

**Modeling Community Care Services for Alternative level of Care
(ALC) Patients:
A Queuing Network Approach**

Pedram Noghani Ardestani

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
for the MSc degree in Systems Science

University of Ottawa

© Pedram Noghani Ardestani, Ottawa, Canada, 2014

Abstract

One of the impacts of the rising demand for community health services, primarily used by seniors, is that hospitals are often faced with the challenge of having patients finish the acute phase of their treatment and yet are unable to discharge them due to the lack of a bed in a more appropriate community care setting. The frequency of this challenge has led to the designation of “alternative level of care” (ALC) being ascribed to patients who remain in the hospitals due to insufficient capacity downstream. The thesis focuses on a model that seeks to address patient flow through the community care network (CCN) and finding capacity allocation policies for the different facilities that resolves the ALC challenge using scenario analysis. A queuing network model with general routings and nodes’ blocking has been developed and a heuristic approximation method has been employed for solving the model. Blocking probabilities and the number of blocked patients are derived as performance metrics of the CCN. We test the accuracy of the queuing model through a simulation model and the behaviours of the system in different scenarios are investigated in the simulation model and our policy insights and conclusions are provided.

Acknowledgements

I would like to express my deepest gratitude to all who made this thesis possible and helped through this learning phase of my academic career.

First and foremost, I would like to offer my sincerest gratitude to my supervisor professor Jonathan Patrick for his extraordinary supervision and contribution to this research. I benefited immeasurably from Professor Patrick's invaluable assistance, support, and encouragement. His insightful direction, and support in various ways will never be forgotten. My special thanks go to Jamie Stevens and Bruce Libman from CCAC and Champlain LHIN for providing necessary data and their kindly contributions and consultations on the details of the modeling process.

At last, I would like to thank my friend and partner, Anoosheh, for her tremendous support.

Contents

Introduction.....	1
1.1. The ALC Challenge.....	1
1.2. ALC Destinations	3
1.3. Research Objectives.....	7
2. Literature review.....	8
2.1. Related studies.....	8
2.2. An Introduction to Queuing Theory	10
2.2.1. Basic Queuing Models.....	10
2.2.2. Queuing Networks	12
2.2.3. Open Jackson Networks	13
2.3. Queuing Models in the Literature.....	14
2.3.1. Queuing Network Models with Blocking.....	15
2.4. Approximation Methods for Open Networks with Blocking	18
2.4.1. Node-by-Node Decomposition Method by Takahashi et al.	18
2.4.2. Heuristic Method by Bretthauer et al.....	20
3. Research Methodology	24
3.1. The CCN as an Open Queuing Network	24
3.2. Queuing Network Model	24
3.3. Simulation.....	27
4. Results	28
4.1. Data.....	28
4.2. Queuing Model Results	39
4.2.1. Base Case Scenario:.....	40
4.2.2. Scenario 1: Increasing LTC capacity.....	41
4.2.3. Scenario 2: Increasing demand	43
4.2.4. Scenario 3: Mixed policy.....	44
4.3. Simulation Results.....	46
4.3.1. Base Case Scenario:.....	46
4.3.2. Scenario 1: Increasing LTC capacity.....	48
4.3.3. Scenario 2: Increasing demand	51

4.3.4. Scenario 2: Mixed policy.....	52
5. Limitations of the model.....	54
6. Conclusions and Further Extensions.....	55
Appendix I: ALC Discharge Destinations' Description.....	57
Appendix II: Abbreviations.....	58
References.....	59

List of Figures

Figure 1-Discharge Destination of ALC Patients in Ontario, 2009 (%) [3]	5
Figure 2- ALC queuing network.....	6
Figure 3- Basic queuing system.....	10
Figure 4- Open Queuing Network	12
Figure 5- Inflows and outflows between stations in Koizumi et al.'s [27] model.....	16
Figure 6- Inflows and outflows between stations in Bretthauer.'s et al. [28] model.....	17
Figure 7- Acute care arrival's probability distribution	30
Figure 8 - Acute care length of stay's probability distribution	31
Figure 9- LTC arrivals.....	32
Figure 10-Probability distribution of CCAC patients' waiting days before discharging to LTC	33
Figure 11- Number of patients added the LTC waiting list.....	34
Figure 12- CCAC regular service admissions probability distribution.....	35
Figure 13- Weekends CCAC regular service admissions probability distribution	36
Figure 14- Weekdays CCAC regular service admissions probability distribution	36
Figure 15- CCN and routing probabilities	38
Figure 16- Average utilization of the CCN facilities- Base case scenario	47
Figure 17- Average utilization of the CCN facilities- Scenario 1- Case I.....	49
Figure 18- Average utilization of the CCN facilities- Scenario 1- Case II.....	50
Figure 19- Average utilization of the CCN facilities- Scenario 2	52
Figure 20- Average utilization of the CCN facilities- Scenario 3	53

