applied sciences

MDPI

Article

Data Analytics and Its Advantages for Addressing the Complexity of Healthcare: A Simulated Zika Case Study Example

Lily Popova Zhuhadar * and Evelyn Thrasher

Information-Systems Faculty, Western Kentucky University, Bowling Green, KY 42101, USA; evelyn.thrasher@wku.edu

* Correspondence: lily.popova.zhuhadar@wku.edu

Received: 25 December 2018; Accepted: 26 March 2019; Published: 29 May 2019

check for updates

Abstract: The need to control rising costs in healthcare has led to an increase in the use of data analytics to develop more efficient healthcare business models. This article discusses a simulation that uses data analytics to minimize the number of physicians and nurses needed in healthcare facilities during a crisis situation. Using a hypothetical emergency scenario, the hospital uses a healthcare analytical system to predict the necessary resources to govern the situation. Based on historical data regarding the flow of patients through the facility, a discrete-event simulation estimates resource scheduling and the resulting impact on both wait times and personnel demand. Furthermore, the value of multiple replications for discrete-event simulation models is discussed and defined, along with factors that enable greater control of multiple design points with this simulated experiment. The results of this study demonstrate the value of simulation modeling in effective resource planning. The addition of only a single doctor significantly reduced predicted wait times for patients during the crisis. Further, the findings support the use of data analytics and predictive modeling to mitigate rising healthcare costs in the United States through efficient planning and resource allocation.

Keywords: healthcare analytics; data analytics; healthcare; information systems; hospital data analytics; healthcare crisis; discrete-event simulation; data analysis simulation; simulation

1. Introduction

The United States has embarked on significant healthcare reform within the last 10 years, specifically with the incorporation of the Patient Protection and Affordable Care Act of 2010 (ACA) [1]. It is no secret that the healthcare system in the United States is complex, multifaceted, and filled with human error [2]. A multitude of regulations have been used to control the healthcare system and its accelerated costs in the United States and yet these regulations (i.e., payments, penalties, and certifications) can lead to unintended consequences because the complexity surrounding the healthcare system both in design and in its non-linear unpredictable nature [1]. In order to guide future policies and reduce unanticipated consequences, those individuals who influence regulation need to understand healthcare as a complex system and apply principles of complexity science (component parts act non-linearly over multiple scales) to achieve the goals regarding healthcare reform [1]. The output of complex systems is dynamic and behaves differently depending on initial conditions, feedback, and error (skill-based, knowledge-based, and rule-based) [1,2]. The healthcare system is comprised of networks of components (hospitals, rehabilitation homes, patient homes, clinics, etc.) that interact nonlinearly on different scales and often produce unintended consequences, such as adverse drug reactions, infections, and re-hospitalization. Essentially, the greater the regulation created to control the behavior, the more the system may deviate from the desired outcome [1]. Methods in place to use disease protocols and financial levers to manage healthcare are significantly flawed and will continue to lead to these unintended consequences [1]. For example, pay-for-performance and value-based payment models aiming to improve hospital care at lower costs may actually encourage treatment that is too aggressive without thought to life expectancy or adverse effects since these models are not centered on mortality rates or Medicare spending [1]. Viewing healthcare as a complex system with several intrinsic properties (i.e., emergent self-organized behavior and dependence on simple rules) can be more effective for influencing its dynamic behavior, thereby guiding the outcomes in a more desired direction.

As noted in past research, complex systems can transform without a grand plan to do so. For example, in the twentieth century, agriculture changed from disorganized, costly, and inefficient to a coherent, productive, and cost-effective system of farming. Those changes have been attributed to a cyclical process of experimentation, measurement, and learning. The same opportunity now exists within the healthcare industry to enable the emergence of a more functional system via innovations and waivers supported by the ACA [1]. This will require effective interactions among patients, families, physicians, and other healthcare providers to reach mutual objectives established within the ACA. The system should provide a global payment for patient care across settings and allow clinicians and healthcare centers to self-organize in a way that better achieves the ACA objectives. The U.S. healthcare system will need to continue to depart from a mechanical, regulatory approach and move toward the complex systems approach, which promotes self-organization [1,3].

The use of analytics in healthcare has increased rapidly and has been shown to reduce costs and increase efficiencies in both the business and clinical areas [4,5]. Research has noted that incorporating solutions using data analytics has highlighted some of the difficulties in providing healthcare solutions including uneven data loads, having heterogeneity in applications, ranges of user expectations that are not consistent and latency sensitivity for validating models [6]. However, improvements in algorithms, data and text mining, and the use of big data offer tremendous potential to reduce much of the complexity that surrounds decision making in healthcare [4,5]. For example, recent research has shown that using a discrete event simulation model to assist in data driven decision making for populations with dementia in Australia generated valid estimates of dementia as well as future resource requirements needed based on those estimates [7]. To this end, the current study demonstrates the use of discrete-event simulation (DES) as a valuable tool to aid in mitigating the complexity of resource planning and prediction in healthcare required for a Zika Outbreak. The current study has created a hypothetical scenario in which a Zika Virus outbreak occurs around a local health clinic, using a healthcare analytical system to predict the resources needed to control the situation. More specifically, based on previous data such as the average amount of time a patient spends waiting to see a nurse and/or physician, a DES of resource scheduling is estimated using SAS Simulation Studio.

In the sections that follow, we present several healthcare challenges and opportunities that have emerged in the United States because of changes in regulations, implementation of electronic health records, cloud based storage, advances in data mining and business intelligence, and the innate complexities of the healthcare service industry. The DES is described, and the results are presented. Finally, conclusions of the study and challenges for future research are discussed.

2. Background

2.1. Complexity of Healthcare and Data Management

U.S. healthcare is a complex system, both from a clinical perspective and from a business perspective [4]. Governmental regulations, insurance policies, billing and payment agreements, and internal entity policies have certainly led to a complex business organization inside every healthcare entity. Adding to this system the uniqueness of every patient, creates an extremely complex clinical process as well. One product of the complex healthcare industry are the large volumes of data captured, much of which have the potential to improve performance on both the business and clinical side if used appropriately. The management of this increasing amount of healthcare data is crucial for having efficient healthcare management and decision making [4,6]. Therefore, information systems are needed to organize and catalog the massive amounts of information in a way that makes the data easily accessible and highly efficient for anyone using a simple search criteria [4–6]. The evolution of information technology, particularly related to data analytics, has the potential to reduce some of the complexity of healthcare, however, researchers must be willing to explore that potential to its fullest [6].

Part of the issue with healthcare information systems is the diversity across entities which creates significant obstacles for successful implementation [8]. Some of those obstacles include system failures, costs, fear regarding information confidentiality, security, and privacy, changes in work processes, and a lack of acceptance by clinicians [8]. For example, clinicians often stress their dislike of such systems, expressing concerns around technical (hardware/software) issues, utility (cannot meet the task needs) issues, and usability (is rejected/not accepted by user) issues. More specifically, many clinicians worry that patients will perceive a poorer quality of care if patient wait times increase due to system issues. In addition, they worry that the quality of their work may decline if the technology interferes with their established routines or negatively impacts their relationships with other healthcare entities [9]. There is no clear consensus on how to address these concerns within the healthcare information system because of the complexities of the issues and the diversity among contexts of use [8]. Implementing health information systems that have the same type of rigorous software development and engineering methods found in other 'critical environments', such as air traffic control or railway signaling, may serve as an important step toward eliminating these obstacles [8].

