

Size does matter: a simulation study of hospital size and operational efficiency

T. Bogomolov^a, J.A. Filar^b, R. Luscombe^c, Y. Nazarathy^b, S Qin^a, P. Swierkowski^c, I. Wood^b

^a Flinders University, ^b University of Queensland, ^c Sunshine Coast Hospital and Health Service
Email: shaowen.qin@flinders.edu.au

Abstract: Hospitals come in different sizes and operate under different conditions. The problem confronting healthcare planners for a long time has been: *How does hospital efficiency relate to its size?* Due to their complex organisational structures, it would be unrealistic, costly and risky for hospitals to conduct field experiments to test the effects of any major organisational change. Mathematical modelling and computer simulations offer an effective and risk-free approach to assess likely impacts of any proposed change.

This study replicates and extends the “Hospital Event Simulation Model: Arrivals to Discharge (HESMAD)” (Ben-Tovim et al., 2016), originally built for the Flinders Medical Centre (SA), to use simulation to model patient flow in another hospital – Nambour on the Sunshine Coast, Queensland. The study will take advantage of the unique opportunity – a natural experiment occurring in the Sunshine Coast, where Nambour hospital has been downsized by more than 50% – for subsequent validation of the HESMAD model, and simulation modelling in general, with data collected before and after the downsizing.

The simulations investigate hospital efficiency by examining the following organisational metrics: (1) length of stay (LOS) by diagnosis-related group (DRG); (2) total hospital utilisation; (3) emergency department (ED) utilisation, where utilisation is synonymous with occupancy rate, that is, number of occupied beds divided by total number of available beds.

The model is used to run experiments on different hospital sizes, especially that of Nambour Hospital before its restructure – 100%, then 50% and 200% of the original size. The base scenario is set to be a reduction/increase in arrivals in line with hospital size while preserving the proportion of ambulance and walk-in (self-presented) patients and the pre-downsizing mix of DRGs. Alternative scenarios tested also include changing the mix of ambulance/walk-in arrivals, as well as DRGs, while maintaining the total rate of arrivals. Such scenarios are realistic because, while ambulance arrivals can be controlled and diverted to other hospitals, self-presentations are much more difficult to control.

Simulation results demonstrate that for the base scenario (proportional reduction) there is no change in average figures of hospital utilisation and LOS (both overall and for individual DRGs). However, there is a dramatic change in variability of the utilisation. A smaller-size hospital has a much greater dispersion in distributions and, consequently, a much higher propensity for becoming overcrowded.

Alternative scenarios (corresponding to disproportional reductions) can be considered as new and very different hospitals. Changing the proportion of ambulance and walk-in patients affects the DRG mix. As a result, the overall LOS, hospital occupancy and waiting times, all change considerably.

The main conclusions are: (1) Smaller size hospitals are operationally more risky; they are at a higher risk of overcrowding, while large hospitals have better ability to absorb spikes in arrivals. This may mean that provision of additional ‘surge’ capacity may be required when commissioning smaller facilities. (2) While deciding on the size of a hospital, attention should be given to the clinical function that the facility is to perform (i.e. the potential mix of DRGs) rather than just the size of a population that the hospital is serving.

Keywords: *Simulation, hospital size, length of stay (LOS), diagnosis-related group (DRG), efficiency, patient flow*

1. INTRODUCTION AND BACKGROUND

Hospitals provide essential services to the community. Yet, the ever-growing cost of hospital care continues to be of great concern in light of demographic trends, such as ageing population and population growth. Therefore, hospital efficiency remains a key goal of health service planning and resource allocation. While hospitals come in different sizes and operate under different conditions, a fundamental problem confronting funders and health service planners for a long time has been: *How does hospital efficiency relate to its size*

Hospital efficiency might have multiple definitions and cover different aspects. In this research, the following organisational metrics are considered: (1) length of stay (LOS) by diagnosis-related groups (DRGs); (2) total hospital utilisation (i.e. rates of occupancy); (3) emergency department (ED) utilisation.

