Proceedings of the 2005 Winter Simulation Conference
M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, eds.

A SURVEY OF DATA RESOURCES FOR SIMULATING PATIENT FLOWS IN HEALTHCARE DELIVERY SYSTEMS

K. Preston White, Jr.

Department of Systems and Information Engineering
P.O. Box 400747
University of Virginia
Charlottesville, VA 22904-4747, U.S.A.

ABSTRACT

Modeling and simulation studies of patient flows in healthcare systems have been reported consistently in these *Proceedings* for over a decade. Our ongoing research in this area is motivated by our desire to develop models which will illuminate the causes and remedies for repeated area-wide ambulance diversions experienced in a metropolitan hospital system. In this paper we summarize our background research on the sources of data available to calibrate patient-flow simulation models, including time series for patient admission, discharge, diagnoses, length-of-stay, and inpatient census for emergency departments and hospitals. Specifically, we review the input analyses reported for various prior simulation studies, including data capture and technical difficulties in reducing data for model calibration. We also suggest alternative sources of data that could prove especially useful in simulation studies of mass ambulance diversions, as well as heavy, area-wide patient loads that might be associated with emergency responses.

1 BACKGROUND

When an emergency department (ED) reaches capacity, non-trauma cases inbound via ambulance or rescue squad typically are *diverted* to hospital with remaining capacity. On January 9, 2001, and again on January 8, 2003, each of the eleven hospital ED's in the Richmond, VA, metropolitan area reached maximum capacity in rapid succession. With the entire metropolitan emergency medical system (EMS) saturated, diversions were made to hospitals outside of the immediate area, at least an hour away by ground transport. Clearly, the waiting time to receive medical attention during a mass diversion may be unduly lengthy, whether the patient is admitted to an overcrowded facility, or transported to a distant care facility.

Following the first area-wide diversion, a systems study was undertaken (Cohort, 2002) to assess causes and potential remedies. A high-level, discrete-event simulation (DES) of patient flows, tracking patients from the receipt of a 911 call through release from one of the eleven area ED's, was developed and analyzed. Results indicated that improving staffing utilization, reducing the number of non-critical patients entering the ED, and reducing the time a patient spends in the ED would provide significant benefits.

The DES results also appeared to confirm abundant anecdotal evidence that diversions were *not* the result of ED resources inadequate to the volume of emergency care cases, *per se*. Instead, "Emergency department stays are often prolonged when patients await hospital admission. It is not uncommon for a significant percent of the total ED stay to include waiting for a hospital bed after the decision to admit to the hospital has been made. The waiting patient has already received ED treatment and is unnecessarily occupying sorely needed staff and bed resources. A recurrent cause of this delay is the unavailability of receiving unit beds." Additionally, "A significant percent of [hospital] inpatient beds are occupied by patients who have been cleared for hospital discharge but are awaiting family or friends for transport."

In order to test, refine, and expand on the many recommendations provided by the Richmond diversion study, we undertook to enhance the core simulation model by providing additional details regarding patient flows within the metropolitan system, particularly among key units in the receiving hospitals. Data sources for the original study—a general report on bed utilization, logs from the Richmond Ambulance Authority and one participating ED, and extensive interviews with appropriate hospital and EMS personnel—could not provide the requisite detail on patient flows needed for the enhancement. It was anticipated, therefore, that the acquisition and reduction of the data necessary to calibrate the area-wide model would represent a formidable challenge.

Accordingly, we undertook a literature survey to understand the sources of data that had been used in prior patient-flow simulations. Partial results of that survey are presented in Section II. Not surprisingly, we found that all (save one) of the prior studies had obtained data directly from the subject care facilities, either from existing data repositories maintained by the facility, or from necessarily limited direct observation of facility operations, or from interviews with facility personnel, or in most cases some combination of these sources. In Section III, we describe an alternative source of U.S. data—the Health Care Cost and Utilization Project—that has proven useful in modeling patient flows in metropolitan region.

2 LITERATURE REVIEW

Jun, et al. (1999) survey the use of DES in healthcare systems, listing over one hundred studies to that date. In the present paper we survey health-care DES studies published in these *Proceedings* from 1997 through 2004, available at <http://www.wintersim.org/pastprog.htm>. For each of the thirty-five WSC papers cited, we describe the simulation application and objectives; then discuss the data issues as reported in each.

2.1 General

Pitt (1997) describes a generalized simulation system developed in liaison with health authorities in West Yorkshire, UK, and aimed at addressing the growing demand for healthcare services. To ease model specification, a baseline default setting was established for each of the input variables (including demographic controls, demand fluctuations, admissions, hospital ward configuration, length of stay (LOS), and day-case percentages), set according to historical healthcare data supplied by Calderdale Healthcare Trust. A case study illustrates the use of this generalized system to conduct experiments and validate results against real healthcare data.

In Sanchez, et al. (2000), Sepulveda speaks specifically of lessons learned regarding data resources for health care simulation, noting that in his experience data availability for healthcare facilities are often "either non-existent, or excessive. In some (rare) instances, notably emergency departments, the information is electronically stored and it is relatively easy to retrieve from existing databases just about anything that may be needed for a given patient"—arrival times, transportation mode, initial assessment, vital signs, age, gender, triage evaluation, diagnosis, medications given, disposition, laboratory/radiology results, consultation times, attending personnel, patient disposition, and insurance coverage.