Another barrier to healthcare information systems is the cost associated with implementation and maintenance [8]. Many researchers note the difficulty of demonstrating a return on investment for these systems. The fact that information technology is not a low risk capital investment further highlights the other noted risks, including financial costs to the organization, uncertain cost benefit tradeoffs, uncertainty in the presence of confidentiality, security and privacy fears, and concerns regarding the level of balance between security/privacy/confidentiality and having access that is acceptable to consumers [8]. Many clinicians have found paper records more convenient than electronic records, stating that electronic integration has not adequately allowed for efficient use of the data within the systems [9]. It is imperative to identify tools and systems that can overcome these obstacles and earn the trust of clinicians, while also reducing healthcare costs.

2.2. *Ecology of Medical Care Framework*

There have been major changes in medicine, financing, and the organization of healthcare since 1961 and some of these changes (i.e., new medications, technological innovations, increased expenditures, managed care, and changes in medical work force) have altered the ecology of medical care [10]. Substantial improvements in the collection and reporting of data on healthcare in the United States have altered the ecology of medical care. However, despite these changes, healthcare relationships have remained stable for 40 years. The ecology of medical care has provided a framework for thinking about the organization of healthcare, medical education, and research [10].

Considering a hypothetical population with 1000 adults at risk, in a typical month 750 will report one or more illnesses or injuries, 250 adults will consult a physician at least once, nine adult patients will be admitted to a hospital, five adult patients will be referred to another physician, and one adult patient will be referred to a university medical center [10]. The Medical Expenditure Panel Survey and the 1996 National Health Interview Survey further indicated that a large portion of the United States noted health problems [10]. Furthermore, only 25% will visit a physician's office, one third will visit a complementary or alternative medical care provider, and less than one person in 1000 will be admitted to an academic-medical-center hospital. These results have changed little in 40 years, even when considering children in the analysis [10].

The ecology model highlights the need for comprehensive healthcare information systems that span all sites of care and identifies the potential missed opportunities when limiting quality and safety programs to hospitals. The model shows the need for alternative types of research laboratories (i.e., practice-based research networks) that enable the study of patients where they receive their care. Yet, the model does not establish causal pathways and may be misinterpreted due to its nested appearance [10]. Simply put, to improve patient care and reduce costs, more is needed to address the complexity of healthcare than the ecology of medical care framework can address [6].

2.3. Patient-Centered Healthcare Analytics

Patient centeredness may be necessary for developing better healthcare models, as it differs from the biomedical model in terms of key dimensions (bio-psychosocial perspective, patient as a person, sharing power and responsibility, the therapeutic alliance, and the doctor as a person), each representing a particular aspect of the doctor-patient relationship [4,11]. The biomedical model was noted as the conventional way of 'doing medicine' by dictating how medicine was practiced. The biomedical model exerts heavy influence on the content and style of the doctor-patient relationship, which notes that patients' reports of illness are taken to indicate the existence of disease processes. This, in turn, dictates that a clinical method focused on identifying and treating standard disease entities be used so that the patient's illness is reduced to a set of signs and symptoms investigated and interpreted within a positivist biomedical framework [11]. Using the patient-centered model, the bio-psychosocial perspective allows for broadening the explanatory perspective on illness to include social and psychological factors such that health bodies/biographies are different from one another. The idea that the patient is a person is concerned with understanding the individual's experience of illness and that patients cannot be characterized completely by a diagnostic label. Instead, full understanding comes from the patient's presentation and the doctor's understanding that the patient is a unique personality [11]. In looking at patient-centeredness, sharing power and responsibility is a key element regarding the doctor-patient relationship. By interrupting the patient's voice with response constraining questions, it strips away the personal meaning of illness; so [11] argues for a mutual participation model. This model addresses the benefit of mutual participation, thus making it often therapeutic for the patient while also recognizing the doctor as a person.

2.4. Data Mining and Knowledge Discovery in Databases

Data mining and knowledge discovery in databases (KDD) have gained considerable attention in both research and media [12]. The KDD field in particular is concerned with developing methods and techniques that make sense of big data, such as addressing the problem of mapping voluminous low-level data into more understandable, useful formats, including short reports, easy descriptions of the modeling process, or predictive models for estimating the value of future cases [12]. At the root of the process are data mining methods for pattern discovery and extraction. Since databases are increasing in terms of the number of records or objects in the database and the number of fields or attributes for an object, databases with exponential numbers of objects are becoming common, leading to the obsolescence of manual data analysis [12]. For example, data cleaning techniques like the merge-purge system have been used to identify duplication of welfare claims for the U.S. government [13]. Furthermore, IBM's Advanced Software has helped NBA coaches organize and interpret data from games [12]. An increasingly important type of discovery is one based on the use of intelligent agents to navigate through information rich environments. With the advent of the Internet, these systems ask the user to specify a profile of interest and search for information across all relevant public domains and proprietary sources [12].

KDD has evolved through the intersection of research from machine learning, pattern recognition, statistics, artificial intelligence, knowledge acquisition, and data visualization with a goal to extract high level knowledge from low level data in very large data sets [4]. The data-mining component of KDD relies on the techniques used in machine learning, pattern recognition, and statistics to discern patterns from the data. Yet, KDD is different from machine learning and other related fields because it focuses on the overall process of knowledge discovery from data, such as how the data are stored and accessed and how algorithms are scaled for massive data sets [12]. KDD can be used in healthcare to model and automate the data analysis process and the art of hypothesis selection [12]. By using representative examples from the database to approximate a model (predictions on new examples derived from the proprieties of similar examples in the model whose prediction is known), more predictions can be made regarding the model.

2.5. Business Intelligence Systems

The demand for Business Intelligence (BI), a popular tool in KDD, has been increasing since its inception roughly 40 years ago [14]. BI systems combine operations data with analytical tools to present complex and competitive information to planners and decision makers. The overall objective is to improve the timeliness and quality of inputs to the decision process [14]. The emergence of the data warehouse as a repository, advances in data cleaning, and increased capabilities of hardware and software combine to create a richer BI environment. Yet, research is limited with regard to how these systems can help healthcare [4]. As Negash [14] notes, often BI systems refer to on-line decision making, instantaneous responses, and the shrinking time frame needed for an answer, making it more proactive than previous decision-making tools. The essential components of smart BI systems are real time data warehousing with big data collection, data mining, automated anomaly and exception detection, proactive alerts for automatic recipient determination, automatic learning and refinement, and geographic information systems [14]. BI systems are used extensively in corporate performance management, optimization of customer relations and monitored business activity, and decision support. BI is not an inward-looking function but a natural outgrowth of a series of previous systems designed to support decision making. BI simply pulls from other systems using both structured and semi-structured data. Semi-structured data is data that doesn't fit into relational/flat files, such as emails, letters, phone calls, web pages, etc.

Data volume is a serious issue for consideration in any system. Google estimates that the Net is doubling in size every 8 months [14]. BI for the masses has shown a new class of emerging analytic tools that serves the broader population. In order to deal with more complex documents, roughly 80% of which are semi-structured in nature, these new tools provide reports and analysis capabilities at all levels of an organization with easy creation and consumption of reports, secure delivery of information, and Internet browsers that are user friendly [14].

