PRIMARY RESEARCH   **OPEN ACCESS**

Modeling Patient Flow among Hospital Wards Using Non-Diagnostic Data

Check for updates

Marco Lussetti, BCS [1], Piper Jackson, PhD [2]*

[1] Department of Computing Science, Thompson Rivers University, Kamloops, BC, Canada, V2E2A5
[2] Department of Computing Science, Thompson Rivers University, Kamloops, BC, Canada, V2E2A5

*Corresponding Author: pjackson@tru.ca

Abstract
Hospital bed capacity is a limited resource and a key concern in health care planning. Using discrete-event simulation modeling and the MIMIC-III data set, this paper produces a demographic and metadata-only model of patient transfer within hospital wards. The model successfully approximated the underlying transfer dynamics (95.63% accuracy measured using RMSE). The accompanying visualization may be used to examine patient flow. The simulation will be used as a test bed for future work concerning flow of artificially generated patient admissions and can in general be useful in simulating patient flow in cases where demographic information is available but transfer records are not.

Keywords: modeling and simulation; health informatics; computational social science; data science

Introduction

In health care planning, the hospital occupies a central role. Hospital resources are limited, and particularly so in Intensive Care Units (ICU)-type wards due to the high cost associated with ICUs [1,2]. Traditionally, a hospital's capacity was estimated in terms of the number of beds available in the hospital, but increasingly additional measures (e.g., patient flow) are considered useful in modern hospital capacity planning [3]. Modeling in the field of health care has many applications, including capacity planning, epidemiology, and understanding doctor-patient relationships [4]; however, planning and resource utilisation has been highlighted as the most widely modeled area of health care [5]. Our work builds upon a well-established tradition of applying simulation modeling to the problem of improving the use of hospital resources [6]. This is often referred to as "bed modeling" due to the capacity limitation of a hospital being defined (at least in principle) by the number of beds available. Various methods have been used to model the movement of patients through a hospital, including discrete-event simulations, system dynamics, and Monte Carlo methods among others, as described in [7]. Previous research has also specifically investigated the use of discrete-event simulations to estimate ICU bed capacity to explore occupancy rates [2].

Discrete-events simulation (DES) is a popular modeling technique in which individual and independent entities are subject to events which may be as varied as the subject modeled [8,9]. In the context of the choice of modeling technique for this kind of problem, DES may be distinguished from system dynamics by virtue of its non-deterministic and non-continuous nature, allowing events rather than explicit equations to represent the flow through the system [5]. Another commonly used simulation technique, agent-based modeling, allows for incorporating "human behavioral aspects" by allowing for rules of varying complexity that the entities or agents in the system will follow. DES allows us to build a model of relative simplicity that is not focused on individual entities but rather on their flow through the system [10].

This paper aims to investigate the use of discrete-events simulations to produce an abstract overview of patients' movements between hospital units. In particular, the aim is to construct a flexible model of the flow of patients among hospital units via the use of demographic information and patient arrival metadata without any regard to diagnostic information concerning a patient's condition(s). This model may then be used to simulate desired demographic or arrival condition mixes and observe their impact on the system's occupancy. Because a patient's movement within a hospital is largely assumed to be a function of the patient's condition and treatment received which is not directly modeled, it is hoped that the model will yield some insight as to how much of the real-world flow of patients could be inferred without having to consider the primary causes. The model's nature as centered in demographic and arrival methods/times has the potential to be informative for hospital layout planning and in general planning for cases where the detailed diagnostic and transfer information might not be available to planners.

This paper is organized as follows: the next section discusses related works; followed by a description of the methods including the source material used and the process employed; then in-depth analysis of the results of the simulation; a discussion of notable aspects of the model and results; and finally conclusions drawn from this stage of our research.