Patient LOS for different DRG, already used for hospital utilisation management and budgeting, could and should be used for regional planning (Fetter *et al.*, 1980). The concept of hospital and ED utilisation is characterised by two reciprocal constraints. On the one hand, it may be unacceptable for utilisation to approach 100%, as it would mean that a hospital is not able to treat new patients arriving. On the other hand, very low utilisation means that valuable resources (beds, staff, medical equipment) are wasted. The question of “optimal utilisation” is actively debated in the literature but a conclusive answer has been elusive (Bain *et al.*, 2010). The problem of cost efficiency of the different size hospitals also has a long history of research (Carr & Feldstein, 1967; Choi *et al.*, 2017) but it lays beyond the scope of this study.

Past studies have linked hospital occupancy, overcrowding in ED and LOS. For example, Richardson (2006) showed how ED overcrowding experience affects overall LOS and even increases mortality rate; Forster *et al.* (2003) demonstrated that ED overcrowding depends on hospital beds availability, that is hospital occupancy; Harris and Sharma (2010) studied LOS in ED and concluded that increased hospital capacity could reduce overcrowding.

Most of these studies have been carried out using real data. However, data availability, complexity and potentially confounding factors make the results observed with such data of limited generalizability and applicability to other situations. Due to the complex organisational structures of modern hospitals and the large number of confounding influences, it would be challenging to conduct experiments to accurately test the effects of any actual facility re-sizing. Mathematical modelling and computer simulations offer an effective and risk-free approach to assess likely impacts of any proposed change, while keeping potentially confounding variables constant (e.g. Hoot *et al.*, 2008; Brailsford *et al.*, 2009).

Computer simulation is an excellent tool for studying systems with complex structures, such as large hospitals where interactions are not easily modelled using closed-form analytical approaches. The aim is to develop a comprehensive model, which can be used to investigate the dynamics of various scenarios in a hospital. While many studies model only some departments, predominantly EDs, a small number of recent studies have modelled more complete hospital structures (Holm *et al.*, 2013; Ben-Tovim *et al.*, 2016). From these studies, it is evident that hospital size is a key variable that could influence hospital efficiency. Yet, to the best of the authors’ knowledge no studies have directly modelled a relationship between hospital size and its efficiency measured as a LOS for different DRGs. This is the focus of this article.

This study employs an adaptation of the Hospital Event Simulation Model: Arrivals to Discharge (HESMAD) (Ben-Tovim *et al.*, 2016) to model patient flow in Nambour Hospital, Queensland. It is inspired by the unique opportunity – a natural experiment taking place in Sunshine Coast, where the capacity of Nambour Hospital is downsized by more than 50%. The model is used to assess efficiency of Nambour Hospital before its restructure (100% size) and then 50% and 200% of the original size. The “base” scenario with a 50% reduction in arrivals is characterised by preserving the proportion of ambulance and walk-in (self-presented) patients and the mix of DRGs. Alternative scenarios that are subsequently investigated include changing the mix of ambulance/walk-in arrivals (which automatically changes the DRG mix), while maintaining the total rate of arrivals at 50% of the original one.

2. METHOD

2.1. Model

The HESMAD model (Ben-Tovim *et al.*, 2016) has been tailored for Nambour Hospital. Figure 1 shows the model structure. Due to space constraint, only an indicative view of the model logic is provided. The model captures events, queues and treatment durations with their associated control rules. Connections between model elements create possible routes for a patient’s hospital journey - from arrival to discharge. For example, all patients arrive at “triage” where each of them receives a triage score that sets ED service priority. Patients

queue up for the service in accordance to triage score and then a first-come-first-serve regime whenever triage scores are equal.

Patient journey might be affected by different factors, including (1) patients' personal characteristics (gender, age, health status, type of arrival); (2) hospital conditions (overall hospital occupancy level, emergency department overcrowding, availability of beds in specialised units); and (3) some external parameters (day of the week and time of the day).

It is worth noting that the level of detail that can be included in the model is largely dictated by that of the hospital data record. Parameters of the events or procedures not recorded cannot be calibrated and validated with any real data, and, as a result, cannot be included in the model. For example, the physical lay-out of wards and corridors might have an effect on the total length of stay in the hospital. However, this potentially important factor cannot be simulated and investigated. Furthermore, factors such as the increased structural complexity, the increased number of interactions related to greater number of organisational sub-units, the potentially more complex referral and communication pathways, the greater distances between units, the greater emphasis on training/education/research as well as various cultural and sociological influences cannot be modelled without data or realistic assumptions. All these factors could be seen as playing a progressively greater role on efficiency, with increasing size of the hospital.