List of Tables

Table 1-Scope of Alternate Level of Care by Province, 2007-2008	3
Table 2-Summary of available data on ALC and LTC patients	28
Table 3- Discharge destinations of ALC patients- Apr. 2009- Mar 2012	29
Table 4- Discharge destinations of LTC patients- Apr. 2009- Mar 2012.....	31
Table 5- CCAC patients routing probabilities	34
Table 6- CCN patients routing probabilities.....	37
Table 7- LOS and Capacity of the CCN nodes.....	38
Table 8- Blocking probabilities (%) in different scenarios introduced by Bretthauer et al model and modified heuristic model	39
Table 9- Blocking probabilities for CCN facilities- Queuing model –Base case scenario.....	40
Table 10-Steady states number of blocked patients at each stage-Queuing model Base case scenario	40
Table 11- Blocking probabilities for CCN facilities- Queuing model -Scenario 1-Case I.....	41
Table 12- Steady states number of blocked patients at each stage-Queuing model–Scenario 1case I..	42
Table 13- Blocking probabilities for CCN facilities- Queuing model -Scenario 1-Case II.....	42
Table 14- Steady states number of blocked patients at each stage-Queuing model–Scenario 1caseII..	42
Table 15- Blocking probabilities for CCN facilities- Queuing model –Scenario 2.....	43
Table 16- Steady states number of blocked patients at each stage- Queuing model –Scenario 2	44
Table 17- Blocking probabilities for CCN facilities- Queuing model –Scenario 3.....	45
Table 18- Steady states number of blocked patients at each stage- Queuing model –Scenario 3	45
Table 19- Blocking probabilities for CCN facilities- Simulation model – Base case scenario	46
Table 20- Average number of blocked patients at each stage-Simulation model-Base case scenario	46
Table 21- Blocking probabilities for CCN facilities- Simulation model -Scenario 1-case I.....	48
Table 22- Average number of blocked patients at each stage - Simulation model- Scenario 1-caseI..	48
Table 23- Blocking probabilities for CCN facilities- Simulation model- Scenario 1 case II.....	49
Table 24- Average number of blocked patients at each stage- Simulation model Scenario 1-case II..	49
Table 25- Blocking probabilities for CCN facilities- Simulation model- Scenario 2.....	51
Table 26- Average number of blocked patients at each stage - Simulation model- Scenario 2	51
Table 27- Blocking probabilities for CCN facilities- Simulation model- Scenario 2.....	53
Table 28- Average number of blocked patients at each stage - Simulation model- Scenario 2	53

Introduction

The age structure of a population is a function of fertility and mortality rates as well as immigration. In Canada, over the last few decades, increasing life expectancy and a decreasing birth rate have resulted in an aging population. The number of Canadian seniors (over 65) increased from 2.7 million in 1986 to 4.8 million in 2010 and at the same time their proportion of the Canadian population rose from 10% to 14%. By 2031, the baby boom generation (born between 1946 and 1965) will all be over 65 with the number and proportion of seniors continuing to increase at a significant pace until that time. After 2031, population aging is also expected to continue, but with a lower growth rate[1].

Population aging is in part because Canadian seniors are living longer. This trend highlights the prudence of assessing the health care needs of seniors and the resulting health system demands in strategic planning. One important aspect of the health system design is patient movement across the system. It is reasonable to assume that population aging directly impacts on patient flow. Compared to younger people, seniors remain in emergency departments and acute care settings more frequently and for longer periods. In part, the longer lengths of stay of seniors in hospitals are because they require more assistance after discharge and quite often that additional assistance is not readily available. For these reasons, it is important to consider the issue of an aging population and its future health care demands through modelling and simulation of patient flow.

1.1. The ALC Challenge

One of the impacts of the rising demand for community health services, primarily used by seniors, is that hospitals are often faced with the challenge of having patients finish the acute phase of their treatment and yet are unable to discharge them due to the lack of a bed in a more appropriate community care setting. The frequency of this challenge has led to the designation of “alternative level of care” (ALC) being ascribed to patients who remain in the hospital due to insufficient capacity downstream. According to the Canadian Institute for Health Information (CIHI), the ALC designation pertains to an individual who “has finished the acute care phase of his/her treatment but remains in the acute care bed”[2]. When the

physician discharges a patient, who no longer requires hospital care, under the Public Hospital Act, the patient should leave the hospital within 24 hours. Some patients, however, remain in hospital awaiting placement in another facility. The ALC waiting period starts at the time of designation and ends at the time of discharge. According to CIHI data, in 2007-2008, the median ALC length of stay was 10 days but there was significant variation. Some patients spent only a few days as ALC while others spent several months. Data shows that 59% of ALC patients stayed at hospital more than a week and 20% experienced more than a month of ALC days¹. Ministry of Health and Long-Term Care reports that about 85% of ALC patients in Canada are over 65 and 35% are older than 85[3]. The ALC issue has been growing in proportion to the aging of the population and thus is not expected to improve for quite some time. This research aims to address the issue of ALC patient flow in by addressing the larger issue of capacity planning for community health services for the elderly.

ALC patients occupy a significant number of hospital beds and cause several challenges for hospital management. According to CIHI statistics, in 2007-2008, there were more than 74,000 ALC hospitalizations and more than 1.7 million ALC days in Canada excluding Quebec and Manitoba (This data is obtained by removing obstetric and pediatric patients from the analysis because they had distinctive demographic profiles and few reported ALC days.)². This is equal to 5% of all hospitalizations and 14% of all hospital days in acute care hospitals[4]. Table 1 shows the scope of ALC patients and the percentage of hospitalizations related to ALC across Canada.