According to Bonney [15], healthcare clinical datasets provide an excellent environment where both structured and unstructured datasets can be analyzed to present useful information. For instance, Electronic Health Records (EHR) are repositories of information regarding the health status of a subject of care. EHRs can be processed by a computer, stored and transmitted securely, and are accessible by multiple authorized users. Additionally, Electronic Medical Records (EMR) are computerized medical records designed for collecting, storing, sharing, and displaying patient information. Along with the increasing demand for EHRs and EMRs, there is also a growing need to use data mining technologies to extract quality data and inference rules from the information stored in these datasets to provide real time decision support. Since BI has emerged as the technology with the most potential to operationalize the repository content of EHRs and EMRs to improve the quality and safety of healthcare delivery, its integration into the healthcare field is a necessity. Healthcare providers stand to gain significantly from the use of BI because of its strength for data mining and knowledge discovery and its advantages for analyzing both structured and unstructured data [15]. BI has the capability to transform information into knowledge and put the 'right information in the hands of the right user at the right time to support a decision' [15]. Specifically, BI technologies are used to shorten the time lag between data acquisition and decisions, a capability that is of great importance in the healthcare industry [15].

The EHR can provide caregivers the relevant information about every patient, encouraging the sharing of medical knowledge through computer-assisted clinical decision support and healthcare delivery. The standardization of healthcare using systems such as the EHR will provide the foundation for future research and increasingly vital quality and regulatory reporting [15]. While the EHR has emerged as an integrated healthcare information system, there are challenges to providing consistent data in patient-centered care that affect the appropriate integration of the BI technology [15]. It is difficult to integrate the BI technology with the EHR without considering the relevant contextual issues, such as data roles and governance, which are not adequately addressed in the EHR. Establishing data governance involves the creation and management of the organizational structures, policies, and processes needed to define, control and ensure the quality of the data [15]. Other issues come in the form of data integrity, data semantics and data security; for example, maintaining the validity and confidentially of the data is critical to ensuring the adoption of the EHR [15]. Yet, even though challenges persist in the implementation of BI technology for the EHR in clinical practice, the platforms are becoming more widespread and sophisticated as many corporate organizations continue to integrate BI with records management solutions [15].

2.6. Case Study—Netherlands

Spruit, et al. [16] completed a case study of BI systems using the KDD approach in the Netherlands asking "How can knowledge discovery techniques support Dutch long-term care institutions to manage their internal organization?" The researchers structured their knowledge discovery process on exploratory research according to the steps of the Cross Industry Standard Process for Data Mining (CRISP-DM) to identify patterns in incident information, patterns in risk assessment information, the relationships between the two, and to identify and predict Care Intensity Package combinations [16]. From a knowledge discovery process perspective, they observed that some data mining goals were not achieved due to the complexity of the data or the lack of data standardization. Further, based on the results of their analysis, it was clear that predictive models are not yet directly valuable for care institutions because the predictions are too complicated and strongly dependent on exogenous factors and contingencies [16]. The authors note that this is due to political choices and policy considerations that have led to changes in existing laws and the introduction of new laws and regulations, which have direct consequences for the information requirements posed on care institutions and the data they need to collect [16]. In summary, their research showed that by further exploring knowledge discovery techniques based on large datasets and guided by representative information needs, researchers can contribute to the domain of long-term care to better manage both quality and cost of care as potentially a next step towards healthcare business intelligence [16].

3. Materials and Methods

3.1. Advantages of Simulation Modeling

Researchers use simulation modeling in situations where experiments are not possible because either the process simply does not exist or it is not cost effective to perform in a real-world setting [17]. In some cases, traditional mathematical methods such as queuing theory, differential Equations, and linear programming have been and can be used successfully. However, the complexity and randomness found in healthcare make such methods inadequate because these methods cannot represent the system by just deriving an analytical model [17]. Crisis management and the healthcare model call for simulation modeling, as it is the preferred method for handling the complexity found in real-world settings.

An event history is a record of when events occurred to a sample of individuals [18]. While event histories are ideal for studying causes of events, they have two varying explanatory variables—censoring and time—that create difficulties for standard statistical practices [18]. As most methods assume that time is measured as a continuous variable (having non-negative values), in some situations discrete-time models and methods may be more appropriate [18]. Some events can occur at any point in time, but available data only records the interval of time within which each event actually occurs [18].

Computer models can be dynamic or static, depending on the criteria set for the event. In case studies involving emergency room visits, as the state changes, countable points in time can be measured, such as arrival, completion of exam/triage, departure, etc. This falls under discrete event simulation (DES) [4]. DES is the process where behavior found in a complex system is coded as an ordered sequence of clearly defined events [19]. In this context, the event includes a change in the system's state at a specific point in time (i.e., arrival to triage). When a simulation model has random inputs, the output is also random; so assumptions cannot be based on one simulation, but rather replications that include multiple runs. This modeling can be explained as follows: X_1, X_2, X_i … are each a random process representing the output from a single simulation run. By conducting a fixed amount of simulation runs or replications with a length of m observations, the parameter of interest is the average across each simulation. As each measurement of X_i is not independent from the next and each are identically distributed, this will give a sample from different random variables. Since the set of averages across these simulations are also independent of each other and identically distributed, the mean and variance of the sample can be found.

Current studies that includes DES include Alkhaldi, et al. [4] which demonstrated the use of DES in healthcare for the reduction of unplanned hospital readmissions. In their study, discharged patients and care providers were inputs into the system. Symptom communication was key to the successful use of the DES, and it allowed for better symptom monitoring and behavior analysis [4]. More recently Standfield, et al. [7] used a DES combining models of epidemiology of dementia with models of health and aged care sectors to estimate accurate future demands of dementia related illness on their healthcare system. This study highlighted the growing demand for resources as well as provided support that individual patient simulations (IPS) methods such as DES allowed for greater realistic depiction of the complex process of unique individuals with compounding complex health histories [7].

3.2. Case Study

As with previous analyses, this study explores DES as one tool that holds promise for addressing some of the complexity of the healthcare industry. Though many factors contribute to this complexity, this study focuses specifically on resource scheduling in the event of a crisis situation. Patient wait times were noted as an outcome directly impacted by the processes of care that can influence the quality and cost of care [4]. In an industry plagued with high costs and inefficiencies, a crisis can greatly exacerbate those costs and inefficiencies, more accurate resource need predictions are required for planning to address the requirements.

Since an urgent care clinic is equipped to treat most individuals including both children and adults, it was used for this case study. As shown in Figure 1, on a typical day, 6 child patients arrive to the clinic per hour ($\mu = 10$); and 8 adult patients (including pregnant women) arrive per hour ($\mu = 7.5$). All patients must wait to see a nurse for an initial vital sign evaluation (triage specialist). The time is triangularly distributed in minutes. For a patient coded as a child the minimum wait time was 5 min, the mode (most occurring time) was 8 min, and the maximum was 12 min. For a patient coded as an adult, the wait times were noted as a minimum of three minutes, mode of five minutes, and a maximum of 9 min. Afterward, there is an additional waiting period for children to see a pediatrician and for adults to see a doctor. The wait time to see a pediatrician had a minimum of 10 min, mode of 14 min, and maximum of 25 min. For adults waiting to see a doctor, the minimum was 8 min, the mode was 12 min, and the maximum was 20 min.

For the purposes of this study, the clinic is operating with two nurses, three pediatricians, and two doctors. The clinic's hours of operation are from 8:00 a.m. to 5:00 p.m., but it does not actually close until all patients have been treated.

