Klein (1989), "Statistical Issues in a General-purpose Simulation Modeling Language," In *Proceedings of the 1989 Winter Simulation Conference,* E.A. Macnair, K.J. Musselman, and P. Heidelberger Eds. IEEE, Piscataway, NJ, 325-333.

Wilson, J.R., D.K. Vaughan, E. Naylor, and R.G. Voss (1982), "Analysis of Space Shuttle ground operations," *Simulation 38*, 187-203.