2.2. Data

The model has been calibrated with the patient flow data from the Nambour General Hospital. The latter is a large regional hospital located in the Sunshine Coast, Queensland. In March-April 2017, the hospital underwent a major downsizing (more than 50%) due to a new hospital opening at Birtinya, Sunshine Coast.

The data set includes information about every patient that arrived from January 1, 2013, to December 31, 2016. The details in the data cover: type and time of arrival, triage score, demographic information, diagnosis data, admission decision, admission and discharge wards/units (if the patient was admitted), timestamps for major events (being seen by a doctor/nurse in ED, discharged from ED, admitted to hospital, discharged). As part of the ethics approval process, information related to patient's identity was removed before the data was provided to the researchers.

In total, there were 286,354 patient records over the four-year period in the data set. More than 60% of the arriving patients were admitted to the hospital. All patients treated in the hospital were classified to 718 Diagnosis-Related Groups (DRGs) according to the Australian Refined Diagnosis Related Group (AR-DRG) classification. The distribution of DRGs was heavily skewed: the top 57 DRGs were responsible for more than 50% of admissions. The top 87 DRGs covered 60% of all patients.

Multiple probability distributions were estimated from the data and used as inputs in the model. These included: rates of arrival, probabilities for type of arrivals (ambulance, self-presentation, etc.), DRG,

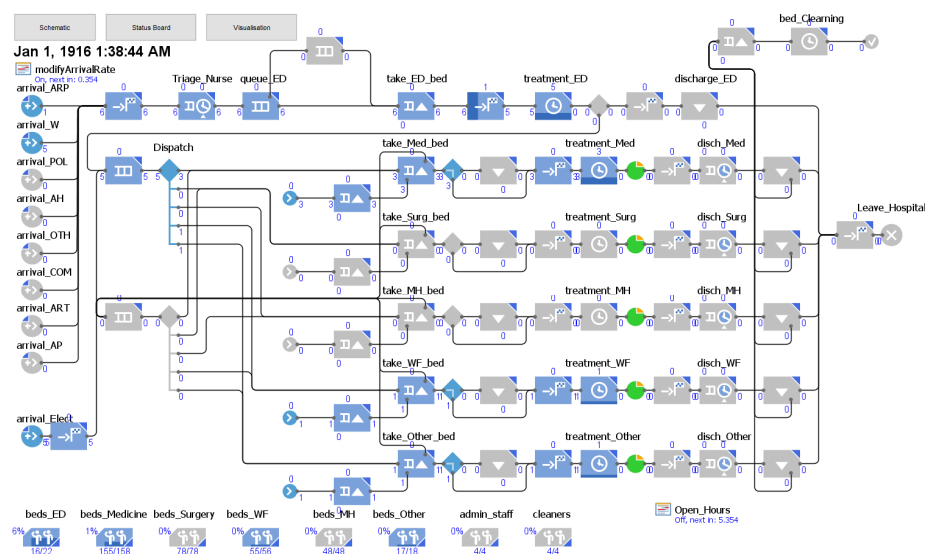


Figure 1. Logical view of the Nambour-HESMAD simulation model

probability of admission, admission and discharge wards, treatment time in the ED and hospital. Some distributions, e.g., the rate of patient arrivals were discretised to one-hour intervals for every day of the week. Other distributions depended on patients' health conditions hence these were estimated for individual DRGs.

With these distributions available, the model generates a triage score, a probability of admission, treatment time in ED, and for those admitted to hospital, admission ward for each patient based on their DRG. DRG codes that accounted for less than 500 cases over the full history in the data set were combined into one group, referred to as "other" as there would otherwise not be enough statistical power for reliable inference.

Besides the variables presented in the data set and used as input variables for the simulation model, there were a number of emergent variables resulting from the simulation. For example, patients' length of stay, boarding time, and ED queue size were largely determined by treatment required as well as bed availability at the time of arrival and admission. These variables were used to first validate the model and then as performance indicators while comparing different simulation scenarios.

3. RESULTS AND DISCUSSION

3.1. Simulation of hospitals of three sizes

The model was run for three different hospital sizes with the following relation to Nambour hospital's original size: 100% – the same size; 50% – half-size; and 200% – double size. These scenarios assumed proportional change in hospital capacity and arrival rates. That is, a hospital with 50% capacity of Nambour hospital had only 50% of typical arrival rate for any given day and time. At the same time, the proportional mix of DRGs remained the same for all hospital sizes.

