Health Care Manag Sci (2010) 13:210–221
DOI 10.1007/s10729-009-9124-9

A data-integrated simulation-based optimization for assigning nurses to patient admissions

Duraikannan Sundaramoorthi · Victoria C. P. Chen · Jay M. Rosenberger · Seoung Bum Kim · Deborah F. Buckley-Behan

Received: 13 July 2009 / Accepted: 22 December 2009 / Published online: 21 January 2010

Abstract The health care system in the United States has a shortage of nurses. A careful planning of nurse resources is needed to ease the health care system from the burden of the nurse shortage and standardize nurse workload. An earlier research study developed a data-integrated simulation to evaluate nurse-patient assignments (SIMNA) at the beginning of a shift based on a real data set provided by a northeast Texas hospital. In this research, with the aid of the same SIMNA model, two policies are developed to make nurse-to-patient assignments when new patients are admitted during a shift. A heuristic (HEU) policy assigns a newly-admitted patient to the nurse who has performed the least assigned direct care among all the nurses. A partially-optimized (OPT) policy seeks to minimize the difference in workload among nurses for the entire shift by estimating the assigned direct care from SIMNA. Results comparing HEU and OPT policies are presented.

Keywords Nurse assignment · Patient assignment · Simulation-based optimization

D. Sundaramoorthi (✉)
Steven L. Craig School of Business,
Missouri Western State University, St. Joseph, MO, USA
e-mail: dsundaramoorthi@missouriwestern.edu

V. C. P. Chen · J. M. Rosenberger
Dept. of Industrial & Manufacturing Systems Engineering,
University of Texas, Arlington, TX, USA

V. C. P. Chen
e-mail: vchen@uta.edu

J. M. Rosenberger
e-mail: jrosenbe@uta.edu

S. B. Kim
Division of Information Management Engineering,
Korea University, Seoul, Korea
e-mail: sbkim1@korea.ac.kr

D. F. Buckley-Behan
School of Nursing,
University of Texas at Arlington, Arlington, TX, USA
e-mail: dbehan@uta.edu

1 Introduction

The health care system in the United States is severely strained because of a shortage of nurses and nurse burnout [49]. Health care policy makers have responded to this crisis in many ways. For instance, significant financial resources were made available to expand nursing education during last few years [21, 25]. Recently, hospitals have been actively thinking of strategies to recruit and retain nurses. Such strategies often call for a state-of-the-art work environment and easy access to career development. A Wall Street Journal article reports a projected spending of $200 billion on construction and renovation of hospitals through 2014 [32]. As part of developing nursing careers, hospitals are launching residency programs and short-term courses enabling easy access for working nurses [10]. Due to commendable effort in different initiatives, there are early signs of the easing of the nurse shortage in selected hospital systems [43].

While significant progress has been made in different aspects of nursing, few efforts have been made to manage nurse-to-patient assignments and balance nurses' workload for a given shift. In an earlier research, Sundaramoorthi et al. [47] developed a data-integrated simulation to evaluate nurse-patient assignments

(SIMNA) at the beginning of a shift based on a real data set provided by a northeast Texas hospital. SIMNA utilized tree-based models and kernel density estimation to extract important knowledge from the real data set. In this current research two policies are developed to make nurse-to-patient assignments for newly admitted patients during a shift and they are evaluated with the aid of the SIMNA model.

There are two major contributions made in this research:

– During a shift when new patients arrive, nurse supervisors often assign the new patient to the nurse who has the least number of patients. This way of assignment may not balance the workload of nurses for the entire shift. This research enhances SIMNA by adding a feature that assists in assigning a nurse to a newly admitted patient during a given shift. The enhanced SIMNA model can aid nurse supervisors to make better decisions by simulating different new-patient assignment policies and quantifying the workload measures from them.
– This research develops and compares a partially-optimized policy (OPT) with a heuristic policy (HEU) to make nurse-to-patient assignments when new patients are admitted during a shift. The HEU policy assigns a newly-admitted patient to the nurse who has performed the least assigned direct care among all the nurses 15 minutes prior to a new patient admission; while the OPT policy seeks to minimize the difference in workload among nurses for the entire shift by estimating the assigned direct care from SIMNA.

The rest of this paper is organized as follows. In Section 2, a literature review on nurse resource planning and simulation-based optimization is provided. In Section 3, a brief review of the SIMNA model is provided. In Section 4, the assignment policies OPT and HEU are developed. Section 5 compares the assignment policies (OPT and HEU) using SIMNA. In Section 6, concluding remarks and future research directions are presented.

2 Literature review

There are two major components in this research— — — nurse resource planning and simulation-based optimization. This section gives a brief literature review on these two topics.

2.1 Nurse resource planning

Nurse burnout issue in health care was reported as early as 1979 [46]. Cullen [12] identified the factors that were embedded within health care, institutional, societal, and nursing systems that caused stressful conditions and burnout for nurses. As a result of burnout, nursing profession has a chronic problem of high turnover, absenteeism, and reduced productivity [11]. Staffing level is one of the key factors that contribute to the nurse burnout [19]. Staffing levels were also found to have a positive correlation with the patient outcomes [22]. In the last couple of decades, several research works addressed determining staffing levels and schedules. Miller et al. [38] developed a constraint-based, artificial intelligence nurse scheduling prototype by incorporating nurses' preferences for Rouen University Hospital. Jaumard et al. [26] presented a 0–1 column generation method for nurse scheduling by maximizing the nurse preference and team balance, and minimizing the total nurse salary for the schedule. Bard and Purnomo [3] formulated and solved the nurse scheduling problem as a multi-objective problem which considered individual nurse's preference. In the past couple of decades, several patient classification systems and acuity systems were developed to aid determination of nursing care, staffing level, and schedule ahead of a shift [7, 8, 16, 23, 28, 36, 50]. It has to be noted that four levels of acuity were considered in this research depending upon the amount of care received by the patients in the north Texas hospital. The top 25 percent of patients who needed the most nursing care was given an acuity level of four while the bottom 25 percent got an acuity level of one. The other two groups got acuity levels two and three. None of the patient classification systems and acuity systems went as far as assigning patients to nurses for a given shift. To the best of our knowledge, apart from this research, only Punnakitikashem et al. [41], Vericourt and Jennings [48], Mullinax and Lawley [39], and Sundaramoorthi et al. [47] address the nurse to patient assignment problem. Punnakitikashem et al. [41] formulated a stochastic programming problem to assign nurses to patients while balancing the nurse workload and solved it using Bender's decomposition approach. Vericourt and Jennings [48] determined nurse-to-patient assignment ratios utilizing queuing theory. Sundaramoorthi et al. [47] developed the SIMNA model to evaluate nurse-to-patient assignments policies by considering hospital specific factors. Mullinax and Lawley [39] developed an acuity system for a neonatal intensive care unit to determine nursing care for each patient and assigned them to nurses by balancing nurse workload using an integer linear program. Apart from Sundaramoorthi et al. [47] and this research, none of the other methods discussed above used real data to reflect the actual system as extensively and developed a tool to evaluate nurse-to-patient assignments to make decisions in real time. This research extends the SIMNA model of Sundaramoorthi et al. [47] by embedding nurse-to-patient assignments policies for new patient admits during a shift. A brief review of SIMNA is provided in Section 3.