More commonly, however: (i) important information is stored, but can not be used directly in modeling (e.g., available arrival and departure times yield the total time in a unit, but do not reveal how much of this time was spent in value-added activities (and should be included as distributions and/or parameters of the model) and how much was spent in non-value added steps (and should be a result of the interactions of entities and resources in the model); (ii) available data is abundant but it is stored on a non-electronic form (e.g., patient's files); and/or (iii) data does not include crucial elements (e.g., procedure start and ending times). "It thus becomes necessary to take and analyze sample data. It is a common assertion in the literature that data and distributions can be easily estimated by asking an expert for educated estimates. The old adage, garbage in-garbage out, is still very valid. A good sample is better than a thousand experts' opinions. It is our experience that small variations in some distribution parameters, or (worse) a change in the distribution itself, usually leads to significant changes in model results, e.g., may lead to different conclusions."

In the context of describing the practical realities of DES, Centeno and Carrillo (2001) provide a thoughtful review of several issues related to data resources, with emphasis on their experiences in health care applications. They note that organizations are continuously collecting data, but that elicitation of useful information may represent a major bottleneck, because the information system collecting such data typically is not designed for stochastic modeling. They point to (i) difficulties associated with the large number of raw data records that must be extracted, filtered to ensure quality, and then transformed into the specific information required to calibrate simulation; (ii) the pervasive need to explain high variability within these data sets; and (iii) the lack of data on some activities altogether. With respect to variability, specifically, they cite a study of hospital patient flows in which they identified that special and neurology procedures do not occur in a steady fashion and had to be treated as special cases that occur with a certain percentage. Further, that found that the data varied significantly due to the "subjective nature of the procedures"—some physicians take longer than others and may have different methods of performing the same procedures.

Eldabi and Paul (2001) propound a modeling and simulation approach to enhance understanding and intercommunications among the many and varied stakeholders for healthcare systems. However, data sources and issues are not reported.

2.2 Emergency Departments (ED's)

Jenkins, et al. (1998) use a simulation study of the Emergency Ward of Campbelltown Public Hospital in Australia to illustrate model verification and validation concepts. Data issues and sources are not reported.

Rosetti, et al. (1999) and Trzcinski (1999) describe the application of DES to inform decisions concerning the efficient allocation and use of ED staff resources at the University of Virginia Medical Center in Charlottesville, VA. The study objectives were to test alternative attending-physician staffing schedules; analyze the corresponding impacts on patient throughput and resource utilization; help identify process inefficiencies; and evaluate the effects of staffing, layout, resource, and patient flow changes on system performance, without disturbing the actual system. The authors provide an unusual level of detail on the nature and range of the mix of data resources used to calibrate their model.

"While the majority of information was already being collected and stored by the ED's computerized patient tracking system, the radiology department's computer database, and the various lab computerized databases, these systems were only partially used to obtain input data for this study." The data collection effort was separated into four different phases: (i) Detailed information about the visit-time delay each stage (i.e. registration, triage, etc.) of a patient's flow through the ED was gathered on all ED patients for the one week, using self-reported work sampling techniques. A total of 1,175 patient visit data sheets were completed over this period. (ii) Information on service-times that ED doctors, nurses, and nurse practitioners spent for various patient care activities was compiled, using a time study restricted to ED patient visits in specific areas during the same week between the hours of 12 p.m. and 8 p.m. The sample data set consisted of 115 complete patient visits (and 30 partially completed visits, which were omitted). (iii) To determine appropriate arrival rates for each of the three patient arrival processes (walk-in, ambulance, and helicopter), information was extracted from a computerized patient tracking system database, using a customized query to sort the information into the necessary arrival types by hour and day of the week. A total of 17 weeks (November-February) were analyzed to determine each average arrival rate by hour and day. This data was modeled as a non-stationary arrival process for generating patient entities. (iv) Estimates on transport and routing times for patients and caregivers between various arrival stations and ED areas and between different areas within the ED were determined from based on approximated distances, divided by a random walking velocity distribution. This approach provides random transport/routing times into or within the ED (based on means of arrival and wing assignment) for use within the model.

Samaha, et al. (2003) report a study conducted for the ED at Cooper Health System in southern New Jersey. The simulation model depicted current operations and was used to evaluate concepts for reducing LOS for ED patients. An extensive set of input data where required, including arrival rates and arrival percentages disaggregated by arrival mode (car or ambulance), as well the duration of each of the steps in the patient process flow. Data were collected on site continuously, over a seven-day period, for every patient treated.

Baester, et al. (2003) developed a simulation model to estimate the maximum demand increment that could be sustained by emergency room of a private hospital in Chile, without increasing the waiting time beyond an acceptable level. The information required as input for the model, such as arrival rates, type of diagnosis, and the type and duration of treatments, were collected from existing hospital databases.

Wiinamaki and Dronzek (2003) describe an emergency-care center simulation project at Sarasota Memorial Hospital in Sarasota, FL. The objective was to project bed requirements for an ED expansion, as well as to project impacts on downstream departments at the center. Data were captured from the hospital on-line system to calculate the patient age and LOS. These data were incorporated into the model as a patient age-distribution applied to forecast arrivals for patients. An age-based multiplier, based on historical LOS, was used to determine LOS for each arrival.