Related Works

The use of modeling techniques and simulation to examine hospital capacity is well established in literature, with Markovian models being reviewed among other techniques by McClain [11]. Bed modeling research has since incorporated demographic characteristics and other ancillary information to simulate interactions in the context of hospital systems with Dumas' work [12]. Different modeling methods have been used explored for bed modeling including neural networks [13], queue-based modeling [11,14,15], probability-based mathematical models [16,17], system simulations [18], and discrete event simulation [19-23]. Specific research has been conducted looking at patient flow in the context of elective surgeries [24], admission policies [25], neuroscience departments [26], cardiovascular surgeries [23], bloodwork and other medical scans administered [22], and among wards [20].

Methods

To build the model, the MIMIC-III dataset was used as source material. The MIMIC III dataset originates with the Beth Israel Deaconess Medical Center in Boston, Massachusetts and spans critical care unit stays between 2001 and 2012. The MIMIC III dataset is "is the only freely accessible critical care database of its kind", and provides detailed information about patient demographics and movement between wards of the hospital [27]. The dataset has been previously used for similar analyzing stay duration length, including predicting stay length duration using neural networks [28].

Patient entities are generated using the demographic data from MIMIC-III. Each entity has the following characteristics:

- **Admission location:** which unit the patient enters the hospital from, e.g., Emergency Room Admission, Clinical Referral, Physician Referral, etc.
- **Gender:** male or female.
- **Age range:** Converted from exact ages to one of four age ranges (under 15, 15-24, 25-64, 65 and over).
- **Admission hour:** 0-23, i.e., from 12:00 a.m. (midnight) to 11:00 p.m.

Patients randomly transfer from one unit to another specific unit based on weighted probabilities reflecting the underlying MIMIC-III data. This is triggered after passing the median timeslot in terms of waiting time in the same underlying dataset. Median time spent in each unit was

calculated based on the MIMIC-III data. Using these characteristics of patients and transfers within the hospital, we simulate patient admissions over a single day and their time within hospital up to a predetermined end time (two weeks in our experiments).

The model's basic time unit is a single hour. Simulating patient movement followed this process for each timeslot:

1. New patients are added to the system based on admission hour (no patients added past the first day).
2. For each location and patient, check if the patient needs to move. If so, use the transfer probabilities to stochastically select a destination and update the patient's location accordingly.
3. If not at the end time, continue to the next time slot.

Thus, to create the model, two key metrics were used: probability of transfer between states (where states are the wards patients are treated in), and the amount of time spent in each state. The former is a relatively straightforward exercise, as the dataset is broad enough to be able to simply rely on transition probabilities from the underlying dataset without alterations. The latter was created from the medial time spent in between the given transitions based on demographic data and other metadata considered.

The timestamps present in the source database are internally consistent between each entry, however there is some small portion of the data with inconsistent entries. This however represents only a very small portion of the overall dataset and should have no significant effect on the overall results. Once the basic model is built as matching probabilities and durations, input data can be run through it to simulate the patients' arrival.

Results

The results of running any given input data are exported as a series of timeslots where the placement of patients is tracked by ward and unit combination. This can then be visualized by replaying the succession of timeslots in a basic visualization module. It can also be compared to real non-generalized hospital data for validation purposes. While the model represents an obvious generalization of the underlying data, to ascertain just how much generalization was occurring it was run through with the MIMIC III dataset and compared to the real transfer data on a timeslot by timeslot basis. A simulation was run for patient permanence in the hospital over 14 days (patients n=58952), clustering all arrivals in the first 24 hours but maintaining the hour of arrival. The real transfer data's granularity was reduced to match the one-hour timeslot duration of the model.