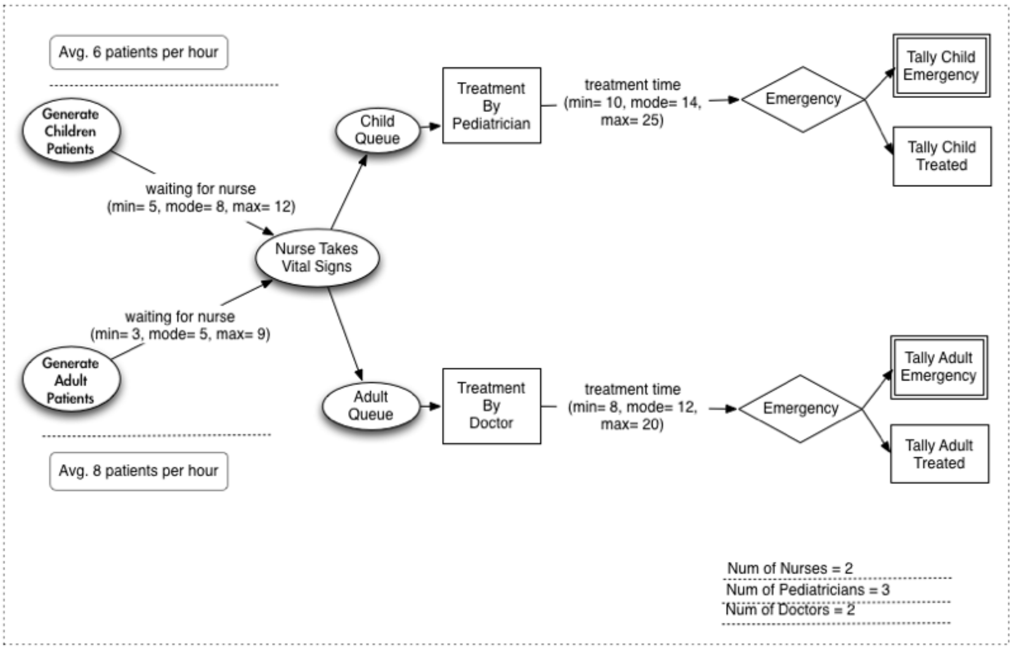


Figure 1. Process flow diagram for urgent care clinic.

3.3. Research Approach (Discrete Event Simulation)

Beginning with a PFD, the first step is to formulate the problem and define the simulation objectives (see Figure 1). The second step applies the conceptual model to a computer model which involves data collection and input analysis, along with the process of implementing entity and information flows. Once the experiment is fully designed, the appropriate models are analyzed. If there are multiple input factors, statistical experimental design techniques (such as full or fractional factorial designs) can be used. Lastly, the data generated from the model is interpreted and used for testing hypotheses, drawing inferences, and/or evaluating alternatives.

3.4. Discrete Event Simulation (DES)

Classic statistical methods are based on independent, identically distributed observations; however, in a simulation run, individual wait time data are not independent or identically distributed. For example, the first patient coming into the clinic has a wait time of zero to check-out; the second patient likely has a wait time greater than zero; and the third person has yet a different wait time distribution from the first two. Furthermore, when a simulation model contains random inputs, the simulation output is also random. Therefore, assumptions concerning system performance cannot be based on a single simulation, but rather on an average across multiple replications. This modeling can be represented as follows: Let $\{X_1, X_2, \ldots, X_m\}$ be a random process representing the output from a single simulation run (i.e., the first row in the matrix below). By conducting n runs of the replications, each a length of m observations, the performance parameter of interest is the average across all runs.

$$
\begin{array}{ccccccccc}
x_{1,1} & x_{1,2} & x_{1,3} & \cdots & \mathrm{x}_{1,\mathrm{k}+1} & \cdots & \mathrm{x}_{1,m} & \mathrm{X}_1 & \leftarrow & \text{Run 1} \\
x_{2,1} & x_{2,2} & x_{2,3} & \cdots & x_{2,k+1} & \cdots & \mathrm{x}_{2,m} & \mathrm{X}_2 & \leftarrow & \text{Run 2} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \\
x_{n,1} & x_{n,2} & x_{n,3} & \cdots & x_{n,k+1} & \cdots & \mathrm{x}_{n,m} & \mathrm{X}_n & \leftarrow & \text{Run n} \\
 & & & & & & & \uparrow & & \\
 & & & & & & & \text{Average Across} & & \\
 & & & & & & & \text{Single Run} & &
\end{array}
$$

Because $\{X_1, X_2, \ldots, X_m\}$ are not independent and identically distributed, this process will give different random variables in a single sample from m. In addition, the set of averages across the n

replications is independent and identically distributed and the sample mean and sample variance can be computed using *the following Equations:*

$$\overline{\overline{X}} = \frac{\sum_{i=1}^{n} \overline{X_i}}{n}, \quad (1)$$

$$s^2 = \frac{\sum_{i=1}^{n} \left(\overline{X_i} - \overline{\overline{X}}\right)^2}{n-1}. \quad (2)$$

The classic $(1-\alpha)100\%$ confidence interval for the parameter of interest can be computed as follows:

$$\overline{\overline{X}} \pm t_{\frac{1-\alpha}{2,n-1}} \frac{S}{\sqrt{n}}. \quad (3)$$

Finally, the classical confidence interval formula, $t_{1-\alpha/2}$, n−1 is defined as the 1− $\alpha/_2$ quantile of the variable's *t*-distribution.

3.5. Implementing Discrete Event Simulation Models

The first step in mapping the PFD into SAS Simulation Studio is to create two separate blocks (arrivals_children, arrivals_adults). In each block, the following parameters are defined: (1) the number of patients arriving per hour; and (2) the minimum, mode, and maximum values of wait times, as shown in Figure 2.

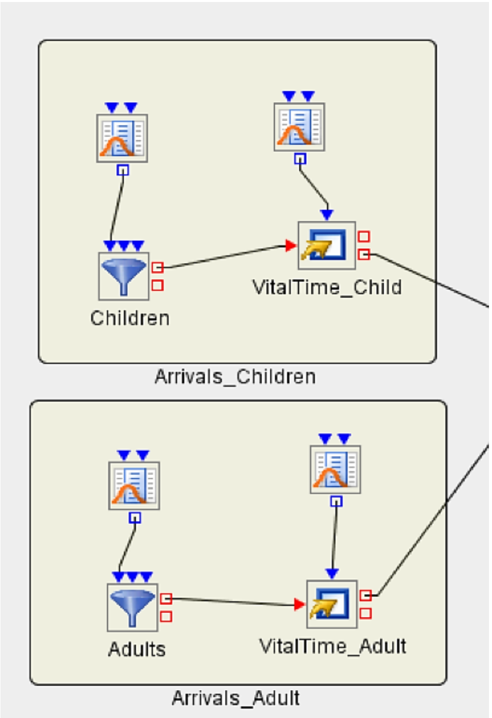


Figure 2. Simulation for an urgent care clinic (first component).

Because children and adults see the same individual for the vital sign evaluation (triage) and since children have different distribution times than adults, separate modifier blocks are added to the flow (as shown in Figure 2). The values of these triangular distributions are set as follows: (minimum = 3, mode = 5, maximum = 9) for an adult patient and (minimum = 5, mode = 8, maximum = 12) for a child patient. In addition, *VitalSigns* capacity is set to 2 due to the number of nurses in house, as shown in Figure 3.