Mahapatra, et al. (2003) developed a decision support system (DSS) using the Emergency Severity Index (ESI) triage method to drive improvements in the care delivery process for an academic ED in York Hospital, PA. The DSS pairs the ESI case mix with simulation to support resource deployment, improve service metrics, and support strategic decision making. Thirty months of patient data provided a total of 160,000 patient arrivals. According to the authors, "York Hospital is a progressive, forward-looking institution that has had extensive data collection systems in place for a long period of time. These data proved invaluable for model building, verification, and validation. The data comprised the key service times that go into delivery of clinical service, including patient arrival times, waiting times at various stages (used for model validation), the case mix of the patients according to the five-level triage, the times spent by the patient at various stages of the service delivery process, and staff schedules of ED personnel. Interviews with emergency physicians, residents, nurse administrators, and technicians were conducted to gain valuable insights into the process of ED service delivery in this teaching environment. Direct observation and the judgment of staff directly involved in the service delivery operations were used to derive average processing times and to represent the fluid interactions between the different service delivery processes."

Centeno, et al. (2003) developed a tool that integrates DES with an integer linear program (ILP) to investigate optimal staffing schedules for Baptist Health South Hospital Coral Gables, FL. The simulation establishes the staffing requirements for each period and the ILP determines how many staff members should start at each shift. All patients that came in to the ED (for an unreported time period) were logged in with all of their properties, including arrival mode, disposition, and auxiliary procedure received. The arrival pattern yielded seven different distributions. Service times were collected for the two main stages of patient flow (RN and MD) and ancillary activities.

Blasak, et al. (2003) present a DES of the ED and Medical Telemetry (Med Tele) Units at Rush North Shore Medical Center in Skokie, IL, which allows management to see the operations of both units, as well as interactions between the two. The model depicts current operations and evaluates possible alternatives to reduce ED LOS and improve operations. An extensive list of data requirements, including arrival times by mode and delay times for a large number of processes, is provided for both model components. In response to the inability to gather a large enough data sample over the course of a week's time, "the study opted to gather as much actual data as possible and rely on internal systems to supplement the data that was gathered."

Miller, et al. (2004) describe a reusable product developed to model and test alternative ED design scenarios. This product responds to the need by hospital administrators to improve key performance indicators, such as patient LOS, bed utilization, and elimination of bottlenecks. Data requirements and resource issues are not reported.

Sinreich and Marmor (2004) assert that, in order to bolster the credibility and acceptance of simulation studies in healthcare systems in general (and ED's in particular), (i) hospital management should be directly involved in the development of simulation models, (ii) these models should be simplified as much as is reasonably possible, and (iii) the use visual aids should be emphasized. This study lays the foundation for the development of a DES which is general, simple to use, and contains default values for most of the system's parameters. Meetings were held with senior physicians and head nurses of each of five ED's to define specific procedures ("elements") routinely performed by the ED staff; teams of supervised students, equipped with standardized code lists, conducted time and motion studies for these elements. A total sample size of 16,250 elements was gathered.

In addition to the data gathered, three hospitals provided historical patient data (about 24 months) from their computerized information systems. The data covered the main sites that handle patients at each hospital: the ED, the imaging centers, and the labs. A unique process chart (available from the authors by request) was developed for each of nineteen patient types at each of the five ED's, which included the duration (mean and variance) of each elements and the frequencies of each of the connections between the different elements. A unified process chart, comprising all the different elements and transitions, was constructed. The authors also provide technical details on the transformation of the data collected into information needed in the model, including a cluster analysis of the principle patient types, the nonstationary time series for patient arrivals at the ED and imaging center, and staff walking model.

Takakuwa and Shiozaki (2004) simulate the ED of the Gifu Prefectural Hospital in Japan to examine patient flows. Times needed for both outpatients and patients arriving via ambulance to be processed in the ED were examined and "a special-purpose data-generator was developed to create experimental data for executing a simulation." It was found that the patients spend the longer part of their system-time waiting; that waiting time depends on the number of patients to be processed; and that the waiting time for available emergency-treatment beds, doctors, drips, and stretchers accounts for the major part of all the waiting time in the ED. A stepwise procedure of operations planning was proposed to minimize patient waiting times and numerical examples are given which illustrate the procedure.

2.3 Other Hospital Units and Services

Rossetti, et al. (1998) discuss the use of DES to analyze the costs, benefits, and performance tradeoffs related to the installation of a fleet of mobile robots within mid-size hospital facilities, in lieu of human curriers. Results show that for clinical laboratory deliveries, a fleet of six mobile robots can achieve significant performance gains over the current system of three human couriers, while remaining cost effective. Entities representing clinical laboratory delivery items are generated according to a non-stationary Poisson process and sent to the appropriate units for pickup according to a demand-probability distribution. Data for mean arrival time and demanding unit were based on historical data supplied by Distribution Services at the University of Virginia Hospital.

A combination of more patients and increased LOS stressed the capacity of the maternity unit at Miami Valley Hospital, Dayton, OH. Johnson (1998) simulated patient flows and census to determine the effect of configuration and policy changes. Data issues and sources are not reported.

Healthcare systems must balance the need for access and availability of intensive care unit beds (ICU) against the high costs of excess ICU capacity. Cahill and Render (1999) developed a DES of patient flows through the ICU, telemetry, and medical-floor beds under current bed allocation for the Cincinnati Veterans Administration Medical Center. Under this allocation, ICU beds are unavailable nearly one third of the time, eliminating new ICU admissions and requiring diversion of ambulance traffic. The simulation was then used to evaluate the effects of the phased construction, intended to relieve the problem. Arrival rates, the average LOS for each unit, and bed downtimes (while a bed is being prepared between occupancies) were determined from historical means, augmented by expert opinion concerning distribution form and variability in each case.