2.2 Simulation-optimization models

Studying industrial systems using simulation was prevalent as early as the late 1950's and early 1960's. Simulation modeling has been used to study a wide range of problems in health care [13, 15, 29, 35, 45]. In recent years, Zenios et al. [51], Kreke et al. [31], and Shechter et al. [44] utilized simulation models even to study organ allocation systems. A comprehensive review of health care simulation models can be found in Klein et al. [30] and Jun et al. [27]. In the literature, most of the health care simulations modeled patient flow and analyzed patient scheduling, admissions, routing, and availability of resources. Very few simulation research works like Duraiswamy et al. [14], McHugh [37], and Sundaramoorthi et al. [47] had staffing as the primary focus. In recent years, combining simulation and optimization has been made possible due to powerful computers. In simulation-optimization, the goal is to find simulation inputs (decision variables) in the allowable range (constraints) that optimize an objective function expressed in terms of the simulation outputs. For a comprehensive review of different simulation-based optimization methods refer to Fu [17], Fu and Hu [18], Hurrion [24], Law and Kelton [33], Law and McComas [34], Olafsson and Kim [40], and Robinson [42]. Simulation-based optimization is still at its early stages of development and to the best of our knowledge this is the first research that utilizes simulation-based optimization to address nurse-to-patient assignments.

3 SIMNA review

Sundaramoorthi et al. [47] developed SIMNA based on the data set obtained from a northeast Texas hospital. At the northeast Texas hospital, each nurse wears a locating device that transmits data to a repository from where the data was collected for this research. The hospital also provided information on admit dates, discharge dates, room numbers, and diagnoses for each patient. The data set with 570,660 observations contained information on nurse movements and patient characteristics of a Medical/Surgical care unit. The following variables were included in the data set:

1. Current location and previous two locations for each nurse.
2. Time spent in each nurse visit to a location.
3. Nurse types.
4. Shift.
5. Hour.
6. Diagnoses codes of patients in each patient room.
7. Acuity levels of patients in each patient room.
8. Nurse-to-patient assignment.

SIMNA utilized four classification trees to estimate probability distributions of nurse movements based on the current state of the system determined from the above listed variables; while a regression tree with kernel density estimates in each terminal node estimated the amount of time spent by nurses at different locations for any given simulation state in SIMNA. The simulation process, which involves repeated traversing of the tree structures, was written in C++.

The first use of SIMNA was to assess the balance of nurse workload that results from the nurse-to-patient assignment policies at the beginning of a shift. Specifically SIMNA tested four assignment policies: clustered, heuristic, stochastic program, and random assignments. In the clustered assignment, patients were assigned by location; that is, patients in consecutive rooms were assigned to the same nurse. In the heuristic assignment, all of the nurses got the same number of patients when the number of nurses divides into the number of patients evenly. The patient with the highest expected direct care time was arbitrarily assigned to a nurse. The patient with the second highest expected direct care time was then arbitrarily assigned to a second nurse, and so on. After assigning one patient for each nurse, in the second cycle of assignments, the patient with the lowest expected direct care time was assigned to the first nurse. The patient with the second lowest expected direct care time was assigned to the second nurse, and so on. This process of assignment was repeated until all of the patients were assigned. The stochastic program assignments were obtained from Punnakitikashem et al. [41]. Finally, the random assignment assigned equal number of patients to nurses randomly. The four policies were compared by quantifying each nurse′s workload. A test problem in Sundaramoorthi et al. [47] resulted in a superior performance of the clustered assignments among all assignments from the four policies. It should be noted that the superior performance of the clustered assignments is confined to the test problem and could differ for other problems. The purpose of SIMNA in Sundaramoorthi et al. [47] was to help hospital managements evaluate different assignment policies prior to a given shift and aid them decide the policy they would like to adapt for that shift. Identifying desirable nurse-to-patient assignment policies at the beginning of the shift for different circumstances would require designing an experiment with large number of treatments (discussed in Section 6) and would be an interesting research by itself. SIMNA utilized structures and pointers to reconstruct tree structures, and efficiently executed the simulation of an entire shift. It took less than three minutes on a Dual 2.4-GHz Intel Xeon Workstation to run 1000 scenarios of the shift when the above four policies were tested to evaluate the balance in nurse workload at the beginning of the shift. A prototype consisting SIMNA was evaluated by two groups of registered nurses enrolled in a north Texas University. 73% of them liked to utilize such a prototype in their work place. Based on their feedback SIMNA was enhanced by including acuity levels and more diagnoses codes. Refer to Baker et al. [2] for more information about the feedback obtained from the evaluation. In this research, we utilize the same SIMNA model to develop new-patient assignment policies in order to help the hospital management determine nurse-to-patient assignments when new patients are admitted during a shift. Similar to the assignments at the beginning of the shift, SIMNA produced the new-patient assignment results of 1000 scenarios, discussed in Section 5, in less than three minutes. Hence, it is possible to use this tool in real time to make nurse-to-patient assignment decisions when new patients are admitted.

4 Simulation-based optimization

4.1 Markov decision problem

Unlike Sundaramoorthi et al. [47], which evaluated initial assignments at the beginning of a shift, the topic of the present research is the assignment of new-patient admissions during the shift. It is assumed in this research, and also common in reality, that the time of admit, patient diagnosis, and patient acuity are known to the decision maker at least 15 minutes prior to the actual admission. A simple decision rule is to simply assign a newly-admitted patient to the nurse who had the least TADC among all the nurses 15 minutes prior to a new patient admission. This is referred to as the heuristic policy HEU. It has to be noted that the HEU policy is different from the initial assignment heuristic policy presented in Sundaramoorthi et al. [47].

More complex to develop is an optimized decision rule. Recently, formulating and solving Markov decision problems using a simulator have become common and successful [6, 20]. A typical Markov decision problem (MDP) would have the following components:

1. State: The state describes the status of a system under consideration. For example, specific values of the shift, the time of day, the nurse type, the current and previous locations of the nurse, the nurse-patient assignments, the patient diagnosis, the patient acuity, and the patient location variables can be considered as the state that describes our nurse-patient system.
2. Action: This is the decision that we desire to optimize. Our decision is the assignment of a newly admitted patient to a nurse.
3. Transition Probability: Transition probabilities determine transitions of the system from one state to another. Assume an action a selected for state i transfers the system to state j with probability $p(i, a, j)$, this quantity is an example of a transition probability. Collection of all such transition probabilities for all possible state transitions is required to capture the dynamics of the system modeled.
4. Policy: A policy defines what action to take based on the state of the system. For example, when a new patient is admitted during a shift, there are different policies that can be used to make the assignment based on the state. A policy that maximizes the sum of TADCs of nurses, shown in Eq. (6), would increase patient care. Two policies that balance nurse workload are presented in Section 4.2.
5. Performance Measure: A performance measure quantifies the performance of a policy. For a patient care improvement problem, the sum of TADCs over all nurses could be used to judge the performance of the policy.