The mean square error and root mean square error were calculated for each timeslot. The root mean square error (Figure 1) is of particular interest and represents the number of misplaced individuals in the system at each timeslot, computed as the difference between the expected and found number of individuals present in each ward. It is expected
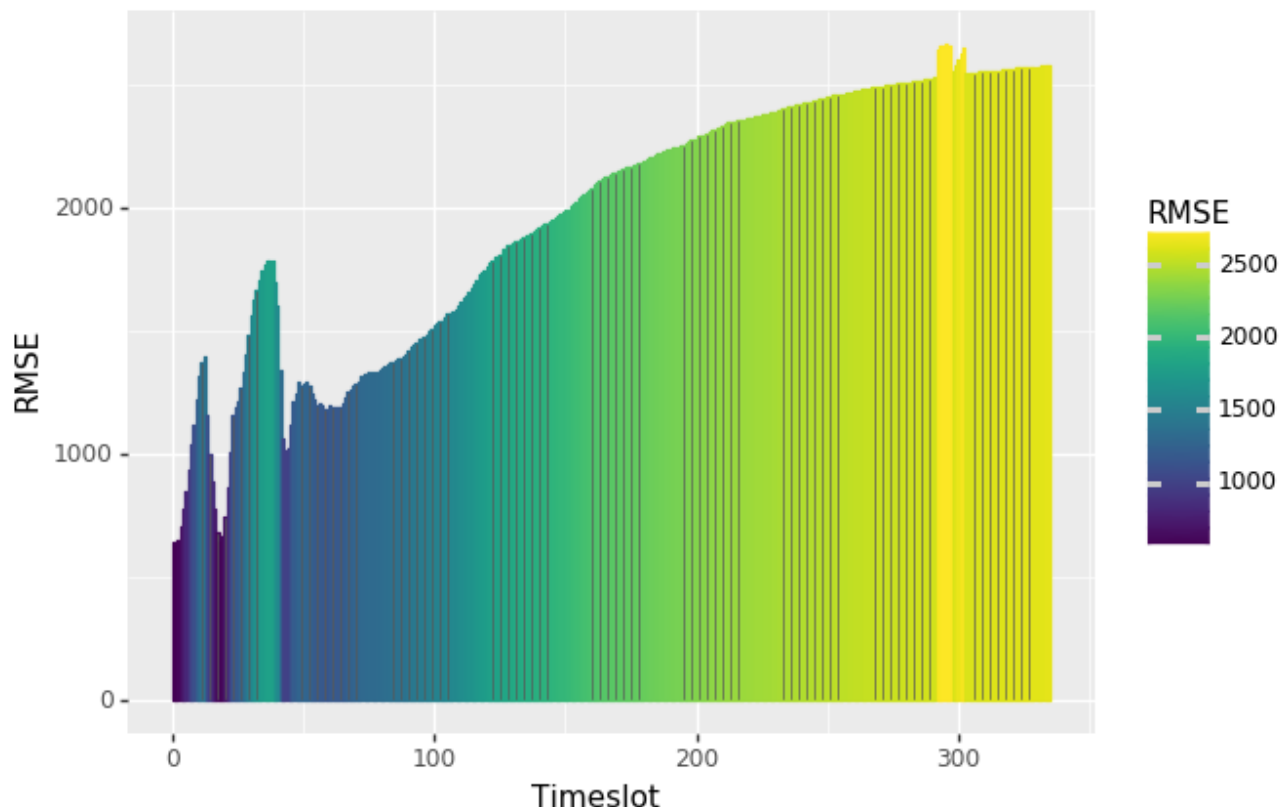
Figure 1. Root Mean Square Error by Timeslot

that as the simulation proceeds the number of misplaced individuals will increase. The results of the simulation can also be visualized on a per timeslot basis, showing the number of individuals in each of the wards, admission locations, and discharge locations modeled. We include sample figures from one execution of the model showing the model's state after one day (Figure 2), two days (Figure 3), and at the end of the run – two weeks (Figure 4).

Discussion

At the end of the sample two-week run, RMSE = 2579, which represents an accuracy of 0.9563 over the sample size. This high accuracy is expected as the model is tightly coupled with the underlying MIMIC-III data used to generate the model. The accuracy value can be thought of as a proxy for the model's expressiveness in explaining patient movements using demographic information and admission metadata.

The visualization allows us to see the traffic levels of the different hospital units over time, with large clusters on admission units (the first row) in early time steps, dispersion into many specialist units (central rows) in intermediate time steps, and various discharge locations (bottom row) at the end of simulation. From this, we are able to clearly see which units get the most traffic, which is critical for debugging (during development and extension

of the model) as well as analysis. For instance, by running intentionally skewed demographic makeups through the model, one may observe the effect this causes on specific wards and at specific time in the entities' permanence in the system.