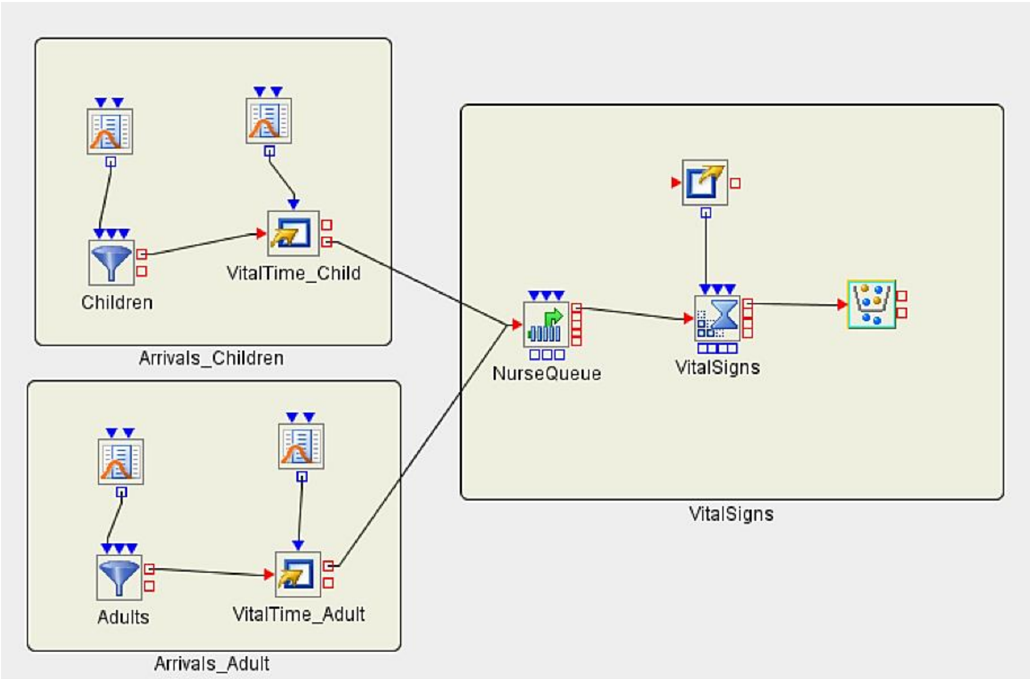


Figure 3. Simulation for an urgent care clinic (second component).

The third component of the simulation illustrates the treatment component (Figure 4). The treatment time for a child patient is triangularly distributed as (minimum = 10, mode = 14, maximum = 25) and for an adult patient the time is triangularly distributed as (minimum = 8, mode = 12, maximum = 20). Finally, *Pediatricians* capacity is set to 3 and *Medical Doctors* capacity is set to 2 to identify how many individual pediatricians and medical doctors are on staff within the clinic.

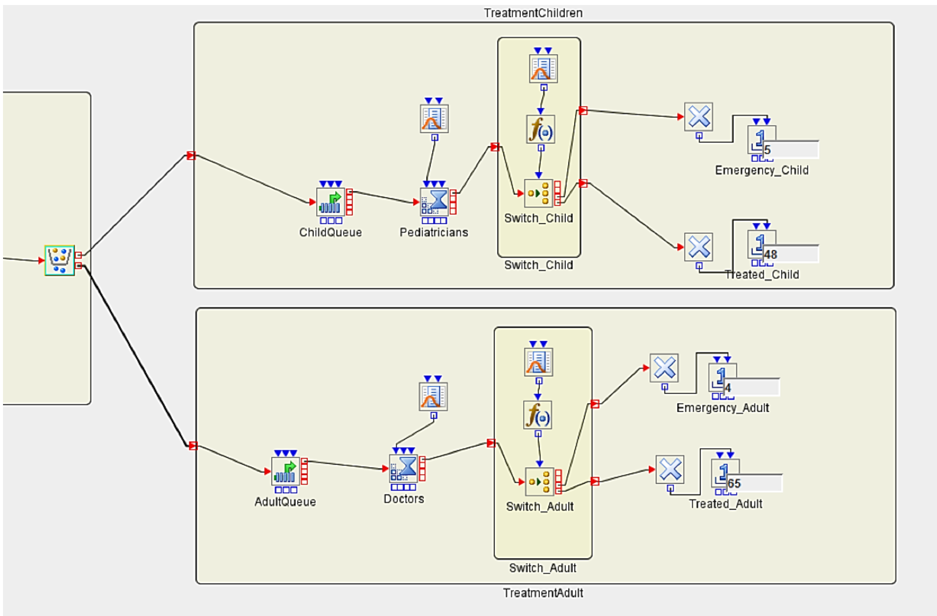


Figure 4. Simulation for an urgent care clinic (third component).

After treatment there is a probability that a patient will go to an emergency room for further treatment. The random assignment of the entity to either *Emergency Room* (value = 1) or *Treatment Complete* (value = 0) is conducted by using a uniform distribution defined as the following for *Emergency_Child*:

- *Cond* (*random value* $< 0.08, 1, 0$); *the probability for going to an Emergency Room is 8% for a child.*

Whereas, in the case of *Emergency_Adult* is defined as following:

- *Cond* (*random value* $< 0.05, 1, 0$); *the probability for going to an Emergency Room is 5% for an adult.*

The clinic's hours of daily operation are from 8:00 a.m. to 5:00 p.m., but the clinic does not actually close until all patients are treated. To accommodate this timeframe, the *End _time* considered in the experiment window is left at the default value of infinity. To shut down patient arrivals after 5 p.m., the *End _time* for both types of patients is set to 540.

After setting the following parameters including *AvgWait_Nurse*, *AvgWait_Ped*, *AvgWait_MD*, *Treated_Child*, *Treated_Adult*, *NurseQueue*, *ChildQueue*, *and AdultQueue* within SAS Simulation Studio, the simulation model is implemented (Figure 5) and the results are reported in Table 1.

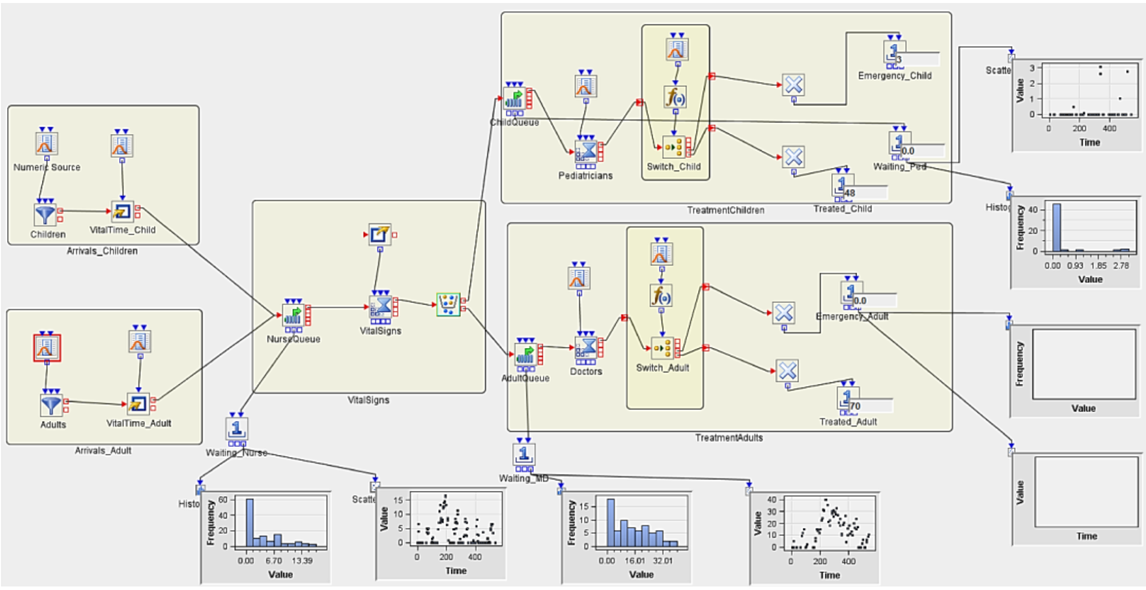


Figure 5. Simulation model when eight adult patients arrive per hour.

