Reducing Patient Delays in a Day Surgery Unit of a Hospital

Victoria Hanna
University of Melbourne, Australia

Kannan Sethuraman
Melbourne Business School, Australia

INTRODUCTION

In this article, we illustrate the relevance of queuing theory principles to the healthcare sector through a case study of a day surgery unit in a hospital in Melbourne, Australia. The hospital has an acknowledged problem: patients are spending longer-than-anticipated periods of time in the day surgery unit, and they are facing excessive waiting times at all stages of their healthcare delivery process. The hospital is operated on a not-for-profit basis, and the executive board members are keen to understand the root causes of the problem, so they can direct their investment in the system to improve its responsiveness.

BACKGROUND

To reduce the time a patient spends waiting in the hospital system, a hospital needs to understand the implications of queuing and the theories that explain the causes and costs of congestion. Little's Law (Little, 1961) states that the average number of patients in a stable system (over some time interval) will equal the product of their average arrival rate and their average time in the system. In manufacturing situations, this law represents the critical link between the three key internal process performance measures: cycle time, work-in-progress (WIP) inventory, and manufacturing lead time (MLT).

- Cycle time characterizes the average time between completions of successive units. In the healthcare context, for instance, the cycle time could represent the average time between completions of successive procedures in a surgical unit.
- WIP inventory represents the number of units contained in the process at any given time. Patients-in-progress (PIP) will be the equivalent measure in the healthcare context, representing the number of patients who are currently being served, and who are waiting for service in the system.
- Manufacturing lead time stands for the total time a unit spends in the manufacturing process. In the healthcare context, the patient resident time (PRT) would be its equivalent. And it represents the total average time a patient spends in the hospital, from the time of his arrival, to his discharge time. This would include all the value-added time and the nonvalue-added time, such as waiting that the patient experiences in the system.

Using the average number of patients residing in the system at any given time, and the surgery throughput capability of the system, Little's Law helps us predict the average time a patient would reside in the system (PRT). In other words, the larger the number of patients waiting for treatment, the longer will be the wait for an arriving patient. An immediate implication of Little's Law is that, for a given level of throughput at any hospital, the only way to reduce the PRT is to reduce the number of patients in the system. For example, assuming patients arrive at the rate of 10 per hour, and stay an average of one hour. This means we should find the average number of patients in the system at any time to be 10. If the arrival rate rises to 20 per hour, the unit must either be prepared to host an average of 20 occupants, or must reduce the time each patient spends in the system to 0.5 hour. Naturally surgery cannot be sped up, so the unit must look to other components of the patients' stay to reduce their time in the system.

This relationship, however, is not obvious to hospital administrators or doctors. The system expects people to wait, and hence provides infrastructure to deal with patients waiting in the system, resulting in escalating costs and poor responsiveness. Long lead times are still accepted by patients, as they have no alternative. What the service providers are missing is an opportunity to simultaneously improve both patient responsiveness and cost effectiveness. Too many patients in the system can be costly in a number of ways. You need more room to hold them, more resources to engage them, monitor, and progress them through the system (hospital), and the excess number of patients in the system acts as a buffer, hiding the linkages between processes. It is critical to observe that the time spent delivering patient care should never be sped up (for example, time for anaesthesia to start working, treatment of wounds, recovery time, and so on), as it compromises quality of care to achieve responsiveness. The focus has to be on simultaneously achieving responsiveness and quality of care through the elimination or minimisation of nonvalue-added activities, such as waiting for service.

How Variability Increases the Time Patients Spend in the System

Variability robs a system of capacity. As a consequence, utilization is reduced, and this affects the time a patient spends in the system. The *Pollaczek-Khintchine* formula (Heyman & Sobel, 1982) provides the fundamental relationship between capacity utilization, variability, and inventory, and is represented by:

$$L = K \times \left(\frac{\rho}{1-\rho}\right) \times \left(V_a + V_s\right)$$

where K is a constant, ρ represents the capacity utilization, V_a characterizes the variability in arrivals, V_s characterizes the variability in the service process, and L stands for the queue length.