During development, a tightly coupled approach to the underlaying data was adopted, taking advantage of the relatively high number of cases available. As the work is extended in future to encompass synthetic and different data (artificially generated patients, patients' records from different sources) the routing model will likely need to be fuzzified to allow for routing paths not explicitly present in the source dataset used to generate the model.

In observing the model's visualization, it was noted that the admission and discharge locations appear to possess a Pareto-like distribution, showing that most patient were either sent home or cared for in their home (Home, Home Health Care) and that likewise most patients hospitalized were admitted via the Emergency Room or Clinical Referral. This is a characteristic of the underlaying data which the model replicates, and it would be interesting to see whether acquiring and normalizing data from other hospitals' admission databases would yield similar distributions. In contrast, patients are most evenly distributed at the intermediate stages inside the hospital, however the dataset's documentation does not provide
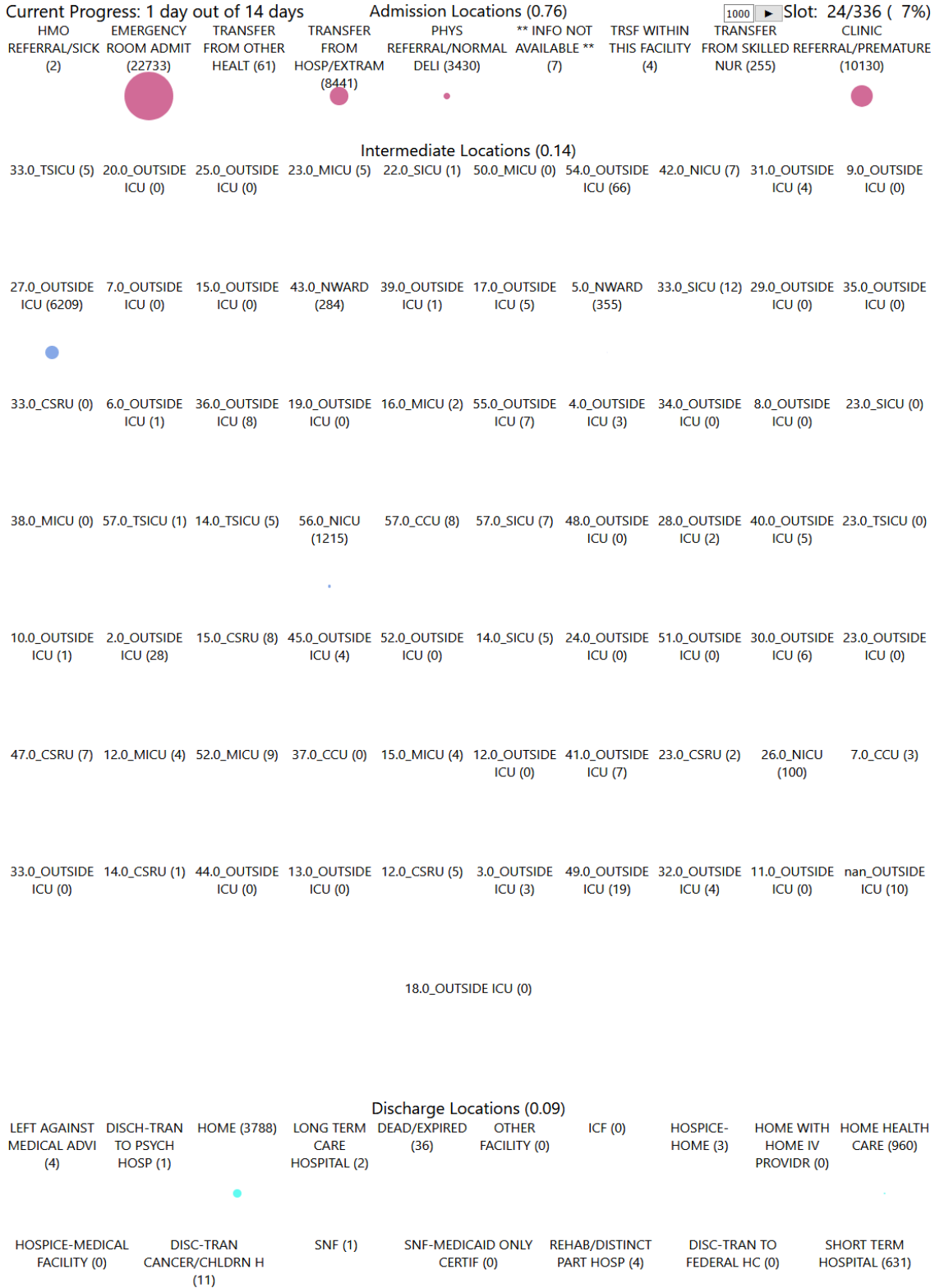
Figure 2. Model State at Timeslot 24 (1 Day)
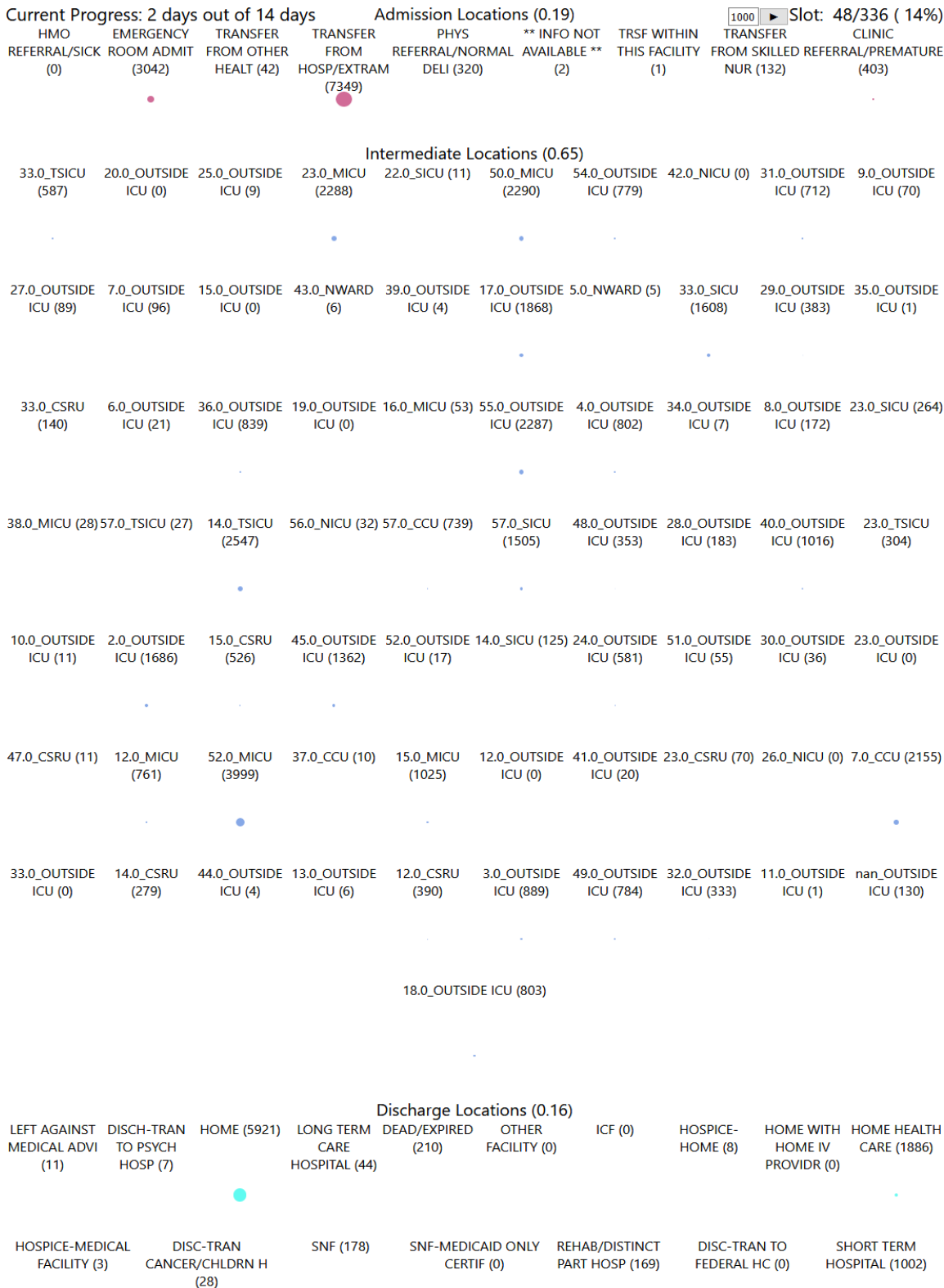
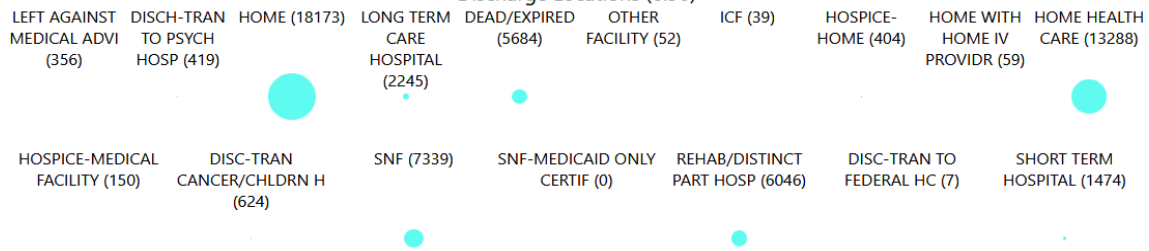**Figure 3.** Model State at Timeslot 48 (2 Days)