Lowery and Davis (1999) developed a DES for Brigham and Women's Hospital in Boston to examine issues related to a planned renovation the surgical suite. The renovation would reduce the number of operating rooms from 34 to 32, but 95% of all outpatient cases currently served would be removed to another facility. The study concluded that the projected changes in surgical workload could be accommodated in operating rooms, if scheduled block times were extended during the weekdays and Saturday blocks were added. Historical data were used to develop distributions for procedure times; turnaround times (i.e., time for case clean-up and set-up); and downtimes (usually due to delays, e.g., staff, patient, or equipment not available, or discrepancies between scheduled and actual surgery starting times).

Centeno, et al. (2000) explored six scenarios for improving efficiencies in patient flows and resource use for the Radiology Department at Jackson Memorial Hospital (JMH), Dade County, FL. The DES was fed activity data from an extensive, existing radiology operations database. Records were analyzed to establish probabilistic models for the DES inputs. Because arrivals at the department are nonstationary, the arrival data were aggregated into one-hour segments from seven in the morning to ten at night for each day of the week. The difficulties encountered in data preparation are identically those cited by Centeno and Carrillo (2001) above, but further complicated by the recognition that many post-operative patients must *return* for additional procedures.

Centeno, et al. (2001) describe a second DES study for JMH, focusing on improving doctor and staff scheduling for the Labor and Delivery Rooms. The Engineering Management Group provided data for 1999 for determining the necessary statistical distributions for arrivals and activity times. Databases for two departments were employed, with 3,309 records in the OB database and another 3,500 records in the OBPAR database. Analysis of these records revealed that, although there is a significant amount of data, some of it is not usable (with some records even showing negative service times). Nonetheless, it was possible to establish relatively good inputs for the DES from the cleansed data. Procedures used to scrub and transform the historical data and correlate records common to the same patient appearing in each database are detailed.

Baesler and Sepúlveda (2001) integrated a DES with a multi-objective optimization heuristic to find the best combination of control variables, considering four performance measures for a cancer treatment center facility. Data issues related to the case study are reported in Sepúlveda, et al. (1999).

Martin, et al. (2003) consider the unmet demand for geriatric care in Norwegian hospitals and the potential and means for increasing patient throughput at the geriatric wards. Data for geriatric patients were acquired from the IT department, including the wards visited prior to admission to the geriatric department, the LOS at each ward, and total LOS prior to admission. Modeling clinical pathways and resource use within the wards was complicated by multiple diagnose codes and inconsistencies in the application of these codes across wards. In the time period examined, there were several hundred combinations of primary and secondary diagnoses. "Since using up valuable physician time to describe all clinical paths was not a viable option," the problem was solved by first identifying the most common primary diagnoses and then using a table of conditional probabilities for secondary diagnoses. Doctors were asked to describe in detail the clinical pathways for each such primary/secondary cluster. Interval estimates were derived to accommodate for uncertainties.

The Telemedicine Program was created to provide medical assistance to people living in extreme poverty in Mexico. Through a satellite connection, a physician located in Mexico City can diagnose the patient who is physically inspected on a mobile unit equipped with telecommunications gear. Lach and Vázquez (2004) describe how the Program's processes were simulated in order to analyze the possible results for different model input parameters. Data sources are not identified.

Ferrin, et al. (2004) discuss the application of DES to assess the value of an incentive program in the operating room (OR) environment at St. Vincent's Hospital, Birmingham, AL. The model also was used to evaluate operational changes, including scheduling processes within the OR and utilization rates in areas such as the Post Anesthesia Care Unit and the Ambulatory Surgery Department. Lessons learned are presented, but data resource issues are not discussed.

Ramakrishnan, et al. (2004) analyzed flows in a teaching hospital (Wilson Memorial Regional Medical Center, Broome County, NY) that was implementing a digital image archiving system within its radiology services. They identified changes to the existing workflow at the CT Scan area that would maximize patient throughput and minimize report-generation time with the digital system. Process mapping was used to identify the initial flow of operations and a DES was built to evaluate the different scenarios. Data from the Radiology Information System were collected over a period of four months and applied to calibrate arrival rates for ED patients, inpatients, and outpatients. A time study yielded distributions for patient CT-Scan waiting times, exam durations, folder retrieval times, image waiting times at the radiologists desk, and report-generation times.

2.4 Outpatient Clinics and Treatment Centers

Swisher, et al. (1997) applied a DES to support the design and development of the Queston Physician Network, encompassing the operations of both a clinic and an Information Center. The reusable DES incorporates standard techniques, such *acceptance/rejection* to describe walk-in patient arrivals and *thinning* to create a non-homogeneous Poisson process for incoming phone calls. "The data collection effort required to obtain the…patient information [needed for categorization] was deemed both too costly and too time-consuming. Some data collection was performed at a nearby clinic, though the modelers relied primarily upon existing data sources and the medical experts' clinical knowledge." For example, since the duration of various medical processes for each patient category were unavailable, medical experts were asked to estimate the parameters of triangular distributions, relying on their own clinical experiences and published medical information. "After a great deal of consultation and reformulation between the modelers and the medical experts, the group as a whole felt comfortable with its definition of the patient population." Of special note, the modelers had to create an additional class of humans (the companions who accompany the patient to the clinic) who do not utilize the clinic's medical resources, but do use its waiting rooms.