In the late 1950's, a mathematical technique called Dynamic Programming (DP) was formulated by Bellman that could solve MDPs [4]. Since then, DP has evolved and been applied for various applications [5, 6, 9]. The theory and solution techniques of DP have also been studied and improved over the years. For a computationally tractable solution, most of the solution techniques reduce to either approximating or simplifying the Bellman optimality equation:

$$J^*(i) = \max_{a \in A(i)} \left[E(r(i,a)) + \sum_{j=1}^{\|S\|} p(i,a,j) J^*(j) \right] \forall i \in S. \quad (1)$$

where:

1. S is the set of all possible states.
2. $A(i)$ is the set of actions available for state i.
3. J^* functions store the unknown optimal values associated with each element in S.
4. $E(r(i,a))$ is the immediate expected reward in i when action a is selected.
5. $p(i,a,j)$ is the transition probability for the state transition from i to j when the action a is selected for state i.

Applying a classical method of solving Eq. (1), for optimizing the assignment of a newly-admitted patient, is impossible due to the high dimensional state space and unavailability of transition probabilities. When transition probabilities are not available explicitly, a Q-factors method uses a simulation model to solve the following equation, which is a mathematical equivalent of Eq. (1):

$$J^*(i) = \max_{a \in A(i)} \left[E(r(i,a)) + E\left(J^*(j)\right) \right] \forall i \in S. \quad (2)$$

Equation 2 can be further simplified as

$$J^*(i) = \max_{a \in A(i)} E\left(r(i,a) + J^*(j)\right) \forall i \in S. \quad (3)$$

Unlike the Bellman optimality equation, each element of Q-factors are associated to state-action pairs. For a state-action pair (i,a), the Q-factor is defined as

$$Q(i,a) = \sum_{j=1}^{\|S\|} p(i,a,j) \left[r(i,a) + J^*(j) \right] \quad (4)$$

By combining Eqs. (1), (3), and (4), we get

$$J^*(i) = \max_{a \in A(i)} Q(i,a) \quad (5)$$

Refer to Bertsekas [5] for a comprehensive review of Q-Factors methods. In the new-admit patient-nurse assignment optimization problem, if the objective is to maximize the sum of TADC across the nurses for the entire shift, the new-admit patient-nurse assignment optimization can be expressed as

$$J^*(i) = \max_{a \in A(i)} \left[\sum_{n=1}^{N} \mathrm{TADC}_n(i,a,i+1) \right] + E\left(J^*(i+1)\right) \forall i \in S. \quad (6)$$

In Eq. (6), N is the total number of nurses working in that shift, the state for the current new-patient-admit is denoted by i, the action a is taken to assign this new patient to a nurse, and then the subsequent state when the next new-patient-admit occurs is denoted by $i+1$. $\mathrm{TADC}_n(i,a,i+1)$ denotes the TADC of nurse n over the period from the current new-patient-admit in state i to the next new-patient-admit in state $i+1$ following the action of assignment a. Note that in Eq. (6), the notation i and $i+1$ represents high dimensional states determined by specific values of shift, time of day, nurse type, current and previous locations of nurses, existing nurse-patient assignments, patient diagnoses, patient acuities, and patient location variables. It is assumed that an action is required only when a new patient is admitted.

As mentioned earlier, when a simulation model is available, a computational optimization technique called Q-Factors is an attractive approach to solve Eq. (6). The fundamental idea of this approach is to store quantities $Q(i,a)$, shown in Eqs. (4) and (5), called Q-Factors for each state-action combination and update them based on the progress of the simulation. In the beginning, these Q-Factors are usually initialized to zero. Then for each action selected, the simulation is allowed to transition to the next state, and the Q-Factors are updated based on the performance measure. For the patient care improvement problem, a state-action pair yielding a larger sum of TADCs of all nurses would be rewarded by increasing the corresponding Q-Factor. State-action pairs yielding smaller sums of TADCs would be punished by reducing the corresponding Q-Factors. The same policy of rewarding and punishing has to be repeated for a sufficiently large number of state-action visits. At the end, the action(s) that produces the highest Q-Factor would be declared as optimum. The key for achieving the true or near optimum in the Q-Factors method depends on the choice of the so-called "sufficiently large number" for state-action pair visits. In the problem of optimizing the assignment of a newly-admitted patient, the number of state-action pairs grows exponentially due to random arrivals of patients (admit times) with the unknown probability distribution for diagnosis and acuity. Such a huge number of state-action pairs makes it computationally impossible to have enough simulation scenarios to obtain reliable Q-Factors.

4.2 Assignment policies

Even though increasing patient care is an important objective, in this research it is implicitly assumed that balancing nurse workload will help improve patient care, and hence the max-min TADC ratio was chosen to be the performance measure. In addition to the computational issues raised in the previous section, the max-min TADC ratio is not additive and consequently, the nurse workload balancing problem cannot be formulated like Eq. (6). For these reasons, methods like simple enumeration, classical DP, and Q-Factors are ruled out for this research.

Among the two expected values in Eq. (2), the first one incorporates the immediate reward i.e., in a sense, it accounts for the past and the immediate present. The second expected value, which approximates the future, for a current decision is impossible to approximate from simulation due to the huge number of potential state-action pairs. In the nurse-patient assignment problem, the difficulty reduces to the estimation of $\text{TADC}(i, a, i+1)$. While solving for the optimal assignment for state i, a huge number of simulation runs will be required to optimize assignments $a(i+1)$, $a(i+2)$, $a(i+3)$, ... For this reason, this research develops an alternate policy that groups both the expected values of Eq. (2) together:

$$J^{\wedge}(i) = \min_{a \in A(i)} E\left(\frac{(\text{TADC}\,(0, a(0), i) + \text{TADC}\,(i, a, T))_{max}}{(\text{TADC}\,(0, a(0), i) + \text{TADC}\,(i, a, T))_{min}}\right) \forall i \in S. \tag{7}$$

We refer to this policy as "OPT" since it is based on the Bellman optimality equation. In Eq. (7), $\text{TADC}_n(0, a(0), i)$ denotes the TADC of nurse n from the beginning of the shift until the current new patient arrival in state i when assignment $a(0)$ is made, and $\text{TADC}(i, a, T)$ is TADC from the current arrival through the end of the shift in state T. $\text{TADC}(i, a, T)$ can be expanded as $\text{TADC}(i, a(i), i+1) + \text{TADC}(i+1, a(i+1), i+2) + \text{TADC}(i+2, a(i+2), i+3)\ldots$; ideally these future assignments and TADC quantities would be obtained via a DP type optimization; however, this is computationally impractical. Instead, the future assignments required to obtain $\text{TADC}(i+1, a(i+1), i+2), \text{TADC}(i+2, a(i+2), i+3), \ldots$ were determined by the HEU policy. In simple terms, the OPT policy considers both the past and the future workload of nurses for a nurse-to-patient assignment decision, while the HEU policy considers only the past workload. The decision maker can use either HEU by itself or OPT to decide which nurse would get the new patient.