¹ Information gathered from CIHI website

² For more information see [4]

Province	Number of hospital bed-equivalents used for ALC	Percent of hospitalizations that were ALC related
British Columbia	910	5%
Alberta	520	3%
Saskatchewan	150	2%
Ontario	2590	7%
New Brunswick	340	5%
Nova Scotia	460	3%
Prince Edward Island	30	2%
Newfoundland and Labrador	180	7%

Table 1-Scope of Alternate Level of Care by Province, 2007-2008 [4]

Ontario has the highest ALC rates among the provinces in Canada[5]. According to CIHI data, in 2009, ALC discharges comprised about 5% of all hospital discharges, but the proportion of days that ALC patients were hospitalized was 16% of the total bed days for the year. Moreover, the total number of days that ALC patients were hospitalized has increased by 75% from 2005/06 to 2009/10 while during the same period the total days of hospitalization for all patients only increased by 7%. A monthly survey conducted by the Ontario Hospital Association shows that ALC patients occupied 16% of all acute care beds in the province as of June 2010[3].

While ALC patients are waiting for a number of destinations (see Figure 2), LTC is consistently the one with the longest wait times. Data suggests that to place all seniors already on the waiting list for LTC, more than 130,000 beds will be required by 2021 - an almost 75% leap in bed capacity[5]. By descriptive analysis, Costa et al. [6] demonstrate that in Ontario most of the ALC patients who are waiting for LTC placement may be more appropriately cared for in the community setting given adequate community resources.

1.2. ALC Destinations

The aforementioned introduction demonstrates that one of the most important challenges of the health care system in Canada is the effective discharge of patients, especially seniors, from hospitals. As indicated before, a large proportion of hospital beds are occupied by older

patients who don't really need such acute facilities. This means that these ALC patients do not receive the appropriate services while those who need acute care are forced to wait, leading to canceled surgeries and overcrowding in emergency departments. Moreover, prolonged stays in the hospital can also lead to deleterious effects on the ALC patient's health. Finally, keeping patients in a hospital is more costly than any other community setting. Estimated daily costs of an ALC patient at hospital is about 450\$[3]. For all these reasons, it is highly desirable for a health system to determine the necessary capacity for community services in order to prevent insufficient capacity downstream causing upstream congestion.

The main destination of ALC patients is a LTC facility accounting for 43% of all ALC patients in 2007-2008. In addition, 27% of ALC patients were transferred to their homes and 12% of them died in hospital[4]. Data shows that ALC patients discharged to rehabilitation facilities or home had shorter median ALC stays than those discharged to LTC (15 days for LTC in comparisons to 6 and 7 days for rehabilitation and home respectively, in 2007-2008)[3]. The main goal of ALC discharge planning is to ensure that patients receive appropriate support and services outside the hospital. Nevertheless, according to CIHI data, in 2007-2008, about 17% of ALC patients discharged to their homes had at least one readmission to hospital within 30 days of discharge. This percentage was 27% for home-discharged ALC patients in Ontario[4].

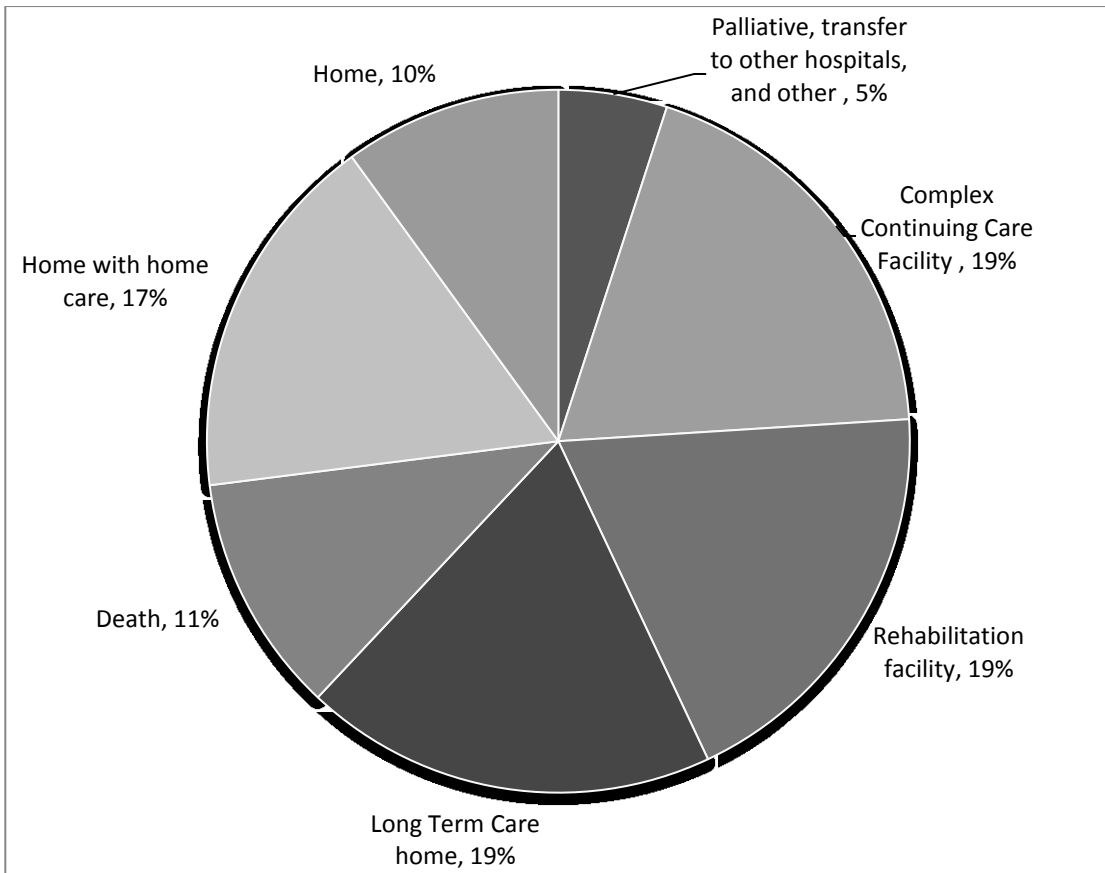


Figure 1-Discharge Destination of ALC Patients in Ontario, 2009 (%) [3]

Discharge destinations of ALC patients include (but are not limited to) one of the following [7] (see Figure 1):

- ALC-CCC (complex continuing care)
- ALC-Home First
- ALC-Rehab
- ALC-LTC
- ALC-Housing
- ALS-HRS
- ALC-Mental health beds
- ALC-Hospice

The interaction among the different discharge destinations of ALC patients is depicted in Figure 2.