Barnes, et al. (1997) review successful health-care applications of DES in pre-op procedures, space utilization, and outpatient studies. The objective of the outpatient studies for University Hospital and Medical Center at Stony Brook, NY, was to explore means to reduce patient waiting time in three clinics (Family Medicine, Ophthalmology and Neurology). A data collection sheet was designed to capture necessary data elements (e.g., patient arrival and departure times; start and end times of all the patient flow steps in the clinic) efficiently and accurately. In each of the studies, two weeks were devoted to data gathering, with staff members logging work and time data during patient visits. The resulting database was used to determine DES input distributions such as arrivals and exam times. The two-week study also established a 'snapshot' of the current situation, which was used to validate the DES. "While the use of theoretical distributions and predetermined industry time standards are valid alternatives, we chose actual data not only because of our confidence in its accuracy but also because, in practical terms, it is 'real' data the clinic staff can relate to. This enhanced the credibility of the simulation process."

In addition to standard performance measures reflecting resource and scheduling requirements (e.g., throughput, system time, and queue times and lengths) Weng and Houshmand (1999) also consider total cash flow in their DES of a local clinic in Cincinnati, OH. Data collected by benchmarking engineers over a two-month period consisted principally of patient process and arrival times. An interview with the staff manager provided the standard staff scheduling times and requirements, as well as the probabilities for process branches. Expense data for the 16,188 cases treated in 1998 provided total revenue and the variable and fixed costs associated with labor, supplies, overhead, depreciation, medical education, and average staff salaries for various job descriptions.

Isken, et al.(1999) present a general framework for simulating outpatient obstetrical clinics in order to explore questions related to demand, appointment scheduling, exam-room allocation, patient flow, and staffing. Modeling challenges are identified, solutions suggested, and examples from a project completed by the authors are used to illustrate the concepts. In particular, the authors discuss the validity of a non-homogeneous Poisson process as a model for demand for appointments, when available data are lumped into finite time buckets and the population being served is large in relation to the number of appointments per time period. The authors also note that a modeling decision must be made related to whether the "atomic level of demand" is a single appointment request, or a sequence of appointments for the same patient. While most prenatal patients have a sequence of increasingly more frequent visits as the pregnancy progresses, data limitations may prevent the use of a sequence of visits as the atomic level of demand for routine prenatal visits. Finally, the authors note that one must be careful to differentiate between demand for appointments and actual visits. Actual visit volume by day-of-week and time-of-day is related directly to the hours of operation, staffing patterns, and appointment templates. At an aggregate level, the number of actual visits is affected by urgent care patients, cancellations, no-shows, and perhaps unmet demand; at a capacity-constrained facility, urgent care volume may be partially linked to the difficulty of obtaining an appointment. At a minimum, urgent care visits can only occur during the hours of operation.

Sepulveda, et al. (1999) analyzed patient flow through the M. D. Anderson Cancer Center in Orlando, FL, a full-service cancer treatment center. The objectives were (i) to evaluate the impact of alternative floor layouts, using different scheduling options, and (ii) to analyze resource and patient-flow requirements for a new building. The study justified relocation of the Center's laboratory and pharmacy, identified changes in scheduling procedures that would allow a significant increase in patient throughput with the same resources, and flagged a waiting room area in the new building design that would be too small for the increased patient flow. The model generates different arrival schedules for Medical Oncology and Ambulatory Treatment Center patients. Historical arrival data were gathered from hospital personnel in charge of patient scheduling, such that the model emulated existing scheduling patterns, as well as typical variations (lateness, earliness) observed in real-life. Because no historical data were available for activity-time distributions for each of nine processes, these were specified as uniform or triangular, using expert opinion derived from interviews with appropriate hospital personnel. Expert opinion, as well as some historical data, was used to build an empirical distribution of chemotherapy treatment times. The amount of data by itself was insufficient to fit a theoretical distribution, but was useful in identifying different treatment length ranges.

Alexopoulos, et al. (2001) observe that, while large healthcare institutions increasingly use DES to meet operational and competitive challenges, small healthcare facilities serving the poor typically lack the resources to do the same. The authors describe an effort to readdress this inequity through the creation of a low-cost, generic DES. A workflow data-acquisition spreadsheet was created that can be completed by clinic staff, facilitating "customizing" the DES for their own purposes. Use of the spreadsheet to collect data for over 400 direct patient encounters is described. Probability distributions were fit to the data, including arrivals disaggregated by varying patient characteristics and service delays associated with check-in, waiting room, pre-exam, exam, checkout, and post-checkout. "These stages involve internally consistent activities, personnel, and functions, the product of which has bearing on the successful completion of activities embedded in the other stages (e.g., failure to initiate appointment reminders increases the likelihood of patient "no-shows").

Ramis, et al. (2002) describe a generic simulator developed for the ACHS-Arauco Health clinical laboratories network in Chile. The simulator was used to standardize service processes, assign personnel, and guide investment decisions, with the overall objective of reducing the time-in-system for outpatients. Different configurations of resources were studied to detect bottlenecks, reallocate personnel to peak hours, and redesign facilities. With the help of laboratory personnel, exams were grouped into seven different families and process times and resources demands were studied for each family. Data were gathered for days having peak demand (Tuesdays, Thursdays and Saturdays). Arrival- and service-time distributions were based on measurements taken in different areas of the labs (over an unreported time period).