Table 1. Average statistics for each response within each experimental run.

Num Nurses	Num MDs	Num Peds	Replicates	AvgWait_Nurse	AvgWait_Ped	AvgWait_MD	Treated_Child	Treated_Adult
2	2	3	5	4.5940061805	0.486359693	11.41747727	51.4	66.2
			1	5.0021294076	0.956373939	4.320797381	53	59
			2	4.7728509958	0.120336954	31.04372502	56	77
			3	3.7125053454	0.259935083	7.304390717	42	57
			4	6.2493173786	0.748680849	6.824577068	58	72
			5	3.2332277751	0.346471642	7.593896164	48	66

3.5.1. Experiment 1

Let $\{X_1, X_2, \ldots, X_m\}$ be a random process representing the output from a single simulation run (first row in Table 2). Five runs of the simulation (replications) are conducted, each having a length of 5 observations, and the performance parameter of interest set as the mean across each run (shown in Table 2). A random stream seed is used to change the stream for one source of randomness.

Each of these replications is independent and identically distributed. For instance, the mean of the average replication responses is still at the top level of the design point. Case in point, the mean wait time to see a nurse (corresponding to the response *AvgWait_Nurse*) for design point-1 is 4.594 min; the mean wait time to see a pediatrician (corresponding to the response *AvgWait_Ped*) is 0.486 min; and the mean wait time to see a doctor (corresponding to the response *AvgWait_MD*) is 11.417 min.

Table 2. Average waiting time for an adult patient without any increase in staffing in minutes.

Avg_Arrival	AvgWait_Nurse	AvgWait_Ped	AvgWait_MD	Treated_Child	Treated_Adult
8/hour (μ=7.5)	4.594	0.486	**11.417**	51.4	55.2
10/hour (μ=6)	8.070	0.403	35.167	51.4	84.2
12/hour (μ=5)	14.961	0.277	71.328	51.4	101
15/hour (μ=4)	36.630	0.246	**126.101**	51.4	122.8

3.5.2. Experiment 2 (Hypothetical Case Scenario)

Let us assume that there is an outbreak of the Zika Virus in the town where this urgent care clinic is located. Based on the information given about the clinic and the simulated model generated in the above section, the new models produced are used to answer *what if* type questions regarding Zika.

Scenario 1: Suppose the arrival rate of adult patients increases with the onset of Zika. What will be, on average, the wait time for the patient (child or adult) to see a nurse and/or physician without increasing the number of physicians?

After adjusting the values of multiple replications for DES and defining along factors that enable greater control of multiple design points with this simulated experiment (Figure 6), the results of these experiments are noted in Table 2. Table 2 demonstrates the results of 3 additional experiments (rows 2, 3, 4) compared to a regular day at this clinic (row 1). More specifically, it illustrates the results of the average statistics for each response within each experimental condition. These conditions are related to the number of adult patients arriving per hour at this clinic (8, 10, 12, and 15). Comparing the first row (8 patients per hour) with the fourth row (15 patients per hour), the mean wait time to see a nurse increased from 4.594 min to 36.630 min. Even more dramatically, the average wait time for an adult patient to see a medical doctor increased from 11.417 min to 126.101 min.

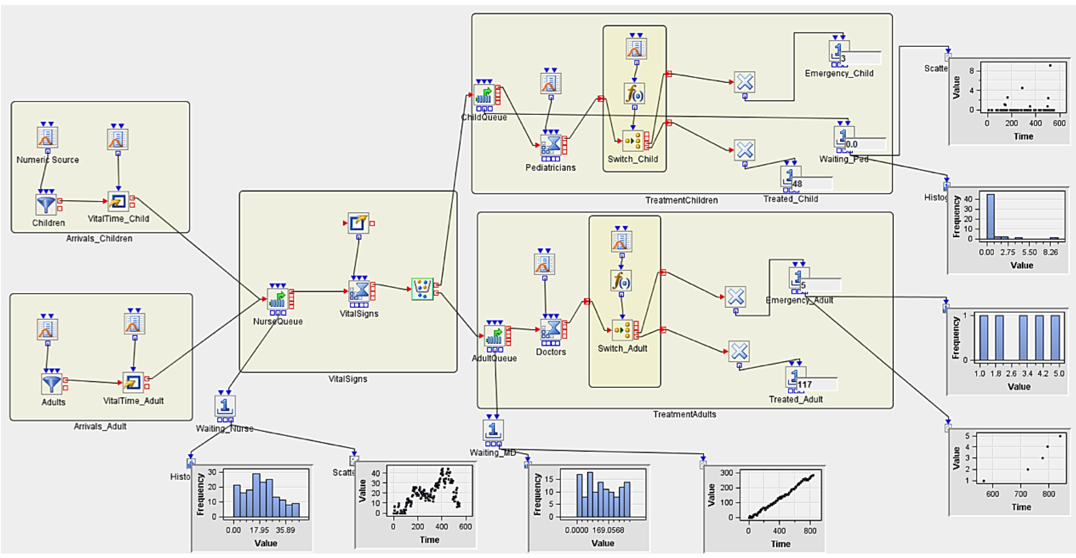


Figure 6. Simulation model when 15 adult patients arrive per hour (no additional staffing).

Scenario 2: Suppose the arrival rate of adult patients increases with the onset of Zika. What will be, on average, the wait time for the patient (child or adult) to see a nurse and/or physician when increasing the number of physicians by one?

While changing the number of doctors and pediatricians has no effect on the amount of time a patient will spend waiting to see a nurse, adding one doctor should reduce the overall adult patient wait time. The following scenario (Figure 7) examines the situation where one additional *Medical Doctor* is utilized. At *Point 2* in Table 3, 5 replications of this model are generated. These models had the same uniform distribution that was utilized in our previous experiment, except the number of medical doctors *(NumMDs)* is now 3.