5 Comparison of policies

5.1 Problem setting

To analyze the performance of OPT and HEU, 50 problems with different initial states were considered. Admissions of two, three, four, five, and six new-patients were considered during a shift. The 50 problems were designed in such a way, shown in Table 1, to have ten problems for each shift and ten problems for each number of admissions. The number of problems for each combination of shift and the number of new admissions were arbitrarily chosen with rates of admission, shown in Table 2, in consideration. It is determined from the north Texas hospital data set that on average there were nine patient-admits for a given day with a maximum of six patients admitted during a shift. While solving an assignment, the future admits were simulated using a Poisson process with the arrival rates determined by the average number of patient admits per day and rates of admit for specific time period shown in Table 2.

There are 26 patient rooms in the Medical/Surgical care unit of the north Texas hospital usually staffed with five nurses. For all the 50 problems considered, the number of empty patient rooms was chosen to be the same as the number of new-patient admits. For a given problem, the empty patient room locations to accommodate new admits were selected randomly. The rest of the rooms were occupied by patients from the beginning of the shift. The diagnosis and acuity of patients present at the beginning of the shift as well

Table 1 Fifty problem instances

Shift (#)	# of New admits				
	2	3	4	5	6
Week					
Day (1)	2	5	3	0	0
Evening (2)	0	0	2	4	4
Night (3)	7	2	1	0	0
Week end					
Day (4)	0	0	0	5	5
Night (5)	1	3	4	1	1

Table 2 Patient admit rate

6am to 2pm	2pm to 6pm	6pm to midnight	Midnight to 6am
12%	70%	16%	2%

Table 3 Outcome of OPT, HEU, and RAND evaluations

# Patients,	Av. ratio			Tukey / bonf		
shift, instance	OPT	HEU	RAND	OPT	HEU	RAND
2, 1, 1	3.206	3.146	3.149	C	C	C
2, 1, 2	2.690	2.945	3.295	B	B	C
2, 3, 1	3.034	3.267	3.183	C	C	C
2, 3, 2	4.292	4.362	4.361	C	C	C
2, 3, 3	3.836	4.141	5.310	B	B	C
2, 3, 4	4.478	5.730	4.469	B	**C**	B
2, 3, 5	4.208	4.496	4.290	C	C	C
2, 3, 6	3.692	3.907	3.949	C	C	C
2, 3, 7	5.871	5.172	6.586	B,C	B	C
2, 5, 1	2.102	2.229	5.511	B	B	C
3, 1, 1	3.069	3.020	3.709	B	B	C
3, 1, 2	3.562	3.564	3.863	C	C	C
3, 1, 3	3.521	3.450	5.412	B	B	C
3, 1, 4	2.712	2.678	2.988	B	B	C
3, 1, 5	4.162	3.770	4.706	B,C	B	C
3, 3, 1	4.007	4.101	4.432	C	C	C
3, 3, 2	6.792	5.584	6.660	**C**	B	C
3, 5, 1	3.201	3.561	3.318	B	**C**	B,C
3, 5, 2	2.439	2.250	5.050	B	B	C
3, 5, 3	2.238	2.225	3.188	B	B	C
4, 1, 1	3.935	4.250	4.790	B	B,C	C
4, 1, 2	2.742	3.131	3.867	A	**B**	C
4, 1, 3	4.213	4.057	7.123	B	B	C
4, 2, 1	2.568	3.758	4.186	A	**B**	C
4, 2, 2	3.499	3.422	3.320	C	C	C
4, 3, 1	2.702	3.043	3.411	A	**B**	C
4, 5, 1	2.657	2.612	4.391	B	B	C
4, 5, 2	2.154	2.165	3.474	B	B	C
4, 5, 3	2.574	2.567	4.402	B	B	C
4, 5, 4	2.341	2.326	4.382	B	B	C
5, 2, 1	4.093	4.080	3.936	C	C	C
5, 2, 2	2.881	2.900	8.267	B	B	C
5, 2, 3	2.946	3.139	3.216	B	B,C	C
5, 2, 4	4.000	4.413	6.720	B	B	C
5, 4, 1	1.972	1.932	4.769	B	B	C
5, 4, 2	1.844	1.888	3.936	B	B	C
5, 4, 3	1.924	1.977	3.650	B	B	C
5, 4, 4	2.084	2.183	5.443	B	B	C
5, 4, 5	2.034	2.041	5.417	B	B	C
5, 5, 1	2.601	2.522	5.110	B	B	C
6, 2, 1	2.635	2.653	3.150	B	B	C
6, 2, 2	3.183	3.749	4.838	B	B	C
6, 2, 3	3.864	3.928	5.059	B	B	C
6, 2, 4	3.309	3.237	3.571	B,C	B	C
6, 4, 1	1.872	1.929	6.645	B	B	C
6, 4, 2	3.017	3.159	10.030	B	B	C
6, 4, 3	1.846	2.388	4.879	A	**B**	C
6, 4, 4	2.468	2.381	8.326	B	B	C
6, 4, 5	2.409	2.743	16.223	B	B	C
6, 5, 1	2.505	2.523	5.762	B	B	C

as newly-admitted patients were chosen randomly. It was assumed five registered nurses work during all the shifts. Admission times of the new patients - for whom assignments have to be determined - were chosen arbitrarily and remained unknown until 15 minutes prior to the actual admit. For simplicity in modeling, it was assumed that there are no patient discharges during the shift. In real life, when discharge occurs, the amount of work load will go down for the nurse who had that patient. It will not affect the relative merit of the nurse-to-patient assignment decisions made by OPT and HEU as discharges impact both policies identically. Hence, it was prefered to ignore discharges in this research.

5.2 Average and spread

The 50 problem instances were simulated on SIMNA with the nurse-to-patient assignments determined by OPT and HEU for each new-patient admit. One thousand scenarios were generated for each problem instance by changing the random seed. The average max-min TADC of the entire shift was determined by averaging max-min TADCs from 1000 scenarios. Assignments from a random policy, referred as RAND, were also simulated to judge whether the "smarter" policies like HEU and OPT yield consistently better results than random assignments. The average max-min TADCs from the 1000 simulation scenarios for each of the 50 assignments are presented in Table 3.