Morrison and Bird (2003) applied DES in a free-standing ambulatory health care setting. Critical success factors for the project were visual mapping and DES tools, process and resource mapping, data collection methods, technologies that minimize on-site time for client and consultant, and the importance of having a clinical healthcare professional on the consultant team. "The model required data distributions for tasks associated with telephone calls, medical records, clinical, and business paperwork for patients." Three types of distributions were required—(i) patient activities directly with administrative and clinical staff, (ii) non-patient contact time that consumed staff time, both clinical and administrative, and (iii) on-site non-patient activities such as mail prescription programs, appointment scheduling, and reminder calls. Time data for these tasks did not exist and, "The challenge was how to get this data without spending an exorbitant amount of time on site." While the center had been through a series of Patient Flow Analyses (using a DOS-based software program developed by the Centers for Disease Control and Prevention), based on one or two days of patient data logs, results did not give statistically valid information for day-to-day and week-to-week variation, and did not show the effect of non-patient work items. "We [ultimately] solved this problem by using PDAs and software that could record multiple events at one time and be used for elapsed-time events as well as random samplings of event occurrences."

The process by which outpatients are scheduled for a doctor's visit is a crucial determinant of the overall efficiency of patient flows. The problem consists of determining prioritization (triage) rules so that adequate patient care is guaranteed, resources (provider schedules) are utilized efficiently, and a service guarantee can be ensured. Guo, et al. (2004) present a DES framework for this problem and summarize their experience with a preliminary implementation for the Division of Pediatric Ophthalmology at Cincinnati Children's Hospital Medical Center. Two years of historical data from the KIDs (Kids Inpatient Database System) hospital information system were used to estimate empirical distributions for the majority of inputs, while only a few parameters had to be estimated for lack of data.

3 HEALTHCARE COST AND UTILIZATION PROJECT (HCUP)

HCUP (2004) is a Federal-State-Industry partnership sponsored by the U.S. Agency for Healthcare Research and Quality. Its mission is to obtain data from statewide information sources, to design and develop multi-state health care databases for health services research and health policy analysis, and to make these data available to a broad set of public and private users. Beginning in 1988, HCUP databases contain patient-level data (PLD) for all payers compiled in a uniform format with privacy protections in place. With the exception of the study by Guo, et al. (2004), this source of data (available under license for a fee) appears to have been untapped by DES practitioners.

The six HCUP databases include: the Nationwide Inpatient Sample (NIS), with inpatient data from a national sample of over 1,000 hospitals; the Kids' Inpatient Database (KID), a nationwide sample of pediatric inpatient discharges; the State Inpatient Databases (SID), the universe of inpatient discharge abstracts from participating states; the State Ambulatory Surgery Databases (SASD), with data from ambulatory care encounters from hospital-affiliated and sometimes freestanding ambulatory surgery sites; and the State Emergency Department Databases (SEDD), with data from hospital-affiliated emergency departments for visits that do not result in hospitalizations.

3.1 SID and VHI

As an alternative to considerable effort, cost, and delay required to obtain data from individual subject hospitals, PLD for hospital inpatients are available as SID (2004) for thirty-six participating states. SID for the Commonwealth were obtained through Virginia Health Information (VHI). While insufficient to support detailed analyses of inpatient flows within hospital units, these data are more than adequate to derive many the time series needed for a more aggregate area-wide, patient-flow DES of the type envisioned for emergency responses.

The SID contain inpatient discharge abstracts translated into a uniform format to facilitate intrastate and interstate comparisons and analyses. Together, the SID encompass about 90 percent of all U.S. community hospital discharges. Some states include discharges from specialty facilities, such as acute psychiatric hospitals. The SID contain a core set of clinical and non-clinical information on all patients, regardless of payer, including persons covered by Medicare, Medicaid, private insurance, and the uninsured. In addition to the core set of uniform data elements common to all SID, some include other elements, such as the patient's race. The SID contain clinical and resource use information included in a typical discharge abstract.

3.2 HIPAA and SID Data Limitations

In the U.S., any release of PLD after 1998 must conform to the 1996 Health Insurance Portability and Accountability Act (HIPAA, 2005). The primary objectives of HIPAA are to assure health insurance portability between employers, reduce healthcare fraud and abuse, enforce standards for health information, and guarantee the security and privacy of health information. To ensure privacy, *public release* versions of SID data are necessarily inadequate to derive detailed hospital-admissions and similar time series. For example, while the VHI public release PLD provide 68 data fields for each patient discharge record (VHI, 2003), the time of a patient's admission and discharge is given only to the quarter and year.

By executing a confidentiality agreement, however, a *special release* of the VHI data for the four years spanning the Richmond mass diversion events (2000-2003, inclusive) was obtained as twelve quarterly files. This special release included both the date and time of each patient's admission and discharge, in addition to the public-release data fields. The acquisition required approximately two months to complete, at an approximate cost of $15K.

While exceptionally clean, one shortcoming of the VHI PLD files is that these are packaged by *the quarter in which the patient was discharged*. Thus records for patients with stays that span multiple quarters do not appear in the data file for the quarter in which the patient was admitted, or in the files for any intervening quarters prior to the discharge quarter. This implies that data files for quarters extending beyond the study period must be concatenated with files spanning this period, in order to obtain a complete picture of admission time series. Moreover, while the time series for the net change in hospital patient census can be obtained directly with this concatenation, the absolute hospital census at any time must be approximated from discharge and LOS data, or obtained from an independent source. Finally, SID only contain records for inpatients. ED data must still be obtained, either from the subject hospitals directly, or indirectly from the SEDD. Fortunately, as noted by many authors in the literature review, complete ED data typically are readily available from existing in-house information systems.

ACKNOWLEDGMENT

Special thanks to Barbara Tawney and Robert Bennett.