All scenarios were run from 20 (size 200%) to 80 (size 50%) years to ensure the total patient turnover in each hospital-size scenario was in excess of 2 million patients before data collection, so that the distributions of the measures of interest are accurately estimated. Then all scenarios were run for 100 years of simulated time (equivalent to 100 repetitions of 1-year scenarios) to collect aggregated data on hospital utilisation.

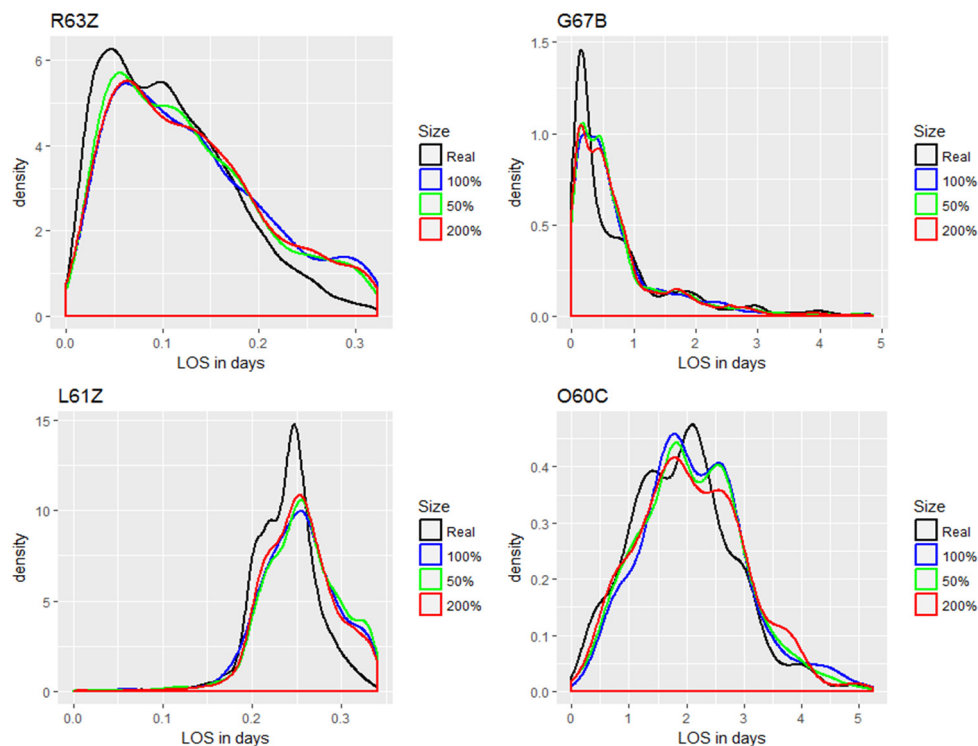


Figure 2. Length of stay (LOS) distribution for four Diagnosis Related Groups

Figure 2 shows distribution of length of stay (LOS) for some popular DRGs truncated at 99% of observations. They are somewhat randomly picked from the most popular DRGs (see Figure 4). The black lines represent the LOS distribution obtained from the real data; blue, green and red lines are 100%, 50% and 200% of hospital size simulations respectively. A close match between the black and the blue lines indicates that the model is

adequate as the 100% – size simulation demonstrates the same overall picture as that represented by the real data. Reasonably good fits between the black and the blue lines are observed for most DRGs.

The green (50%) and the red (200%) lines in Figure 2 are also very close to the 100% – size simulation. This means that the change in hospital size does not affect the LOS distribution. Simplification of the model due to restricted data availability makes it impossible to capture any differences in LOS, which potentially could arise from a hospital size change.