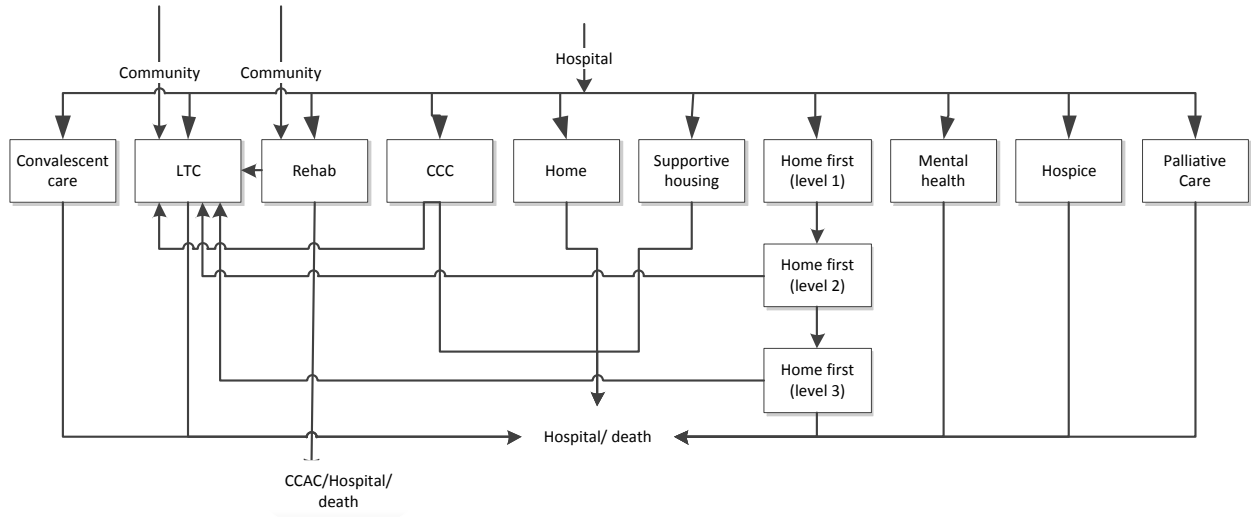


Figure 2- ALC queuing network

As shown in figure 3, there exists at least twelve possible ALC destination facilities (nodes). We will call this the community care network (CCN). The definition of each node is provided in the Appendix. “Home first” is a program developed in order to ease elderly patients back into living at home so as to avoid sending them to LTC. There are three types of “home first” shown in the ALC patient network. Home first level 1 indicates care for the first 60 days post hospitalization (at more than 90 hours of service per week), then patients step down to home first level 2 to receive service for the next 30 days (90 hours of service per week). Finally they step down to home first level 3 (regular community service). The appropriate destination of each ALC patient is determined by their RAI score and service guidelines.

As illustrated in the diagram, the demand for the facilities in the CCN come from the hospitals as well as the community. In other words, patients in the hospital have to compete with others from the community in order to get access to the same resources. These community clients already experience lengthy wait times to access the appropriate care. Any attempt to model capacity requirements for the various facilities in the CCN cannot afford to ignore this important segment of demand. Thus, our model will include both ALC demand for the CCN as well as demand that arises directly from the community. A well performing system must not only maintain reasonable and steady ALC numbers in the hospitals but also maintain reasonable wait times for service in the community.

1.3. Research Objectives

The research question that we seek to answer is the following:

What is the capacity allocation plan for the various facilities in the CCN such that the census of ALC patients in the hospitals is held below a pre-determined threshold?

The imposition of performance measures such as wait time targets or ALC census levels necessarily imposes a capacity requirement but the process of moving from the performance goal to the capacity requirement is far from straightforward. This is especially true in a system that involves multiple patient classes competing for the same pool of resources and where patients move through the network facilities in a stochastic fashion. This thesis presents a mathematical model that will aid decision makers in making the connection between performance and capacity.

Modelling also provides the ability to predict the cost of inefficiency in the current CCN management and as a result, we aim to suggest some effective policies toward a more cost-effective system specifically focusing on capacity allocation across the network of facilities.

This study has been divided in 4 separate sections. In section 2, literature that address the patient flow modelling is reviewed and the approximation methods for solving the queuing network problem will be developed. In section 3, we will introduce the research methodology and the details of the modified queuing model. In Section 4 we begin by providing the details of available data and follows by presenting the results of the queuing and simulation model. Finally, in section 5 the results of the two approaches are compared and conclusion and further studies are discussed.