In Table 3, the first column represents the problem instances presented in Table 1. The second column

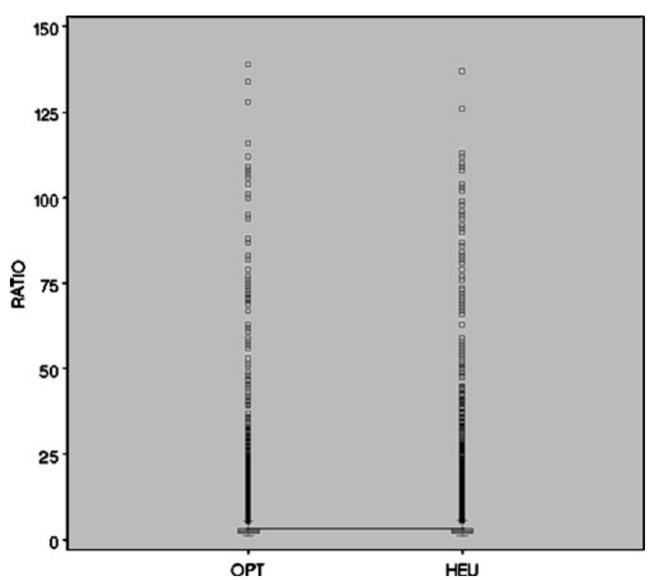

Fig. 1 Boxplots of max-min TADC ratios from OPT and HEU with all 50,000 data points

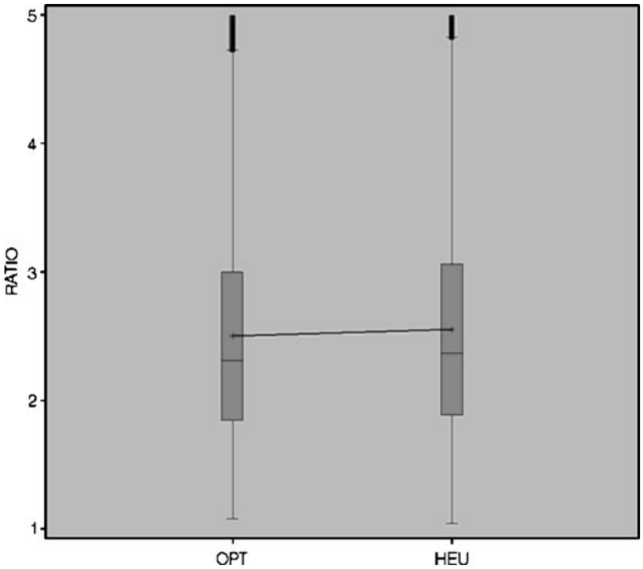

Fig. 2 Boxplots of max-min TADC ratios from OPT and HEU that are less than five

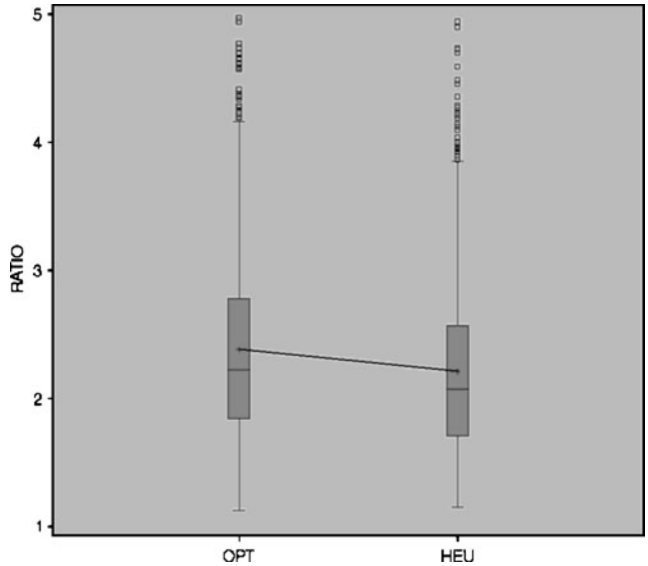

Fig. 4 A boxplot showing HEU win (# New-Patients:3, Shift: 5, Instance: 2)

presents the average max-min TADCs from the three policies evaluated. Ideally, a policy that produces a max-min TADC ratio of one is desired in that it achieves perfect balance in workload among nurses. The policy that yields the smallest average max-min TADC is preferred as it achieves the best possible balance among the three policies. It can be observed that OPT resulted in the least ratio for 30 of the 50 problems, while HEU had 17 smallest ratios. Not surprisingly, RAND managed to be the preferred policy just three times out of the 50 problems. While considering averages to determine the performance of policies, it is important to account for the variability associated with each policy. Boxplots are provided in Figs. 1 and 2 to illustrate the spread of data from the OPT and HEU policies. Because of the outlier scenarios, the scale of boxplots in Fig. 1 is extended leaving it hard for a reader to observe the difference between the plots from OPT and HEU. In Fig. 2, the max-min TADC values higher than five were removed to facilitate the visualization of the boxplots. After removal of outliers, the OPT and HEU policies had, respectively, 45,429 and 45,089

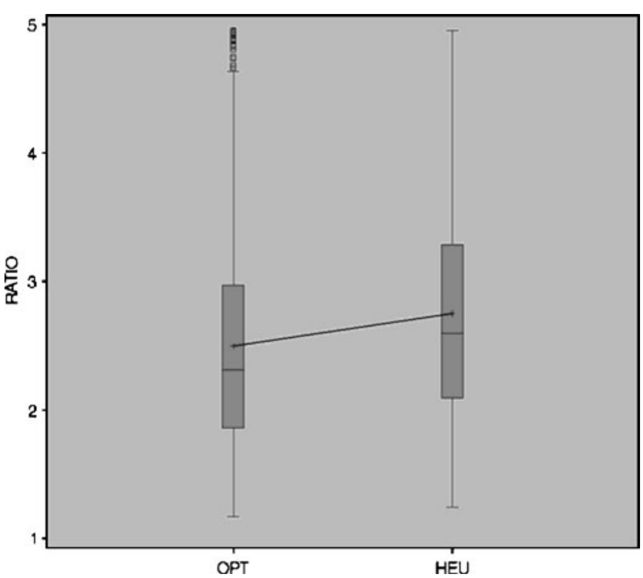

Fig. 3 A boxplot showing OPT win (# New-Patients:4, Shift: 1, Instance: 2)