Figure 4. Model State at Timeslot 336 (2 Weeks)

insight into the purpose and nature of non-ICU wards and notes that the numbering may not be consistent for the duration of the data collection period. This limits the ability to draw conclusions from non-ICU portions of the model. This includes the high early allocation of patients to Ward 27. The model is limited by not investigating waiting times directly, but instead relies exclusively on arrival and departure times for each ward or unit. That means that no queuing system is explicitly implemented in the model, though some queuing is in effect implied in the median time spent in a unit. Future work could refine this to include waiting time and explicit capacity limits on wards to improve the model's explanatory power.

Conclusions

This model is a preliminary attempt to consider demographic aspects of hospital resource utilization with high accuracy in matching patient transfers to the source data. This is notable considering that only four patient characteristics are used to compute transfer, two of which are demographic (age and gender) and two which are based on admission (admission location and time of day). While it still makes sense to consider medical characteristics (such as diagnosis) in a more complex iteration of this model, the current results are a strong argument for the importance of including demographic factors. The MIMIC-III data also includes other patient data, including religion, ethnicity, and marital status. These could also be informative characteristics to add to the model, although some specific considerations need to be made, such as whether or not to consider different religious denominations (e.g., Catholic, Baptist, Lutheran, etc.) separately, or to combine them into one group for each religion as a whole. The next stage of investigation is to test this system both using artificially generated patients and patients from other datasets, and test whether the system behavior as a whole is still close to the MIMIC-III data. One key requirement we are working on is ensuring that the generated patients are of sufficiently high quality, i.e., they are similar to real patients admitted to this hospital. We hope that the system model described here will serve as a good testbed for the next stage of our work.

List of Abbreviations Used

ICU: Intensive Care Unit
MIMIC: Medical Information Mart for Intensive Care
MSE: Mean Squared Error
RMSE: Root Mean Squared Error

Conflicts of Interest

The authors declare that they have no conflict of interests.

Ethics Approval and/or Participant Consent

Research that only uses secondary data does not require ethical review at Thompson Rivers University. Use of the MIMIC-III data set followed the ethical requirements and process set out by the data holder.

Authors' Contributions

ML: made ongoing contributions to the design of the study, accessed and prepared the data, built the simulation model and visualization, ran the analysis, contributed to the manuscript, and gave final approval of the version to be published.