2. Literature review

2.1. Related studies

In the field of patient flow modelling, the majority of papers demonstrate methods for improving surgical scheduling or more efficiently managing emergency departments. For instance, Jun et al.[8] and Jacobson et al.[9] provide a review of discrete event simulation (DES) in health care system. Belien and Demeulemeester [10] use a mixed integer programming (MIP) model and minimize the expected shortage of beds in hospitals. Chow et al. [11] combine a Monte Carlo simulation model and a MIP model to schedule operating rooms and Thompson et al. [12] use a Markov Decision process (MDP) approach to model the emergency department congestion problem as a finite-horizon, discrete-time, stationary Markov Decision Process. In contrast, it seems modelling the outflow of patients from acute care to other facilities is understudied.

There are few studies that develop models for patients flow after discharging from acute care to other facilities such as LTC or home and community care (HCC). The main goals of such studies are how to predict the future demand and determine the optimal capacity of LTC facilities. For instance, Hare et al. [13] develop a deterministic multistate Markov model for HCC in British Columbia (BC), considering publicly funded residential care as well as home care and non-publicly funded care for predicting the future demand. They consider the predicted demographic changes in BC and the relationship between age and health status in their model by dividing the patients in different age categories. Another example is Intrevado et al.[14] that use large scale mixed integer program for expanding the LTC network in Montreal. They incorporate location, modular capacity acquisition and patient assignment decision in their model and determine the necessary capacity levels of each LTC facility in a time horizon to react to future increasing demand appropriately. They have also considered how geographically locating the LTC facility can impact on patient quality of life. Cardoso et al. [15], Lin et al. [16], Zhang et al.[17] and Zhang and Puterman [18] also find the optimal capacity allocation for LTC facilities through different approaches. Cardoso et al. [15] use simulation model based on a Markov cycle tree structure to predict demand for different types of LTC in Portugal for the period of 2010-2015. They use a comprehensive model that

considers the different health and socio-economic characteristics of patients to predict demand for LTC in details. Lin et al. [16] apply optimal control theory to determine the optimal capacity for publicly funded community based LTC facilities. They conclude that such facilities are less expensive than nursing home care in which providing housing for patients are costly. Zhang et al. [17] and Zhang and Puterman [18] integrate demographic and survival analysis, DES and optimization in order to find optimal capacity level of LTC over a multiyear planning to achieve the predetermined wait time service levels. They use demographic and survival analysis for predicting the LOS and patients' arrival rate as inputs of their simulation model. The main contribution of their model is varying the arrival and service rates of different types of patients by classifying the clients based on their age and gender.

All of the aforementioned studies focus on LTC as a separate node in the health care system and does not consider other types of services which are linked to the LTC and make the system as network. Moreover, there are very few studies that consider number of ALC patients as their parameter of interest in their model. One exception is the research presented in Patrick[19] that presents a MDP model to determine a policy for hospital placements to LTC in order to maintain the ALC census in the hospital below a pre-specified threshold. While the MDP policy presents a solution to the congestion in the hospitals by ensuring a predictable and manageable census of ALC patients in the region, Patrick demonstrated that without increased capacity, the wait times for clients attempting to access LTC directly from the community are well above the 90 day target. Patrick developed a simulation model to demonstrate the necessary increase in Supportive Housing (SH) capacity coupled with a reduction in length of stay in LTC in order for the current capacity in LTC to be sufficient to both maintain stable ALC numbers in the hospitals and meet the wait time target of 90 days for community demand.

In our study, the model will expand to consider the capacity allocation for all the facilities in the CCN rather than focusing solely on LTC and SH. With the larger number and types of facilities, an MDP model is no longer appropriate – partly because we are more interested in capacity allocation rather than scheduling and partly because the size of the model would make the MDP intractable. Instead, we will build a queuing network model to help determine the required capacity.

2.2. An Introduction to Queuing Theory

2.2.1. Basic Queuing Models

The history of queuing theory goes back 100 years when A.K. Erlang (1913) analyzed queues in telephone facilities. Since then, queuing theory has been applied extensively in the industrial and retail sectors. The use of queuing theory in health care, however, is recent and its applications include minimizing waiting times (patients, nurses, etc) and maximizing resources utilization (beds, operation rooms, etc).

A basic queuing system involves customers arriving according to a certain distribution who queue for service at a station (consisting of one or more servers). Service times also follow a known distribution. After service, the customers exit the system.

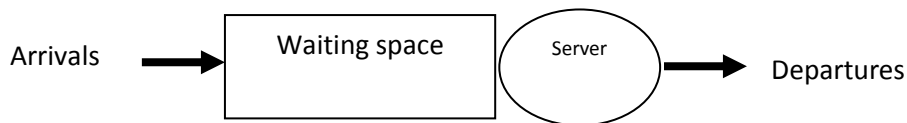


Figure 3- Basic queuing system

A queuing system has six basic characteristics [20]:

- (i) *Arrival pattern of customers:* This is specified by the distribution of inter-arrival times of customers. The mean arrival rate or mean inter-arrival time are two important quantities which can describe the arrival pattern. Since these two measures are related, either of them suffices in describing the system input. A commonly used distribution for the inter-arrival time of customers is an exponential distribution. It is also necessary to know the reaction of a customer upon entering the system. A customer may choose to wait in the queue, may leave if the wait time gets too long (reneging) or may leave immediately if the queue is already a certain size (balking). If there is one or more parallel lines then customers can switch from one to another in order to jockey for position. A time-independent arrival pattern is called a stationary arrival pattern and the one that is time-dependent is called non-stationary.
- (ii) *Service pattern of servers:* This is specified by the distribution of time for a service completion. The service pattern can be described by the number of customers served per unit of time or by the time required to serve a customer. As

with the arrival pattern, a common distribution for the service pattern is the exponential distribution. If the service length is dependent on the number of customers then it is referred to as state-dependent service. A service pattern can also be stationary or non-stationary.