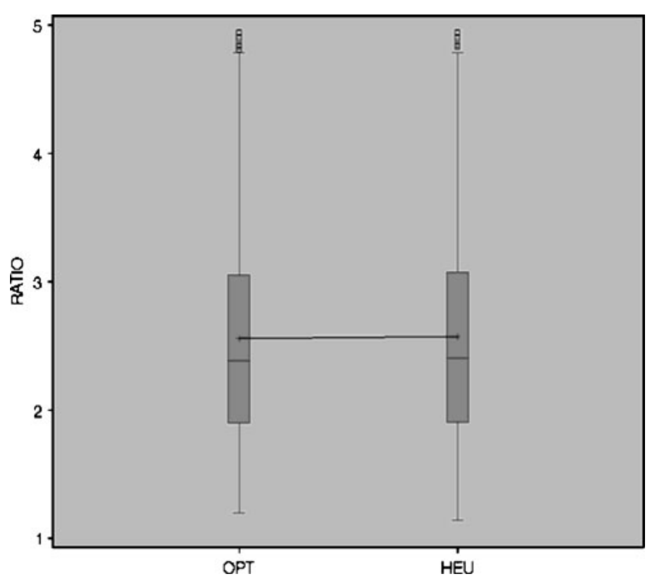

Fig. 5 A boxplot showing tie between OPT and HEU (# New-Patients:5, Shift: 2, Instance: 2)

Table 4 Confidence Intervals for means of HEU-OPT max-min TADC ratios

# Patients, shift, instance	HEU-OPT 95% CI		99% CI		Winning policy 95% CI	99% CI
2, 1, 1	−0.305	0.184	−0.382	0.261	Tie	Tie
2, 1, 2	0.039	0.470	−0.029	0.538	OPT	Tie
2, 3, 1	−0.069	0.536	−0.164	0.631	Tie	Tie
2, 3, 2	−0.485	0.625	−0.660	0.799	Tie	Tie
2, 3, 3	−0.220	0.830	−0.385	0.995	Tie	Tie
2, 3, 4	0.521	1.985	0.291	2.214	OPT	OPT
2, 3, 5	−0.170	0.746	−0.314	0.890	Tie	Tie
2, 3, 6	−0.164	0.595	−0.283	0.714	Tie	Tie
2, 3, 7	−1.385	−0.012	−1.601	0.204	HEU	Tie
2, 5, 1	0.066	0.188	0.047	0.208	OPT	OPT
3, 1, 1	−0.284	0.186	−0.357	0.259	Tie	Tie
3, 1, 2	−0.435	0.438	−0.572	0.575	Tie	Tie
3, 1, 3	−0.497	0.355	−0.631	0.489	Tie	Tie
3, 1, 4	−0.152	0.084	−0.189	0.121	Tie	Tie
3, 1, 5	−0.865	0.080	−1.013	0.228	Tie	Tie
3, 3, 1	−0.336	0.525	−0.471	0.660	Tie	Tie
3, 3, 2	−1.960	−0.456	−2.197	−0.219	HEU	HEU
3, 5, 1	0.151	0.570	0.085	0.635	OPT	OPT
3, 5, 2	−0.263	−0.113	−0.286	−0.090	HEU	HEU
3, 5, 3	−0.072	0.046	−0.090	0.065	Tie	Tie
4, 1, 1	−0.234	0.863	−0.406	1.036	Tie	Tie
4, 1, 2	0.253	0.526	0.210	0.569	OPT	OPT
4, 1, 3	−0.589	0.278	−0.725	0.415	Tie	Tie
4, 2, 1	1.006	1.375	0.948	1.432	OPT	OPT
4, 2, 2	−0.288	0.134	−0.354	0.200	Tie	Tie
4, 3, 1	0.199	0.484	0.154	0.529	OPT	OPT
4, 5, 1	−0.162	0.072	−0.199	0.108	Tie	Tie
4, 5, 2	−0.051	0.074	−0.071	0.094	Tie	Tie
4, 5, 3	−0.096	0.082	−0.125	0.110	Tie	Tie
4, 5, 4	−0.112	0.083	−0.142	0.114	Tie	Tie
5, 2, 1	−0.537	0.510	−0.701	0.674	Tie	Tie
5, 2, 2	−0.128	0.166	−0.175	0.213	Tie	Tie
5, 2, 3	−0.005	0.390	−0.067	0.452	Tie	Tie
5, 2, 4	−0.079	0.905	−0.233	1.059	Tie	Tie
5, 4, 1	−0.087	0.007	−0.102	0.022	Tie	Tie
5, 4, 2	0.003	0.085	−0.010	0.097	OPT	Tie
5, 4, 3	0.008	0.098	−0.006	0.113	OPT	Tie
5, 4, 4	0.042	0.156	0.024	0.174	OPT	OPT
5, 4, 5	−0.046	0.061	−0.063	0.078	Tie	Tie
5, 5, 1	−0.165	0.006	−0.192	0.033	Tie	Tie
6, 2, 1	−0.095	0.130	−0.130	0.165	Tie	Tie
6, 2, 2	0.155	0.977	0.026	1.106	OPT	OPT
6, 2, 3	−0.422	0.550	−0.575	0.703	Tie	Tie
6, 2, 4	−0.269	0.127	−0.331	0.189	Tie	Tie
6, 4, 1	0.012	0.101	−0.002	0.114	OPT	Tie
6, 4, 2	0.028	0.256	−0.008	0.292	OPT	Tie
6, 4, 3	0.489	0.596	0.472	0.613	OPT	OPT
6, 4, 4	−0.158	−0.016	−0.181	0.007	HEU	Tie
6, 4, 5	0.202	0.389	0.172	0.419	OPT	OPT
6, 5, 1	−0.057	0.093	−0.081	0.117	Tie	Tie

max-min TADC ratios, a sufficiently large number of data points to make a comparison of spread. It could be observed that the spread of data in both plots are similar and it would be safe to use average max-min TADC ratio to judge the performance of the policies. Similarly, individual boxplots from each of the 50 instances, not presented here, obtained after removal of five or higher max-min TADC ratios from OPT and HEU had comparable spread. One could well argue that, in reality, it is unlikely to have an imbalance of a magnitude that would result in a value of five or more for max-min TADC ratios. It has to be noted that in all the 50 problems the nurse-to-patient assignments at the beginning of the shift was not balanced and hence, high values for max-min TADCs cannot be ruled out.

Boxplots from three problem instances are provided in Figs. 3, 4, and 5 to illustrate the preferable performances of OPT and HEU in terms of average max-min TADC ratios. In Fig. 3, a typical OPT performance with a lower max-min TADC ratio than HEU is shown. In Fig. 4, a better performance of HEU is shown, while Fig. 5 illustrates an equal performance of OPT and HEU.