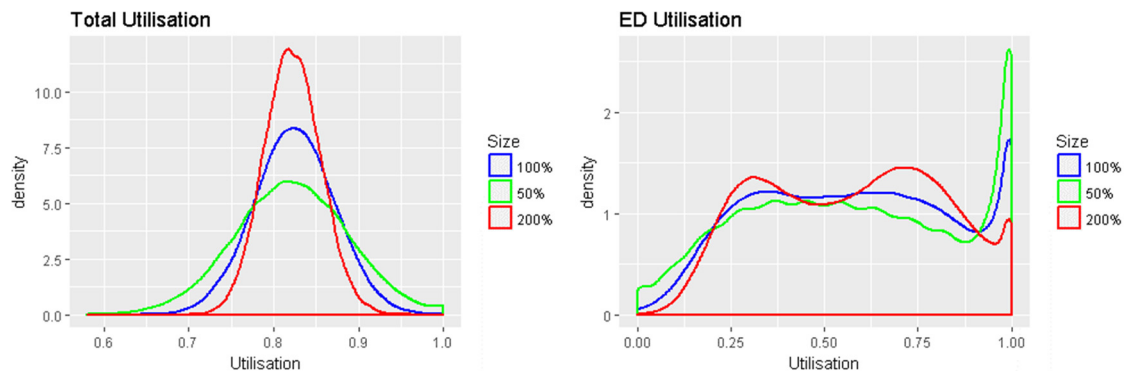


Figure 3. Total hospital utilisation and Emergency Department utilisation

Moving on to the utilisation analysis (i.e. hospital occupancy), the left panel in Figure 3 reports a distribution of the hospital utilisation level for the above three “base” scenarios. Utilisation is defined as the number of occupied beds divided by total number of beds. The distribution in Figure 3 is based on hourly measurements of utilisation, hence can cover the full range of situations in a hospital - from no patients to no free beds. Hourly utilisation data is not available from the real data, hence it is not possible to generate a black curve here for comparison.

Again, on average, there is no difference between 50%, 100% and 200% – sized hospitals: mean utilisation is at 82% for all scenarios. This result is expected, after observing the LOS results. As the hospital sizes and arrival rates are being changed proportionally, and the LOS stays the same, then the average utilisation should also stay the same.

The major difference in hospital utilisation between the three scenarios is the level of dispersion in the utilisation distribution. As the hospital size goes up, dispersion goes down. A larger hospital has the capacity to absorb natural fluctuation in the number of arrivals more efficiently than smaller hospitals, by temporarily redistributing patients to non-designated wards.

The spike on the right-hand side of the green curve indicates that a 50%-size hospital has 100% utilisation on multiple occasions. That is, the hospital was often overcrowded and unable to serve new arrivals.

The right panel in Figure 3 shows distributions of ED utilisation. The big spikes on the right-hand side of the density curves indicates that no one hospital is completely immune to overcrowding in the ED. However larger size hospitals are less likely to be overcrowded. Past academic studies report a close relationship between ED overcrowding and hospital occupancy (Forster et al., 2003; Harris and Sharma, 2010). Hence, it is not surprising that a 50%-sized hospital has a higher probability of 100% hospital utilisation, as well as experiences ED overcrowding more frequently. At the same time, the left-hand end of the green density curve indicates that ED in 50%-sized hospital has a higher probability to have zero patients as well. This is an example of extreme underutilisation – a highly undesirable situation from the managerial and cost perspectives.

The largest queue sizes in ED observed during the simulations were 44, 26 and 26 patients for 50%, 100% and 200%-size hospitals respectively. While these quantities depend on the number of simulation runs, they give an indication of the extreme values that can be observed. It is necessary to stress that the 200%-sized hospital had four times the rate of arrivals of the 50%-sized hospital, but still a smaller ED queue in absolute numbers.

3.2. Down-sized hospitals with different DRG mixes

Scenarios considered in 3.1 used an equal proportional change in hospital size and rate of arrivals, which might be considered as an unrealistic assumption. It might be possible to reduce the number of ambulances arriving to the hospital by 50% – extra ambulances can be diverted to other hospitals. However, self-presentations (or

walk-in patients) are less manageable and might keep arriving at the same or a slightly lower rate, despite the hospital downsizing.

Hence it would be useful to consider scenarios where the 50%-sized hospital receives a “biased” reduction in arrivals. An alternative scenario tested was a 25% reduction in walk-in patients and a 67% reduction in ambulance arrivals. Weights were selected in such a way that the total number of patients stayed the same.

Figure 4 demonstrates a difference in the mix of DRGs between the 50%-sized hospital (“proportional”) with a proportional reduction for arrival rate and a “biased” 50%-sized hospital. Despite the same total number of arrivals, the mix of DRGs is now different. DRGs related to patients predominantly delivered by ambulances have lower representations, while other DRGs are overrepresented. As a result, these two hospitals of equal sizes would operate in very different environments. This change, understandably, leads to a change in all performance indicators. Figure 5 presents distributions of the total hospital utilisation and the ED utilisation for both 50%-sized hospitals – with “proportional” and “biased” reduction in arrival rates/type respectively.