- (iii) *Queue discipline*: The manner by which customers are selected for service is referred as the queue discipline. The most commonly used discipline is first come first served (FCFS) or first in, first out (FIFO). Other queue disciplines include last come first served, or a variety of priority disciplines where customers are given priorities based on some criteria and the ones with the higher priority are served first regardless of the length of wait.
- (iv) *System capacity*: A queuing system may have a finite sized waiting room. Therefore, if the line reaches capacity, no further customers can enter to the system until space becomes available following a service completion. This is called blocking.
- (v) *Number of service channels*: The number of parallel servers that can service customers simultaneously is referred to as the number of service channels.
- (vi) *Number of service stages*: A queuing system may have one or more than one stages. In a multistage queuing system, each customer should proceed through several stages receiving service from a number of servers.

The common notation for a queuing system was first introduced by Kendall. The notation is in the form of $A/B/C/X/Z$ where A and B refer to arrival and service patterns, respectively; C refers to the number of servers; X is system capacity, and Z is system discipline. For example the notation $M/M/2/\infty/FCFS$ indicates a queuing process with exponential inter-arrival times and service times, two servers, an infinite buffer (waiting room) capacity and a first come, and first served discipline.

2.2.2. Queuing Networks

In many real world systems (including the one in question), customers are served in more than one station arranged in a network structure. This network structure consists of nodes connected by paths. A node is a group of servers operating from the same facility. In a queuing network, customers may demand service from more than one server and all customers may not be served by the same set of servers. Figure 4 shows a simple queuing network in which directional arrows show the sequencing of the service.

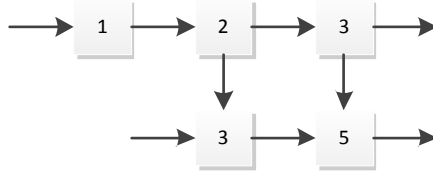


Figure 4- Open Queuing Network

All nodes of the network represent queues and the total number of customers in the network is $\sum_i Q_i(t)$ where $Q_i(t)$ is the number of customers at node i at time t . When the total number of customers in the network is constant, we have a *closed network*. In other words when the arrival rate and departure rate of the network in the system are the same, the system is a *closed network*. If it varies in time, the network is called an *open network* [21].

Consider a network consisting of k nodes. A customer moves to node j with probability P_{ij} ($i, j = 1, 2, \dots, k$) after being served at node i . State 0 denotes the state where customers arrive to the system as well as the absorbing state upon exit. The transition probabilities of this state are as follows:

$$P_{00} = 0$$

$$P_{0i} \geq 0; i = 1, 2, \dots, k$$

$$P_{j0} \geq 0; j = 1, 2, \dots, k$$

The transition probability matrix P which is also known as the routing matrix can be represented as:

$$P = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & \dots & k \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \\ k-1 \\ k \end{matrix} & \left[\begin{array}{cccccc} 0 & P_{01} & P_{02} & \dots & P_{0k} \\ P_{10} & P_{11} & P_{12} & \dots & P_{1k} \\ P_{20} & P_{21} & P_{22} & \dots & P_{2k} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ P_{k-1,0} & P_{k-1,1} & P_{k-1,2} & \dots & P_{k-1,k} \\ P_{k0} & P_{k1} & P_{k2} & \dots & P_{kk} \end{array} \right] \end{matrix} \quad (1)$$

The rate of customers pass through each node is defined as the throughput of that node. Let λ_i be the arrival rate of customers at node i and $\lambda = (\lambda_0, \lambda_1, \dots, \lambda_k)$. At steady state we have $\lambda P = \lambda$ which is the basic equation for obtaining the limiting distribution of the Markov chain in the absence of blocking. The solution to this equation gives us the relative throughput in the network. Blocking occurs when there is a finite waiting room between two stations and customers with completed service in the first have to wait for completion of an ongoing service at the second station because the waiting room between the two stations is full.

2.2.3. Open Jackson Networks

An open Jackson network is a Markovian network in which customers from node i proceed to an arbitrary node j while fresh customers may also join that node from outside the system. In an open Jackson network, customers arrive from outside to node i following a Poisson process with rate $\lambda_i > 0$ and service times at node i are exponential with mean $1/\mu_i$. Each node represents a $M/M/s$ queue with s_i servers. There is no blocking between nodes as each node is assumed to have an infinite sized waiting room. Also, reneging and balking are not allowed. Let α_{ij} be the probability that a customer upon completing service at node i goes to node j , $i \neq j$ and let α_{i0} be the probability that the customer will leave the network after being served at node i . The input to the i th node consists of the outputs from other nodes as well as the external input λ_i . The total average arrival rate of customers γ_i to node i is the sum of the Poisson arrival rate of λ_i from outside plus the arrival rate from arrivals to node i from other internal nodes $\sum_j \alpha_{ij}\gamma_j$.

Therefore we have:

$$\gamma_i = \lambda_i + \sum_j \alpha_{ij}\gamma_j, i = 1, 2, \dots, k. \quad (2)$$

γ_i gives the effective arrival rate to node i or effective rate of flow through node i .