REFERENCES

Alexopoulos, C., D. Goldsman, J. Fontanesi, M. Sawyer, M. De Guire, D. Kopald, and K. Holcomb. 2001. A Discrete-Event Simulation Application for Clinics Serving the Poor, *Proceedings of the 2001 Winter Simulation Conference,* B. A. Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer, (eds.), 1386-1391

Baesler, F. F., and J. A. Sepúlveda. 2001. Multi-Objective Simulation Optimization for a Cancer Treatment Center, *Proceedings of the 2001 Winter Simulation Conference,* B. A. Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer, (eds.), 1411-1405

Baesler, F. F., H. E. Jahnsen, , and M. DaCosta. 2003. The Use of Simulation and Design of Experiments for Estimating Maximum Capacity in an Emergency Room." *Proceedings of the 2003 Winter Simulation Conference*, S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice (eds.), 1903-1906

Barnes, C., C. Benson, and D. McGuinness. 1997. Success Stories in Simulation in Health Care, *Proceedings of the 1997 Winter Simulation Conference*, S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson (eds.), 1280-1285.

Blasak, R. E., W. S. Armel, D. W. Starks, and M. C. Hayduk. 2003. The use of Simulation to Evaluate Hospital Operations between the Emergency Department and a Medical Telemetry Unit, *Proceedings of the 2003 Winter Simulation Conference*, S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice (eds.), 1887-1893.

Cahill, W. and M. Render. 1999. Dynamic Simulation Modeling of ICU Bed Availability, *Proceedings of the 1999 Winter Simulation Conference,* P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans (eds.), 1573-1576

Centeno, M. A., Albacete, C., Terzano, D.O., Carrillo, M., and Ogazon, T. 2000. A Simulation Study of the Radiology Department at JMH, *Proceedings of the 2000 Winter Simulation Conference,* J. A. Joines, R. R. Barton, K. Kang, and P. A. Fishwick (eds.), 1978-1984

Centeno, M. A., and M. Carrillo. 2001. Challenges of Introducing Simulation as a Decision Making Tool, *Proceedings of the 2001 Winter Simulation Conference,* B. A. Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer (eds.), 17-21

Centeno, M. A., R. Giachetti, R. Linn, and A. M. Ismail. 2003. A Simulation-ILP based Tool for Scheduling ER Staff, *Proceedings of the 2003 Winter Simulation Conference*, S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice (eds.), 1930-1938

Centeno, M. A., M. A. Lee, E. Lopez, H. R. Fernandez, M. Carrillo, and T. Ogazon. 2001. A Simulation Study of the Labor and Delivery Rooms at JMH, *Proceedings of the 2001 Winter Simulation Conference,* B. A. Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer (eds.), 1392-1400

Cohort. 2002. *Metropolitan Richmond Hospital Diversions: A Systems Analysis and Change Proposal*, Project report, Systems and Information Engineering Executive Master's Program, University of Virginia, Charlottesville, VA.

Eldabi, T., and R. J. Paul. 2001. A Proposed Approach for Modeling Healthcare Systems for Understanding, *Proceedings of the 2001 Winter Simulation Conference,* B. A. Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer (eds.), 1412-1421

Ferrin, D. M., M. J. Miller, S. Wininger, and M. S. Neuendorf. 2004. Analyzing Incentives and Scheduling in a Major Metropolitan Hospital Operating Room through Simulation, *Proceedings of the 2004 Winter Simulation Conference,* R .G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters (eds.), 1975-1980

Guo, M., M. Wagner, and C. E. West. 2004. Outpatient Clinic Scheduling – A Simulation Approach, *Proceedings of the 2004 Winter Simulation Conference,* R .G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters (eds.), 1981-1987

HIPAA. 2005. Medical Privacy—National Standards to Protect the Privacy of Personal Health Information, U.S. Dept of Health and Human Services [online]. Available via <http://www.hhs.gov/ocr/hipaa> [accessed July 9, 2005].

HCUP. 2004. Overview of HCU, [online]. Available via <http://www.hcup-us.ahrq.gov/overview.jsp> [accessed July 9, 2005].

Isken, M. W., Ward, T.M., and McKee, T.C. 1999. Simulating Outpatient Obstetrical Clinics, *Proceedings of the 1999 Winter Simulation Conference,* P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans (eds.), 1557-1563

Jenkins, R., Y. Deshpande, and G. Davison, 1998. Verification and Validation and Complex Environments: A Study in Service Sector, *Proceedings of the 1998 Winter Simulation Conference,* D.J. Medeiros, E.F. Watson, J.S. Carson and M.S. Manivannan (eds.), 1433-1440

Johnson, W. C. 1998. Birth of a New Maternity Process, *Proceedings of the 1998 Winter Simulation Conference,* D.J. Medeiros, E.F. Watson, J.S. Carson and M.S. Manivannan (eds.), 1429-1432

Jun, J. B., S. H. Jacobson, and J. R. Swisher. 1999. Application of Discrete-Event Simulation in Health Care Clinics: A Survey, *Journal of the Operational Research Society*, 50(2)109-123.