The total hospital utilisation appears much lower in the case of “biased” arrival reduction – on average, only 69%, compared to 82% for the proportional reduction. In general, the walk-in patients have a lower probability to be admitted to the hospital, compared to the ambulance arrivals; hence the overall hospital population is lower, even with the same total number of arrivals.

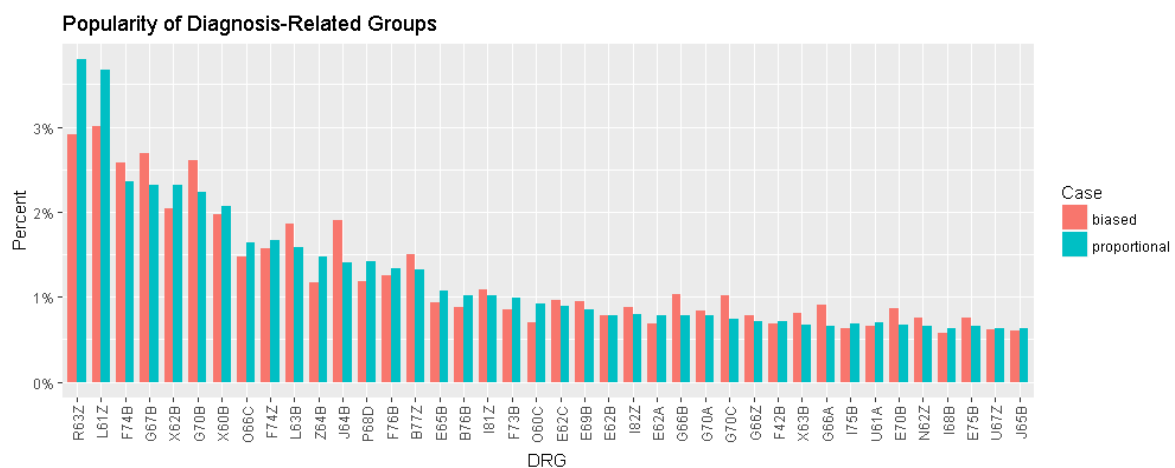


Figure 4. Change in the mix of DRGs for proportional and non-proportional reduction in arrival types.

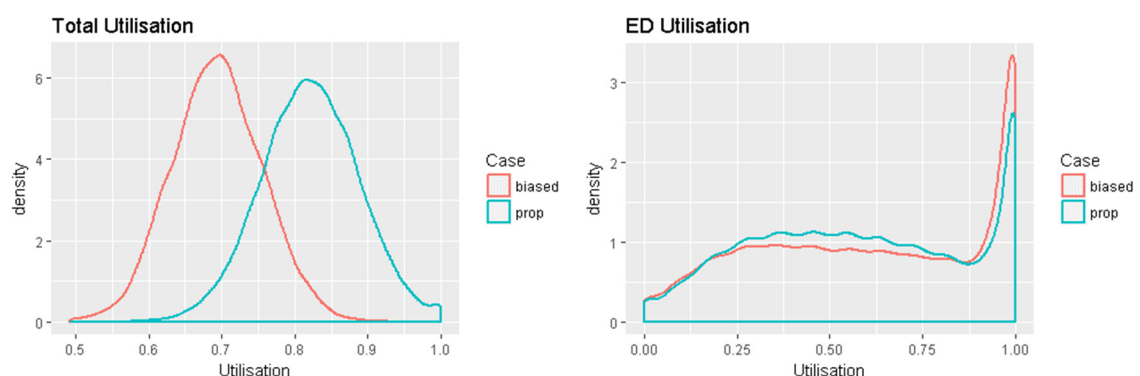


Figure 5. Total hospital utilisation and Emergency Department utilisation for proportional and biased reduction in arrival rates.

ED utilisation levels, on the other hand, appear to be quite similar between the two cases on average – 58% for “proportional”, and 61% for “biased”. However, the ED in the hospital with “biased” arrivals has a higher probability to be overcrowded, as the distribution of walk-in arrivals has much greater variability than distribution of ambulance arrivals. Therefore, an ED reduced to 50% could not properly function on multiple occasions despite a much lower total hospital occupancy. This means that, in addition to the link between the

ED efficiency and the total hospital occupancy established in the prior literature and confirmed by this study, there are other key factors affecting ED performance, such as a type of arrivals (ambulance vs. self-presentation) and the DRG mix.