Jackson [22] demonstrated that for a Jackson network of (Markovian) queues, the limiting distribution of the network is the product of the limiting distributions of $M/M/s_i$ queuing systems in equilibrium, implying the independence of nodes in the network. Let Q_1, Q_2, \dots, Q_k be the number of customers in the k nodes when $t \rightarrow \infty$, and define $p_{n_1 n_2 \dots n_k} = P(Q_1 = n_1, Q_2 = n_2, \dots, Q_k = n_k)$

Jackson demonstrated that

$$p_{n_1 n_2 \dots n_k} = p_1(n_1) p_2(n_2) \dots p_k(n_k) \quad (3)$$

Where

$$p_i(r) = \begin{cases} p_i(0) \frac{\left(\frac{\gamma_i}{\mu_i}\right)^r}{r!}, & r = 0, 1, 2, \dots, s_i \\ p_i(0) \frac{\left(\frac{\gamma_i}{\mu_i}\right)^r}{s_i! s_i^{r-s_i}}, & r = s_i, s_i+1, \dots, \end{cases} \quad (4)$$

Given λ_i and α_{ij} , γ_i can be calculated. By using the normalizing condition (sum of all probabilities should be equal to 1) $p_i(0)$ for $i = 1, 2, \dots, k$ can be determined as follows:

$$\sum_{n_1} \sum_{n_2} \dots \sum_{n_k} p_{n_1 n_2 \dots n_k} = 1 \quad (5)$$

In the next section, we present some examples of queuing models applied to healthcare with a particular focus on those models that allow for blocking - a major aspect of patient flow in our current application.

2.3. Queuing Models in the Literature

Several studies present applications of queuing theory to patient flow through a health care system. As an example, Green and Nguyen [23] use a queuing approach to demonstrate that a reduction in the hospital length of stay leads to a reduction in hospital admittance delays. They allocate hospital beds to gain the potential impact of cost-cutting strategies on patients' waiting time and identify the major factors that have an impact on the trade-off between hospital occupancy levels and delays. They use M/M/s queuing models to investigate two types of hospital services- obstetrics and surgery- to understand the relationship between size, occupancy levels, and patient delays.

2.3.1. Queuing Network Models with Blocking

While many studies use queuing network theory, there are very few that consider finite waiting lines in their models and thus deal with the inevitable downstream blocking that is common within healthcare. Certainly, within the CCN, blocking is likely to play a major role as a patient in an upstream facility cannot move on unless capacity is available at the next station. Weiss and McClain [24] present a queuing system with state-dependent service rates in order to model ALC discharges. Their model consists of two stages in which the first stage has infinite capacity but the second stage is capacitated. They present three categories of patients in a hospital: “acute”, “pre-ALC acute” and “ALC”. The acute category includes patients for whom extended care is not necessary. The second category, pre-Acute ALC, encompasses patients who will need extended care but currently require acute care. The third category comprises patients whose acute phase is complete but who remain in the hospital because of a lack of extended care capacity. Using a state dependent service rate approach, Weiss and McClain [24] predict the probability distribution of the ALC census. They emphasize that any change in the discharge planning budget would clearly impact on the number of ALC patient days and the quality of care delivered. In comparison to their two-stage model, our model presents a more general form of multi stage queuing network with blocking in multiple facilities.

While Jackson [22] shows how to obtain the equilibrium probability distributions of customers in an open (i.e. one with external demand as well as exits from the system) queuing network without any blocking, finding the exact solution for the cases that have more than two stations with blocking is more complex. In these cases, approximation methods can be utilized. One approximation method, introduced by Hillier and Boling[25], decomposes a network into smaller subsystems, analyzes each subsystem in isolation, and uses those results for analyzing the overall network. But application of their approximation method is restricted to tandem networks in which nodes are connected to each other in series. Takahashi et al. [26] use an open queuing network system with feed-forward flows and finite buffers between stations. They introduce an approximation method for open restricted queuing networks in which the service time and arrivals have exponential and Poisson distributions, respectively. The restriction in that network is the maximum queue length. They define pseudo-arrival and effective service rates in order to obtain an approximate

node-by-node decomposition. This idea is based on the inter-dependency nodes indicator: the blocking probability. Each node is treated as an M/M/1 system. They evaluate their approximation method with a triangle open restricted network and a four node tandem network. They compare the derived results of the model with that of a simulation and the exact computational results to demonstrate the validity of their method. The approximation method is capable of providing various performance measures such as blocking probabilities, output rates, etc., of general open restricted queuing networks. As their method is able to solve problems similar to the one studied here, we introduce it in more detail in next section. In a more recent study, Koizumi et al. [27] use a queuing network system with blocking and analyze congestion in a mental health system in Philadelphia. They use a multi-server model to present both mathematical and simulation results. Their system consists of three types of psychiatric institutions: extended acute hospitals (E), residential facilities (R), and supported housing (S). The accommodations outside their system are categorized into two groups: acute hospitals (A) and all other accommodations (X). Their network is illustrated in figure 5. The dotted lines represented a very small patient flow (less than 0.01 % of total patients who leaves S or R) and as a result were omitted in their final model.

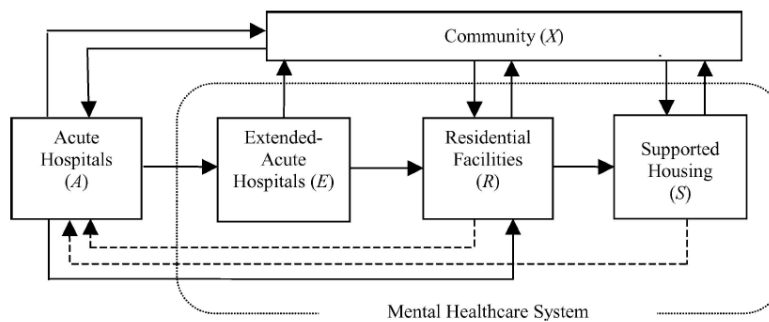


Figure 5- Inflows and outflows between stations in Koizumi et al.'s[27] model.