Lach, J. M., and R. M. Vázquez. 2004. Simulation Model of the Telemedicine Program, *Proceedings of the 2004 Winter Simulation Conference,* R .G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters (eds.), 2012-2017

Lowery, J. C., and J. A. Davis. 1999. Determination of Operating Room Requirements Using Simulation, *Proceedings of the 1999 Winter Simulation Conference,* P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans (eds.), 1568-1572

Mahapatra, S., C. P. Koelling, L. Patvivatsiri, B. Fraticelli, D. Eitel, and L. Grove. 2003. Pairing Emergency Severity Index5-Level Triage Data with Computer Aided System Design to Improve Emergency Department Access and Throughput, *Proceedings of the 2003 Winter Simulation Conference*, S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice (eds.), 1917-1925

Martin, E., R. Grønhaug, and K. Haugene. 2003. Proposals to Reduce Over-Crowding, Lengthy Stays and Improve Patient Care: Study of the Geriatric Department in Norway's Largest Hospital, *Proceedings of the 2003 Winter Simulation Conference*, S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice (eds.), 1876-1881

Miller, M. J., D. M. Ferrin, and M. Messer. 2004. Fixing the Emergency Department: A Transformational Journey with EDsim, *Proceedings of the 2004 Winter Simulation Conference,* R .G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters (eds.), 1988-1993

Morrison, B. P., and B. C. Bird. 2003. A Methodology for Modeling Front Office and Patient Care Processes in Ambulatory Health Care, *Proceedings of the 2003 Winter Simulation Conference*, S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice (eds.), 1882-1886

Pitt, M. 1997. A Generalized Simulation System to Support Strategic Resource Planning in Healthcare. *Proceedings of the 1997 Winter Simulation Conference,* S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson (eds.), 1155-1162.

Ramakrishnan, S., K. Nagarkar, M. DeGennaro, K. Srihari, A. K. Courtney, and F. Emick. 2004. A Study of the CT Scan Area of a Healthcare Provider, *Proceedings* *of the 2004 Winter Simulation Conference,* R .G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters (eds.), 2025-2031

Ramis, F. J., J. L. Palma, V. F. Estrada, and G. Coscolla. 2002. A Simulator to Improve Patient's Service in a Network of Clinic Laboratories, *Proceedings of the 2002 Winter Simulation Conference*, E. Yücesan, C.-H. Chen, J. L. Snowdon, and J. M. Charnes (eds.), 1444-1447

Rossetti, M. D., A. Kumar, and R. A. Felder. 1998. Mobile Robot Simulation of Clinical Laboratory Deliveries, *Proceedings of the 1998 Winter Simulation Conference,* D.J. Medeiros, E.F. Watson, J.S. Carson and M.S. Manivannan (eds.), 1415-1422

Rossetti, M. D., G. F. Trzcinski, and S. A. Syverud. 1999. Emergency Department Simulation and Determination of Optimal Attending Physician Staffing Schedules, *Proceedings of the 2003 Winter Simulation Conference*, P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans (eds.), 1532-1540.

Samaha, S., W. Armel, and D. Starks. (2003). The Use of Simulation to Reduce the Length of Stay in an Emergency Department, *Proceedings of the 2003 Winter Simulation Conference*, S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice (eds.), 1907-1911

Sanchez, S. M., T. Ogazon, D. M. Ferrin, and J. A. Sepulveda. (2000). Emerging Issues in Healthcare Simulation. *Proceedings of the 2000 Winter Simulation Conference,* J. A. Joines, R. R. Barton, K. Kang, and P. A. Fishwick (eds.), 1999-2003

Sepulveda, J. A., W. J. Thompson, F. F. Baesler, M. I. Alvarez, , and L. E. Cahoon, III. 1999. The Use of Simulation for Process Improvement in a Cancer Treatment Center, *Proceedings of the 1999 Winter Simulation Conference,* P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans (eds.), 1541-1548

SID. 2004. Overview of the State Hospital Databases (SID) [online]. Available via <http://www.hcup-us.ahrq.gov/sidoverview.jsp> [accessed July 9, 2005].

Sinreich, D., and Y. Marmor. 2004. A Simple and Intuitive Simulation Tool for Analyzing the performance of Emergency Departments, *Proceedings of the 2004 Winter Simulation Conference,* R .G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters (eds.), 1994-2002

Swisher, J., B. Jun, S. Jacobson, and O. Balci. 1997. Simulation of the Question Physician Network, *Proceedings of the 1997 Winter Simulation Conference,* S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson (eds.), 1146-1154.

Takakuwa, S., and Hiroko Shiozaki. 2004. Functional Analysis for Operating Emergency Department of a General Hospital, *Proceedings of the 2004 Winter Simulation Conference,* R .G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters (eds.), 2003-2011

VHI. 2003. *Virginia Health Information Data Directory*, VHI, Inc., Richmond, VA. Available via <http://www.vhi.org/products_pld.asp> [accessed July 9, 2005].

Weng, M. L., and A. A. Houshmand. 1999. Healthcare Simulation: A Case Study at a Local Clinic, *Proceedings of the 1999 Winter Simulation Conference,* P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans (eds.), 1577-1584

Wiinamaki, A., and R. Dronzek. 2003. Using Simulation in the Architectural Concept Phase of an Emergency Department Design, *Proceedings of the 2003 Winter Simulation Conference*, S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice (eds.), 1912-1915

Note that all of the WSC *Proceedings* are published by Institute for Electrical and Electronics Engineers, Piscataway, NJ. Complete Proceedings since 1997 available at <http://www.wintersim.org/pastprog.htm>.

AUTHOR BIOGRAPHY

K. PRESTON WHITE, JR., is Professor of Systems and Information Engineering at the University of Virginia. Dr. White serves on the Advisory Board of Virginia Modeling, Analysis, and Simulation Center; is a member of the of the INFORMS Simulation Society Council; represents IEEE/SMC on the WSC Board of Directors; and currently serves as Immediate Past WSC Board Chairman and liaison to the WSC Foundation. He can be contacted at <kpwhite@virginia.edu>.