5.3 Statistical comparison

In Section 5.2, performances of OPT, HEU, and RAND were analyzed by comparing the average and spread of max-min TADC ratios. In that analysis, it was found that the OPT policy is the most successful, while the RAND policy is the least successful among the 50 problems considered. However, it is necessary to perform statistical analysis to draw a reliable conclusion regarding the difference in performances among the policies. In order to understand the statistical difference among the policies, Tukey and Bonferroni simultaneous pairwise comparison groupings were generated at 0.05 significance level and shown in the last column of Table 3. The distinct groups are represented by alphabets A, B, and C with A and C being the groups with the smallest and the highest means for the max-min TADC ratio, respectively. It has to be noted that if there is only one group (C), it need not be a high mean group. A policy would not be desirable if it falls in a higher mean group while there is at least one other policy in a lower mean group. Both Tukey and Bonferroni grouped the policies identically. From Table 3, it can be observed that 39 times either or both OPT and HEU were in a lower mean group than RAND. Similarly, it can be observed that HEU was out performed by either or both OPT and RAND six times (highlighted by bold), while OPT was outperformed just once by HEU (highlighted by bold). Clearly, from this analysis RAND is the least desirable policy and proves that the “smarter” policies HEU and OPT yield better results. Also, this analysis showed that OPT results are statistically slightly better than HEU.

To further understand the magnitude of the difference between HEU and OPT (HEU - OPT), 95% and 99% confidence intervals (CIs) were constructed in Table 4. In this table, HEU is declared as the winner if both the upper and lower limits are negative. The negative limits indicate a higher max-min TADC ratio from the OPT policy compared to the HEU policy. Similarly, OPT is declared as the winner if both the upper and lower limits are positive. The instances with zero included in the CIs are declared as a Tie. It can be observed from these tables that OPT won 15 out of the 50 instances, while HEU won only four times with 95% CI. The rest of the 31 instances ended as a Tie between OPT and HEU. With 99% CIs, OPT won ten times, while HEU won only twice. The remaining 38 problem instances were declared as tied because CIs include zero. It can be viewed that OPT performed at least as good as HEU in 46 and 48 instances with 95% and 99% CIs, respectively.

Intuitively, assignments obtained from OPT would perform better than HEU when a reliable estimation of future was used while solving for the assignments. From the above analyses, not surprisingly, the OPT policy performed better than the HEU and RAND policies.

6 Conclusions and future work

This research along with [47] makes a significant contribution to the scientific management of nurse-to-patient assignments. It has introduced a tool to evaluate different new-patient nurse-to-patient assignment policies. When new patients are admitted, nurse supervisors often assign the new patient to the nurse who has the least number of patients. This method need not balance the work load of nurses for the entire shift. This research added a feature to SIMNA that helps evaluating nurse-to-patient assignment policies to identify a nurse assignment for the new patient. The enhanced SIMNA model can aid nurse supervisors to make better decisions by simulating different new-patient assignment policies and quantifying the workload measures from them. This research also developed and compared the OPT policy with the HEU policy to make nurse-to-patient assignments when new patients are admitted during a shift. The HEU policy assigned the newly-admitted patient to the nurse who performed the least assigned direct care among all the nurses 15 minutes prior to a new patient admission; while the OPT policy finds the assignment that minimized the difference in workload among nurses for the entire shift from SIMNA. Results from the HEU and OPT policies were compared, and the OPT policy was found to be the better policy. The following are the other promising directions that can be incorporated to this research.

1. HEU vs OPT: It was found from this research that OPT performed better than HEU. Intuitively, HEU′s solution should get better towards the end

of a shift as workload imbalance information from the past is naturally more important and available at the end of the shift. Similarly, with SIMNA approximating the future accurately, OPT should perform relatively much better than HEU at the beginning of a shift than towards the end. Identifying circumstances suitable for OPT and HEU is another interesting area of research. While making a nurse-to-patient assignment decision for a new-patient admit, factors like the time left in the shift, diagnosis, acuity, shift, empty room location, and existing nurse-to-patient assignments could influence the performance of OPT and HEU. To statistically analyze the performance of the assignment policies, an experiment should be designed with diagnosis, acuity, shift, empty room location, existing nurse-to-patient assignments, and time left in the shift as factors and max-min TADC ratio as the response. With 19 diagnoses codes, four acuity levels, five possible shifts, at least eight time periods in a shift, and 26 patient rooms, the experiment will result in more than 79,040 treatments. To perform such an analysis efficiently and reporting results from them would be an interesting research by itself.

2. "Time Period-Action Q-Factors" method: In this research, a brief discussion about the potential use of the Q-Factors methods was provided especially in circumstances when a simulator is available. However, the existing algorithms of the Q-Factors method is not feasible to implement for the nurse-patient assignment problem because the number of state-action pairs is huge. It will be interesting to explore the possibility of having the Q-Factors for arrival-action pairs instead of state-action pairs. This approach will reduce the number of Q-Factors significantly. It should be noted that with stochastic arrivals, it is still difficult to update all the arrival-action pairs accurately within a reasonable number of simulation runs. For example, the first arrival time in a simulation run is likely to be different from another first arrival simulated in a different simulation run. To tackle this issue, the shift can be divided into smaller time periods to get the Q-Factors for each period-action pair. The actions in this research are to assign the newly-admitted patients to nurses. There is no action required in a time period if there is no new-patient admits. Therefore, with the "time period-action Q-Factors", the number of Q-Factors would be equal to the number of time-periods times the number of nurses. For example, for an eight hour shift broken into one hour periods with five nurses working, there would be just forty Q-Factors. As mentioned earlier, it would take just three minutes to run one thousand scenarios, and it is possible to update the Q-Factors for real time decision making using the proposed "time period-action Q-Factors" method.
3. Optimization: Exploring the applicability of simulation-optimization methods, such as in [1], and [18], is also an interesting topic for future research. The traditional simulation-optimization methods, in general, use an approximated value for the gradient of the simulation. The dynamics of SIMNA in [47] are captured by the static tree structures from CART. Extracting the gradient of the simulation from CART and using it for optimization is potentially feasible and worth exploring.
4. Patient Discharge: It was assumed that there are no patient discharges during a shift for simplicity in modeling. However, it is common to have discharges during a given shift. Incorporating patient discharges in future will enhance practicality of SIMNA's usage in hospitals.

Acknowledgements This research was supported by the Robert Wood Johnson Foundation grant number 053963. The first author gratefully acknowledges summer research grant from Steven L. Craig School of Business which enabled the successful completion of this research. The authors thank three anonymous reviewers and the editor-in-chief for a constructive review process and useful comments.