4. CONCLUSIONS

Computer simulation is a useful tool for modelling complex system behaviours. This study offered a number of novel insights into the relationship between a hospital size and the key performance indicators, including the length of stay and bed utilization.

Our simulations show no change in the averages for the total hospital and ED utilisations as the hospital size changes. However, there was a substantial change in the variability of the outcome measures. Smaller size hospitals exhibited a greater dispersion in the distributions of utilisation and thus would have a much higher propensity for being either overcrowded or under-utilised – both undesirable outcomes.

While it is not possible to categorically conclude that the larger size hospital is more efficient, it is clear that smaller hospitals are at much higher risk. That risk is defined by the higher probability to be overcrowded, which could result in inability to service newly arrived patients and/or provide adequate level of treatment.

Changing the proportion of ambulance and walk-in patients affects the DRG mix of patients presenting at the hospital; this, in turn, changes the overall LOS, the total hospital occupancy and the ED utilisation. The simulated scenario with non-proportional change in arrival rates can be thought of as a new and very different hospital environment. This result highlights the importance of the DRG mix for hospital planning. Different DRG mixes will require different facilities and hospital sizes, even for the same expected arrival rates and/or population served.

A major limitation of the current model is its inability to investigate the possible efficiency-limiting factors that can be logically anticipated with increasing facility size. These are: increased structural complexity, greater educational/training/research burden, slower internal referral and communication pathways, exponentially increasing complexity of interaction as the number of structural sub-units increases, etc.

Further research might focus on developing a variant of the HESMAD model with additional modules and factors. This in turn will enable the impact of more variables in the hospital environment to be evaluated, and provide greater insights for healthcare service managers to deal with complex problems occurring in hospital.

REFERENCES

- Bain, C.A., Taylor, P.G., McDonnell, G., & Georgiou, A. (2010). Myths of ideal hospital occupancy. *Medical Journal of Australia*, 192(1), 42.
- Ben-Tovim, D., Filar, J., Hakendorf, P., Qin, S., Thompson, C., & Ward, D. (2016). Hospital Event Simulation Model: Arrivals to Discharge–Design, development and application. *Simulation Modelling Practice and Theory*, 68, 80-94.
- Brailsford, S. C., Harper, P. R., Patel, B., & Pitt, M. (2009). An analysis of the academic literature on simulation and modelling in health care. *Journal of simulation*, 3(3), 130-140.
- Carr, W.J., & Feldstein, P.J. (1967). The relationship of cost to hospital size. *Inquiry*, 45-65.
- Choi, J.H., Park, I., Jung, I., & Dey, A. (2017). Complementary effect of patient volume and quality of care on hospital cost efficiency. *Health care management science*, 20(2), 221-231.
- Fetter, R.B., Shin, Y., Freeman, J.L., Averill, R.F., & Thompson, J.D. (1980). Case mix definition by diagnosis-related groups. *Medical care*, 18(2), i-53.
- Forster, A.J., Stiell, I., Wells, G., Lee, A.J., & Van Walraven, C. (2003). The effect of hospital occupancy on emergency department length of stay and patient disposition. *Academic Emergency Medicine*, 10(2), 127-133.
- Harris, A., & Sharma, A. (2010). Access block and overcrowding in emergency departments: an empirical analysis. *Emergency Medicine Journal*, 27(7), 508-511.
- Holm, L.B., Lurås, H., & Dahl, F.A. (2013). Improving hospital bed utilisation through simulation and optimisation: with application to a 40% increase in patient volume in a Norwegian General Hospital. *International journal of medical informatics*, 82(2), 80-89.
- Hoot, N.R., LeBlanc, L.J., Jones, I., Levin, S.R., Zhou, C., Gadd, C.S., & Aronsky, D. (2008). Forecasting emergency department crowding: a discrete event simulation. *Annals of emergency medicine*, 52(2), 116-125.
- Richardson, D.B. (2006). Increase in patient mortality at 10 days associated with emergency department overcrowding. *The Medical Journal of Australia*, 184(5), 213-216.