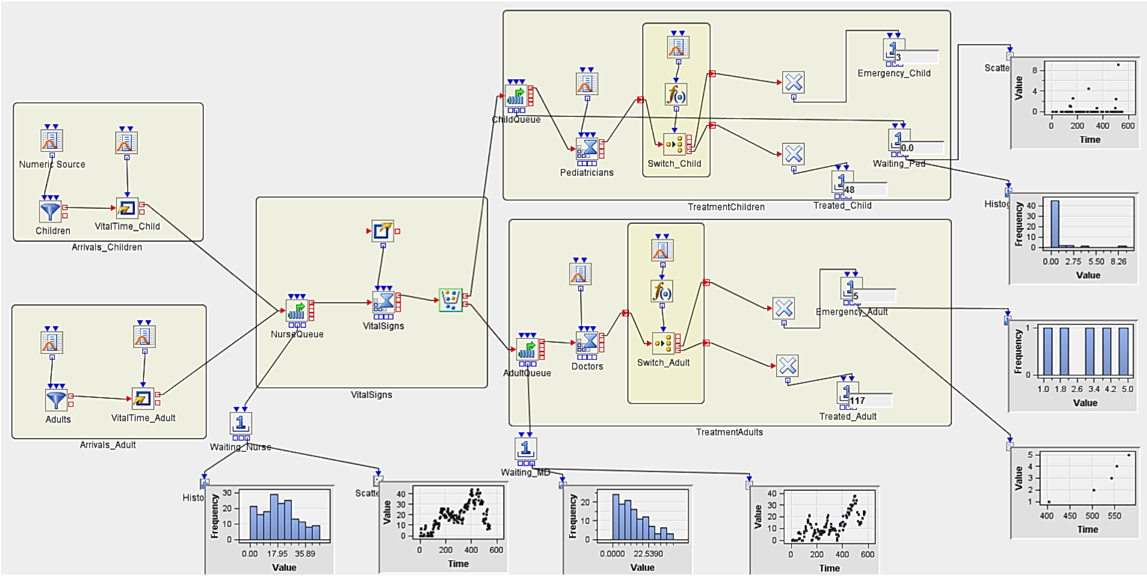


Figure 7. Simulation model when 15 adult patients arrive per hour (with increase in staffing).

Table 3. Avgerage waiting time for an adult patient with and without an additional doctor.

Num Nurses	Num MDs	Num Peds	Replicates	AvgWait_Nurse	AvgWait_Ped	AvgWait_MD	Treated_Child	Treated_Adult
2	2	3	5	36.630793431	0.246454666	126.1016082	51.4	122.8
			1	35.065351482	0.132051062	87.58132401	53	115
			2	54.166290379	0.143129240	165.5447846	56	138
			3	11.778215357	0.287797270	101.0748635	42	113
			4	62.881467827	0.240439458	141.4930868	58	131
			5	19.262642110	0.428856304	134.8139818	48	117
2	3	3	5	36.630793431	0.246454666	11.43105992	51.4	122.8
			1	35.065351482	0.132051062	5.496669125	53	115
			2	54.166290379	0.143129240	23.03052886	56	138
			3	11.778215356	0.287797270	10.15426507	42	113
			4	62.881467827	0.240439455	6.088315225	58	131

Table 4 presents the histograms of wait times for patients during one day for each experiment. The effect of the increased number of adult patients over time is evident.

Comparing Cell 1-3 and Cell 2-3, the wait times for an adult patient to see a doctor during the Zika Virus outbreak increased dramatically, with a maximum wait time of 32.5 min to more than 4 h. Also, as demonstrated by the scatter plots and a comparison of Cell 1-3 and Cell 2-3, the average wait time per adult patient increases quickly at the beginning of the day and climbs, exponentially, until the doors close. Recall that on a typical day this clinic opens at 8:00 am and closes at 5:00 p.m., which corresponds to time ranges from 0 to 540 min. This is no longer the case with this emergency because this clinic has to stay open to accommodate all patients. Cell 2-3 indicates that there are still patients in this clinic at 10:00 p.m., which corresponds to time ranges of 0 to 840 min.

Experiment 2 (Scenario 2), in which the number of medical doctors was increased from 2 to 3, shows a noticeable difference in wait times (Table 5). A comparison of Cell 2-3 and Cell 3-3 shows that utilizing an additional doctor helps, as the wait times for an adult patient to see a doctor during the Zika Virus outbreak decreased dramatically, resulting in a maximum wait time of 37.5 min compared to 4 h in Cell 2-3 (having only 2 doctors). By observing the Scatter Plots (Table 5) and comparing Cell 2-3 and Cell 3-3, it is noticeable that the average wait time per adult patient drastically decreases and the clinic doors close around 5:40 p.m. (which corresponds to time ranges from 0 to 540) compared to 10:00 p.m. in experiment 2.

Table 4. Histograms of waiting time for patients during one day for each experiment.

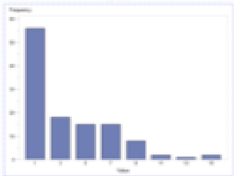
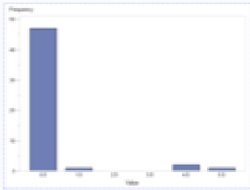
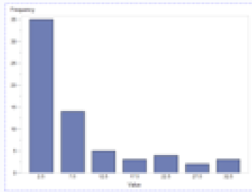
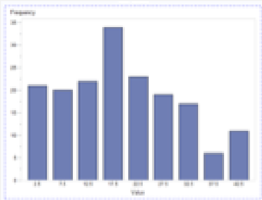
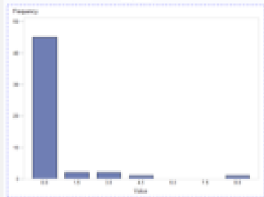
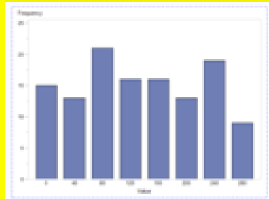
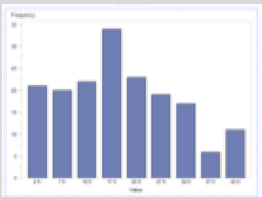
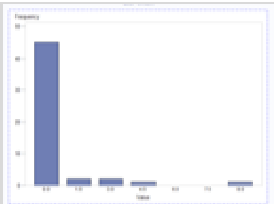
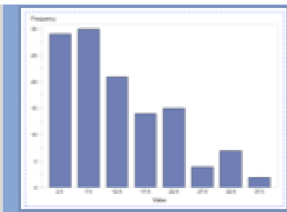
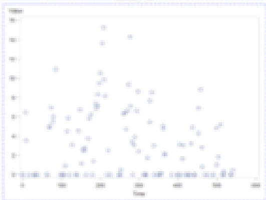
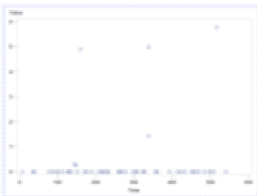
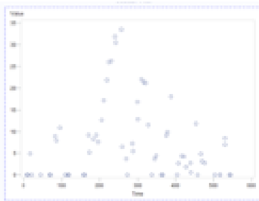
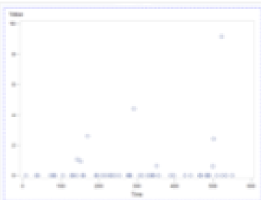
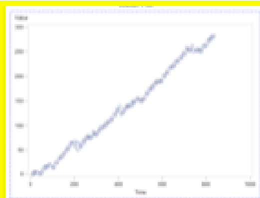
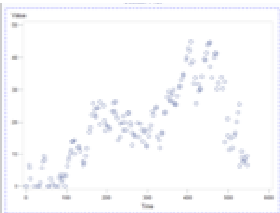
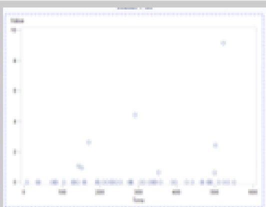
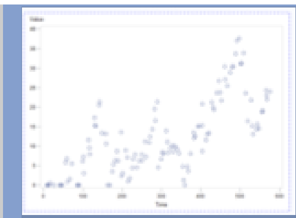
Experiment #1 No Emergency	 1-1 (*for a Nurse*)	 1-2 (*for a Ped*)	 1-3 (*for a MD*)
Experiment #2 *(Scenario 1):* Zika Virus outbreak with no additional staffing	 2-1 (*for a Nurse*)	 2-2 (*for a Ped*)	 2-3 (*for a MD*)
Experiment #2 *(Scenario 2):* Zika Virus outbreak with additional staffing	 3-1 (*for a Nurse*)	 3-2 (*for a Ped*)	 3-3 (*for a MD*)

Table 5. Scatter plots of the values of average waiting time per patient vs. time for each experiment.