References

1. Atlason J, Epelman MA, Henderson SG (2004) Call center staffing with simulation and cutting plane methods. Ann Oper Res 127:333–358
2. Baker RL, Buckley-Behan DF, Goss K, Punnakitikashem P, Turpin P, Rosenberger JM (2010) Phase I: creating an electronic prototype to generate equitable hospital nurse-to-patient assignments. Computers, Informatics, Nursing 28(1):57–62
3. Bard J, Purnomo HW (2005) Preference scheduling for nurses using column generation. Eur J Oper Res 164:510–534
4. Bellman RE (1957) Dynamic programming. Princeton University Press
5. Bertsekas DP (2001) Dynamic programming and optimal control. Athena Scientific, Belmont, Massachusetts
6. Bertsekas DP, Tsitsiklis J (1996) Neuro-Dynamic programming. Athena Scientific, Belmont, Massachusetts
7. Bigbee JL, Collins J, Deeds K (1992) Patient classification systems a new approach to computing reliability. Appl Nurs Res 5:32–53
8. Cameron JM, Baraff LJ, Sekhon R (1990) Case-mix classification for emergency departments. Med Care 28:146–158
9. Chen VCP, Ruppert D, Shoemaker CA (1999) Applying experimental design and regression splines to high dimensional continues state stochastic dynamic programming. Oper Res 47:38–53

10. Childers L (2008) Bay area hospitals, colleges join forces to address nursing shortage. Bay Business times (accessed - September 2008 on eastbay.bizjournals.com)
11. Cordes CL, Dougherty TW (1993) A review and an integration of research on job burnout. Acad Manage Rev 18(4):621–656
12. Cullen A (1995) Burnout: why do we blame the nurse? Am J Nurs 95(11):22–28
13. Dumas, MB (1985) Hospital bed utilization: An implemented simulation approach for adjusting and maintaining appropriate levels. Health Serv Res 20:43–61
14. Duraiswamy N, Welton R, Reisman A (1981) Using computer simulation to predict icu staffing needs. J Nurs Adm 11:39–44
15. Evans GW, Gor TB, Unger E (1996) A simulation model for evaluating personnel schedules in a hospital emergency department. In: Proceedings of the 1996 Winter Simulation Conference Coronado, California, USA
16. Fries BE, Cooney LM (1985) Resource utilization groups: a patient classification system for long-term care. Med Care 23(2):110–122
17. Fu MC (2002) Optimization for simulation: theory vs practice. INFORMS J Comput 14(3):192–215
18. Fu MC, Hu JQ (1997) Conditional Monte Carlo: gradient estimation and optimization applications. Kluwer, Norwell, Massachusetts
19. Garrett C (2008) The effect of nurse staffing patterns on medical errors and nurse burnout. AORN J 87(6):1191–1203
20. Gosavi A (2004) Reinforcement learning for long-run average cost. Eur J Oper Res 155:654–674
21. Gosselin G (2008) Michigan provides $5m for nurse-teacher training. Oakland Business Review (accessed - September 2008 on mlive.com)
22. Griffiths P (2009) Rn + rn = better care? what do we know about the association between the number of nurses and patient outcomes? Int J Nurs Stud 46:1289–1290
23. Hasman A, Wiersma R, Halfens R, Algera-Osinga JT (1993) Evaluation of a patient classification system for community health care. Int J Bio-med Comput 33:109–118
24. Hurrion RD (2000) A sequential method for the development of visual interactive meta-simulation models using neural networks. J Oper Res Soc 51:712–719
25. Jacobs J (2008) East bay nursing programs get injection of state funds. East Bay Business Times (accessed - September 2008 on bizjournals.com)
26. Jaumard B, Semet F, Vovor T (1998) A generalized linear programming model for nurse scheduling. Eur J Oper Res 107(1):1–18
27. Jun JB, Jacbson SH, Swisher JR (1999) Application of discrete event simulation in health care clinics: a survey. J Oper Res Soc 50(2):109–123
28. Kempson D (2008) Scheduling case history: the needs of the many. Health Manag Technol Oct:22–25
29. Kim SC, Horowitz I, Young KK, Buckley TA (2000) Flexible bed allocation and performance in the intensive care unit. J Oper Manag 18(4):365–385
30. Klein RW, Dittus RS, Roberts SD, Wilson JR (1993) Simulation modeling and health-care decision making. Med Decis Mak 13(4):347–354
31. Kreke J, Schaefer AJ, Angus D, Bryce C, Roberts M (2002) Incorporating biology into discrete event simulation models of organ allocation. In: Proceedings of the 2002 Winter Simulation Conference, San Diego, California, USA
32. Landro L (2008) A hospital treatment room with a view. The Wall Street Journal (accessed - September 2008 on reporternews.com)
33. Law AM, Kelton WD (2001) Simulation modeling and analysis. McGrawHill, New York
34. Law AM, McComas MG (2002) Simulation-based optimization. In: Proceedings of the 2002 Winter Simulation Conference, Piscataway, New Jersey, USA
35. Lim T, Uyeno D, Vertinsky I (1975) Hospital admission systems: a simulation approach. Simul Gaming 6:188–201
36. Lundgrén-Laine H, Suominen T (2007) Nursing intensity and patient classification at an adult intensive care unit (icu). Intensive Crit Care Nurs 23:97–103
37. McHugh ML (1989) Computer simulation as a method for selecting nurse staffing levels in hospitals. In: Proceedings of the 1989 Winter Simulation Conference, Washington, D.C., USA
38. Miller HE, Pierskalla WP, Rath GJ (1996) Nurse scheduling using mathematical programming. Oper Res 24(5):857–870
39. Mullinax C, Lawley M (2002) Assigning patients to nurses in neonatal intensive care. J Oper Res Soc 53:25–35
40. Olafsson S, Kim J (2002) Simulation optimization. In: Proceeding of the 2002 Winter Simulation Conference, Piscataway, New Jersey, USA
41. Punnakitikashem P, Rosenberger JM, Behan DF (2008) Stochastic programming for nurse assignment. Comput Optim Appl 40:321–349
42. Robinson S (2004) Simulation: the practice of model development and use. Wiley, Chichester, UK
43. Roser MA (2008) Austin-area hospitals: Nursing shortage easing. Austin American-Statesman (accessed - September 2008 on statesman.com)
44. Shechter SM, Bryce C, Alagoz O, Kreke JE, Stahl JE, Schaefer AJ, Angus D, Roberts M (2005) A clinically based discrete event simulation of end-stage liver disease and the organ allocation process. Med Decis Mak 25(2):199–209
45. Smith EA, Warner HR (1971) Simulation of a multiphasic screening procedure for hospital admissions. Simulation 17:57–64
46. Storlie FJ (1979) Burnout: the elaboration of a concept. Am J Nurs 79(12):2108–2111
47. Sundaramoorthi D, Chen VCP, Rosenberger JM, Deborah BF (2009) A data-integrated simulation model to evaluate nurse-patient assignments. Health Care Manage Sci 12(3):252–268
48. Vericourt Fd, Jennings OB (2006) Nurse-to-patient ratios in hospital staffing: a queuing perspective. http://faculty.fuqua.duke.edu/%7Efdv1/bio/ratios3.pdf (accessed July 2006)
49. Vu P (2008) States work to avert nurse shortage. Stateline.org (accessed - September 2008)
50. Williams S, Robert C (2006) Emergency department patient classification systems: a systematic review. Accid Emerg Nurs 14:160–170
51. Zenios SA, Wein LM, Chertow GM (1999) Evidence-based organ allocation. Am J Med 107(1):52–61