PJ: made ongoing contributions to the design of the study, contributed to and edited the manuscript, and gave final approval of the version to be published.

References

[1] Van Houdenhoven M, Nguyen D-T, Eijkemans MJ, Steyerberg EW, Tilanus HW, Gommers D, et al. Optimizing intensive care capacity using individual length-of-stay prediction models. Crit Care. 2007;11(2):R42. https://doi.org/10.1186/cc5730

[2] Hoon Hen B, Zhu Z, Liang Teow K. Estimating ICU bed capacity using discrete event simulation. International J Health Care QA. 2012 Feb 3;25(2):134–44. https://doi.org/10.1108/09526861211198290

[3] Rechel B, Wright S, Barlow J, McKee M. Hospital capacity planning: from measuring stocks to modelling flows. Bulletin of the World Health Organization. 2010 Aug;88(8):561–640. https://doi.org/10.2471/BLT.09.073361

[4] Complex Systems Modelling Group. Modelling in Healthcare. Providence, R.I: Amer Mathematical Society; 2010. 218 p.

[5] Brailsford SC, Hilton NA. A comparison of discrete event simulation and system dynamics for modelling health care systems. In: Riley J, editor. Planning for the Future: Health Service Quality and Emergency Accessibility. Glasgow Caledonian University; 2001. Available from: https://eprints.soton.ac.uk/35689/

[6] Günal MM, Pidd M. Discrete event simulation for performance modelling in health care: a review of the literature. J Simulation. 2010 Mar 1;4(1):42–51. https://doi.org/10.1057/jos.2009.25

[7] Mielczarek B, Uziałko-Mydlikowska J. Application of computer simulation modeling in the health care sector: a survey. SIMULATION. 2012 Feb 1;88(2):197–216. https://doi.org/10.1177/0037549710387802

[8] Robinson S. Discrete-event simulation: from the pioneers to the present, what next? Journal of the Operational Research Society. 2005 Jun 1;56(6):619–29. https://doi.org/10.1057/palgrave.jors.2601864

[9] Allen M, Spencer A, Gibson A, Matthews J, Allwood A, Prosser S, et al. What is discrete event simulation, and why use it? NIHR Journals Library; 2015. Available from: https://www.ncbi.nlm.nih.gov/books/NBK293948/

[10] Siebers PO, Macal CM, Garnett J, Buxton D, Pidd M. Discrete-event simulation is dead, long live agent-based simulation! Journal of Simulation. 2010 Sep 1;4(3):204–10. https://doi.org/10.1057/jos.2010.14

[11] McClain JO. Bed Planning Using Queuing Theory Models of Hospital Occupancy: A Sensitivity Analysis. Inquiry. 1976;13(2):167–76. http://www.jstor.org/stable/29770996

[12] Dumas MB. Simulation modeling for hospital bed planning. SIMULATION. 1984 Aug 1;43(2):69–78. https://doi.org/10.1177/003754978404300201

[13] Walczak S, Pofahl WE, Scorpio RJ. A decision support tool for allocating hospital bed resources and determining required acuity of care. Decision Support Systems. 2003 Mar 1;34(4):445–56. http://www.sciencedirect.com/science/article/pii/S0167923602000714

[14] Barado J, Guergué JM, Esparza L, Azcárate C, Mallor F, Ochoa S. A mathematical model for simulating daily bed occupancy in an intensive care unit*. Critical Care Medicine. 2012 Apr;40(4):1098. https://journals.lww.com/ccmjournal/Abstract/2012/04000/A_mathematical_model_for_simulating_daily_bed.7.aspx

[15] Cochran JK, Roche K. A queuing-based decision support methodology to estimate hospital inpatient bed demand. Journal of the Operational Research Society. 2008 Nov 1;59(11):1471–82. https://orsociety.tandfonline.com/doi/abs/10.1057/palgrave.jors.2602499

[16] Kokangul A. A combination of deterministic and stochastic approaches to optimize bed capacity in a hospital unit. Computer Methods and Programs in Biomedicine. 2008 Apr 1;90(1):56–65. http://www.sciencedirect.com/science/article/pii/S0169260708000047