Experiment #1 No Emergency	 1-1 (*for a Nurse*)	 1-2 (*for a Ped*)	 1-3 (*for a MD*)
Experiment #2 *(Scenario 1):* Zika Virus outbreak with no additional staffing	 2-1 (*for a Nurse*)	 2-2 (*for a Ped*)	 2-3 (*for a MD*)
Experiment #2 *(Scenario 2):* Zika Virus outbreak with additional staffing	 3-1 (*for a Nurse*)	 3-2 (*for a Ped*)	 3-3 (*for a MD*)

Therefore, based on the above findings, we conclude that adding one medical doctor in this case scenario was sufficient to manage the specific crisis.

With respect to validating simulated models within the field of DSE, commonly used techniques include: (1) comparing SAS simulation models to real-world systems through comparisons of output from the simulation model to actual measurements from the real system, if possible; (2) seeking the input of subject matter experts (in this instance, Zika medical experts); (3) fitting input probability distributions to ensure that these distributions match the expected distributions by running goodness-of-fit tests and plots of the data; (4) performing a sensitivity analysis to identify the factors that have a significant impact on the model outputs; and (5) animation. The Animation Method is the technique used to verify that the logic implemented in the simulated model is steady. This method was applied to generate and collect the datasets at each animated point. Lastly, these datasets were used to generate the charts presented in the previous section.

4. Conclusions and Future Work

This case study described the analysis of "what-if" situations for a local clinic. DES enabled us to examine several hypothetical circumstances without impacting the actual clinic process. Data that is used as input to the simulation model consisted of information about the number of staff this clinic has (two nurses, three pediatricians, and two doctors), in addition to its operational time (from 8:00 a.m. to 5:00 p.m.). However, since this clinic has no definite end time (only closes after treating the last patient), we had to include this characteristic within the design stage of the discrete simulated model. Multiple replications for discrete-event simulation models were applied to enable control on multiple design points within each "what-if" situation. Besides, the utilization of random seeds allowed more control over the simulation runs.

The next stage of this research is to provide a simulated App to this clinic. In this way, input data can be adjusted according to the increase or decrease in staff members, in addition to the variation of the seasoning operation time. The real-time implementation of this App will enable the construction of a chain of trials to understand better the impact of various inputs might have on the essential statistics of the responses of the simulation model. Besides, instead of using the Theoretical fitting model, a Data-Driven model can be used of patients such as (id, age, gender, and severity-level) during the arrival to a clinic in Figure 3. Accordingly, instead of exploiting the uniform distribution fitting model on the input data, we can use a discrete empirical distribution.

Finally, this App will serve as a pilot for collecting the on-going output of these simulated models based on the changes in the input data. The final step is to design a universal simulated App that can be used by any clinic in which input data can be adjusted manually.

Author Contributions: L.P.Z. performed the experiments and analyzed the results. E.T. wrote the literature review and helped in the analyzing the results.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lipsitz, L.A. Understanding health care as a complex system: The foundation for unintended consequences. *JAMA* **2012**, *308*, 243–244. [CrossRef] [PubMed]
2. Yazdanparast, R.; Hamid, M.; Azadeh, A.; Keramati, A. An intelligent algorithm for optimization of resource allocation problem by considering human error in an emergency department. *J. Ind. Syst. Eng.* **2018**, *11*, 287–309.
3. Spruit, M.; Lytris, M. Applied data science in patient-centric healthcare: Adaptive analytic systems for empowering physicians and patients. *Telemat. Inform.* **2018**, *35*, 643–653. [CrossRef]
4. Alkhaldi, F.; Alouani, A. Systemic design approach to a real-time healthcare monitoring system: Reducing unplanned hospital readmissions. *Sensors* **2018**, *18*, 2531. [CrossRef] [PubMed]

5. Islam, S.; Hasan, M.; Wang, X.; Germack, H.D.; Noor-E-Alam, M. A systematic review on healthcare analytics: Application and theoretical perspective of data mining. *Healthcare* **2018**, *6*, 54. [CrossRef] [PubMed]
6. Mahmud, R.; Koch, F.L.; Buyya, R. Cloud-fog interoperability in lot-enabled healthcare solutions. In Proceedings of the 19th International Conference on Distributed Computing and Networking, Varanasi, India, 4–7 January 2018.
7. Standfield, L.B.; Comans, T.; Scuffham, P. A simulation of dementia epidemiology and resource use in australia. *Aust. N. Z. J. Public Health* **2018**, *42*, 291–295. [CrossRef] [PubMed]
8. Croll, J. The Impact of Usability on Clinician Acceptance of a Health Information System. Ph.D. Thesis, Queensland University of Technology, Brisbane, Australia, 2009.
9. Lærum, H.; Ellingsen, G.; Faxvaag, A. Doctors' use of electronic medical records systems in hospitals: Cross sectional survey. *BMJ* **2001**, *323*, 1344–1348.
10. Green, L.A.; Fryer, G.E.; Yawn, B.P.; Lanier, D.; Dovey, S.M. Ecology of medical care revisited. *N. Engl. J. Med.* **2001**, *344*, 2021–2025. [CrossRef] [PubMed]
11. Mead, N.; Bower, P. Patient-centredness: A conceptual framework and review of the empirical literature. *Soc. Sci. Med.* **2000**, *51*, 1087–1110. [CrossRef]
12. Fayyad, U.; Piatetsky-Shapiro, G.; Smyth, P. From data mining to knowledge discovery in databases. *AI Mag.* **1996**, *17*, 37–54.
13. Hernández, M.A.; Stolfo, S.J. Real-world data is dirty: Data cleansing and the merge/purge problem. *Data Min. Knowl. Discov.* **1998**, *2*, 9–37. [CrossRef]
14. Negash, S. Business intelligence. In *Handbook on Decision Support Systems 2*; Springer: Berlin/Heidelberg, Germany, 2004; Volume 13, pp. 177–195.
15. Bonney, W. Applicability of business intelligence in electronic health record. *Procedia Soc. Behav. Sci.* **2013**, *73*, 257–262. [CrossRef]
16. Spruit, M.; Vroon, R.; Batenburg, R. Towards healthcare business intelligence in long-term care: An explorative case study in the Netherlands. *Comput. Hum. Behav.* **2014**, *30*, 698–707. [CrossRef]
17. Pérez, E.; Uyan, B.; Rohde, R.E.; Wehbe-Janek, H.; Hochhalter, A.K.; Fenton, S.H. Assessing catheter associated urinary tract infections prevention interventions in intensive care units: A discrete event simulation study. *IISE Trans. Healthc. Syst. Eng.* **2016**, *7*, 43–52. [CrossRef]
18. Allison, P.D. Discrete-time methods for the analysis of event histories. *Sociol. Methodol.* **1982**, *13*, 61–98. [CrossRef]
19. Fishman, G.S. *Principles of Discrete Event Simulation*, 1st ed.; Wiley: Hoboken, NJ, USA, 1978.

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).