If patients undergoing treatment are a form of inventory (analogous to work in progress), the above relationship has important consequences for a hospital trying to reduce patient delays.

- In steady state, if a hospital system experiences higher levels of variability (either in the patient arrival process, or in the patient care delivery process, or both), it will essentially result in an increased patient queue length. From Little's Law, we can infer that increased patient queue length will lower responsiveness, or in other words, increase average patient resident time.
- In highly unpredictable environments (such as an emergency department), the patient queue length and the average patient resident time tend to increase exponentially with increased capacity utilization. If better information cannot be obtained to lower the level of variability in the patient arrival process, the system will need to operate at lower levels of utilization to achieve the target service levels. That is, they will need buffer capacity. Attempts to fully utilize capacity will result in extremely poor responsiveness—a situation unlikely to be acceptable to patients or healthcare providers.

Litvak and Long (2000) discuss how healthcare systems exhibit variability, and categorise them into three broad types: (i) *clinical variability,* which represents the differences in the patients' degree of illness, choice of treatment alternatives, and responses; (ii) *flow variability,* which captures the randomness in the arrival patterns of these patients; and (iii) *professional variability,* that describes the variability in the ability of the medical practitioners and healthcare delivery systems to provide treatment. The presence of clinical, flow, and professional variability results in a significant increase in complexity, and adds cost to healthcare systems. In addition to these natural variabilities, they also identify an artificial variability, which is an artefact of dysfunctional management and policies. They conclude that eliminating artificial variability in healthcare systems holds the maximum potential for reducing waste and improving responsiveness.

These theoretical underpinnings helped us to explain why the day surgery unit in our case was experiencing excessive wait times for its patients, and our analysis of the situation focused on the following:

1. **System design:** If the patient care delivery process is poorly designed (for example, redundant activities, activities occurring in sequence instead of in parallel, and so on), then the total time it takes for a patient to progress through the system will be longer than ideal.
2. **Waiting times at key stages of the process:** Even when a system is well designed, the wait times could be poor, due to constrained resources acting as bottlenecks. Since all systems exhibit some variation, queuing theory provides key insights

Related Content

The Role of ICTs in the Management of Rare Chronic Diseases: The Case of Hemophilia
Leonor Teixeira, Vasco Saavedra, Carlos Ferreiraand Beatriz Sousa Santos (2013). *Handbook of Research on ICTs and Management Systems for Improving Efficiency in Healthcare and Social Care (pp. 635-649).*
www.irma-international.org/chapter/role-icts-management-rare-chronic/78047

Exploring Linkages between Quality, E-Health and Healthcare Education
Christopher L. Pateand Joyce E. Turner-Ferrier (2010). *International Journal of Healthcare Delivery Reform Initiatives (pp. 66-81).*
www.irma-international.org/article/exploring-linkages-between-quality-health/53873

Development of an Emergency Response Management using Mobile Devices for Hospital Infrastructures Affected by Power Grid Failures
Safa Attia, Abdelhak Boubetraand Manel Saad Saoud (2016). *International Journal of Healthcare Information Systems and Informatics (pp. 36-57).*
www.irma-international.org/article/development-of-an-emergency-response-management-using-mobile-devices-for-hospital-infrastructures-affected-by-power-grid-failures/155116

Information as Change Agent or Barrier in Health Care Reform?
Alan C. Gilliesand John Howard (2013). *Healthcare Information Technology Innovation and Sustainability: Frontiers and Adoption (pp. 1-17).*
www.irma-international.org/chapter/information-change-agent-barrier-health/73811

Emergency Department Logistic Optimization Using Design of Experiments: From Triage to High Quality of Service
Jose Antonio Vazquez-Ibarra, Rodolfo Rafael Medina-Ramirezand Irma Jimenez-Saucedo (2018). *Health Care Delivery and Clinical Science: Concepts, Methodologies, Tools, and Applications (pp. 386-400).*
www.irma-international.org/chapter/emergency-department-logistic-optimization-using-design-of-experiments/192682