[17] Lamiell JM. Modeling Intensive Care Unit Census. Mil Med. 1995 May 1;160(5):227–32. https://academic.oup.com/milmed/article/160/5/227/4843578

[18] Kumar S. Modeling Hospital Surgical Delivery Process Design Using System Simulation: Optimizing Patient Flow and Bed Capacity As an Illustration. Technol Health Care. 2011 Jan;19(1):1–20. http://dl.acm.org/citation.cfm?id=1938499.1938502

[19] Griffiths JD, Jones M, Read MS, Williams JE. A simulation model of bed-occupancy in a critical care unit. Journal of Simulation. 2010 Mar 1;4(1):52–9. https://doi.org/10.1057/jos.2009.22

[20] Bountourelis T, Luangkesorn L, Schaefer A, Maillart L, Nabors SG, Clermont G. Development and validation of a large scale ICU simulation model with blocking. In: Proceedings of the 2011 Winter Simulation Conference (WSC). 2011. p. 1143–53.

[21] Zhu Z, Hoon Hen B, Liang Teow K. Estimating ICU bed capacity using discrete event simulation. International J Health Care QA. 2012 Feb 3;25(2):134–44. http://www.emeraldinsight.com/doi/10.1108/09526861211198290

[22] Brenner S, Zeng Z, Liu Y, Wang J, Li J, Howard PK. Modeling and Analysis of the Emergency Department at University of Kentucky Chandler Hospital Using Simulations. Journal of Emergency Nursing. 2010 Jul 1;36(4):303–10. http://www.sciencedirect.com/science/article/pii/S0099176709003304

[23] Marmor YN, Rohleder TR, Cook DJ, Huschka TR, Thompson JE. Recovery bed planning in cardiovascular surgery: a simulation case study. Health Care Manag Sci. 2013 Dec 1;16(4):314–27. https://doi.org/10.1007/s10729-013-9231-5

[24] Kolker A. Process Modeling of ICU Patient Flow: Effect of Daily Load Leveling of Elective Surgeries on ICU Diversion. J Med Syst. 2008 May 17;33(1):27. https://doi.org/10.1007/s10916-008-9161-9

[25] De Oliveira MJF. A patient-oriented Modelling of the Emergency Admission system of a Brazilian Hospital. In Glasgow, Scotland; 1993. https://www.researchgate.net/profile/Mario_Jorge_De_Oliveira2/publication/304381150_A_patient-oriented_Modelling_of_the_Emergency_Admission_system_of_a_Brazilian_Hospital/links/576d892a08ae0b3a3b7553b1.pdf

[26] Hiltrop J. Modeling neuroscience patient flow and inpatient bed management [Thesis]. Massachusetts Institute of Technology; 2014. https://dspace.mit.edu/handle/1721.1/90769

[27] Johnson AEW, Pollard TJ, Shen L, Lehman LH, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. Scientific Data. 2016 May 24;3:160035. https://doi.org/10.1038/sdata.2016.35

[28] Gentimis T, Alnaser AJ, Durante A, Cook K, Steele R. Predicting Hospital Length of Stay Using Neural Networks on MIMIC III Data. In: 2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech). 2017. p. 1194–201. https://doi.org/10.1109/DASC-PICom-DataCom-CyberSciTec.2017.191

Citation
Please cite this article as follows:
Lussetti M, Jackson P. Modeling patient flow among hospital wards using non-diagnostic data. URNCST Journal. 2019 Sep 30: 3(8). https://urncst.com/index.php/urncst/article/view/153
DOI Link: https://doi.org/10.26685/urncst.153

Do you research in earnest? Submit your next undergraduate research article to the URNCST Journal!
| Open Access | Peer-Reviewed | Rapid Turnaround Time | International |
| Broad and Multidisciplinary | Indexed | Innovative | Social Media Promoted |
Pre-submission inquiries? Send us an email at info@urncst.com | Facebook, Twitter and LinkedIn: @URNCST
Submit YOUR manuscript today at https://www.urncst.com!