The authors assume a single First-Come-First-Served (FCFS) queue for all patients arriving at any given station. Their model can be written as a M/M/C/∞/FCFS queue which represents Markov arrivals and departures, a fixed number of servers (beds), infinite capacity and a FCFS queuing discipline. They use the same approximation methodology (single node decomposition) which Takahashi et al. [26] developed in their paper but extended to a multi server model. Two performance indicators are used in their model: number of patients

waiting to enter each facility of the system and the associated waiting time in steady state. In addition, simulation models are used for analyzing short term transient system performance and testing the robustness of the mathematical model. In their model, when a customer's service is finished at a given stage, the customer is moved to an infinite waiting room, if there is no idle server at the next desired stage; nevertheless, in reality the customer would be blocked and continue to stay at the given stage.

A recent study by Brettahauer et al. [20] presents a new heuristic method for tandem systems with blocking. They use a queuing network system to model patient flow between intensive care units (ICU), step-down (SD), acute care (AC), and post acute care units (PAC) in a large hospital located in the United States. In their study, ICU includes the medical/surgical ICU and coronary care units; SD includes the transitional ICU and telemetry units; and AC includes the surgical, medical, oncology, neurology, and transplant units. The patient flow between different stations is illustrated in figure 6. The solid lines in this figure are the primary patient flow while the dotted lines show limited patient flow.

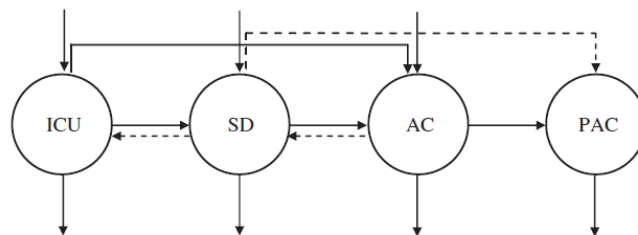


Figure 6- Inflows and outflows between stations in Brettahauer. 's et al. [28] model.

The aim of their study is to find the best mix of beds across the hospital's inpatients units in order to provide the highest quality of care possible given a limited budget. They develop a method for determining the optimal combination of beds at different units that minimizes a weighted sum of blocking probabilities for a given set of resources. In order to solve this optimization problem, the blocking probabilities in the queuing network are derived using a new approximation technique. They claim that this new technique overcomes the limiting assumption of Koizumi's et al. [27] method. To verify their method, an n-stage tandem network (for which the exact solution can be obtained) has been deployed and the results obtained from the new heuristic method are compared to Koizumi's et al. [27] method and the exact solution of the system. They use their verified method to solve the more complex

network shown in figure 5 for which an exact solution is unattainable. In their heuristic method, the number of servers and the service rate are adjusted at a stage to accommodate for potential blocking effects. Their method will be presented in detail in the next section.

In our study, a queuing network system with blocking will be used to model ALC patient flow through a system of service facilities. To our knowledge, there have been no applications of this kind of model to the analysis of congestion in the field of ALC patients. In the next section approximation methods for solving the open networks with blocking are explained.

2.4. Approximation Methods for Open Networks with Blocking

Approximation solution techniques can be used to analyze queuing networks when obtaining the exact solutions of the performance measures is impossible or inordinately expensive. There are several studies that deal with the approximation methods for queuing networks with blocking (Perros [29], Balsamo et al. [30]) and a few that have applied these methods to the health care system. Recalling from the previous section, we present two approximation methods that are introduced by Takahashi et al.[26] and Bretthauer et al. [28] to address a queuing network model in the health care system.

2.4.1. Node-by-Node Decomposition Method by Takahashi et al.

Takahashi et al. [26] present an approximation method for open queuing networks with blocking in which the service time and arrivals have exponential and Poisson distributions, respectively.

Takahashi et al. develop the idea of Pseudo-arrival rates and effective service rates to overcome the challenges arising from blocking. Blocking complicates the picture as it means that queues are no longer independent and thus cannot be treated separately and yet to solve them as a single unit is intractable due to the number of possible states to the system. We say that node i is blocked by node j if a customer, having completed service at node i , cannot proceed to node j because all servers at node j are currently occupied and the maximum queue size is reached. Suppose d_i denotes the rate of actual arrivals from other nodes to node i and λ_i^* denotes the rate of actual arrivals for non blocked time intervals. So $\lambda_i^* = d_i / (1 - \epsilon_i)$ where ϵ_i called the blocking probability, is the probability that node i blocks the

preceding nodes. λ_i^* is the pseudo-arrival rate and is used to approximate the arrival process from other nodes by a Poisson process with rate λ_i^* [26]. The effective service rate μ_i^* reflects the actual service time plus the expected time spent waiting at node i due to downstream blocking. Consider two nodes i and j where i precedes j . The mean service time at node i is equal to $1/\mu_i$ when node j is not full with customers and is equal $\frac{1}{\mu_i} + \frac{1}{\mu_i^*}$ when node j is full. A recurrence relation between successive nodes is obtained as follows:

$$\frac{1}{\mu_i^*} = \frac{1-\epsilon_j}{\mu_i} + \epsilon_j \left(\frac{1}{\mu_i} + \frac{1}{\mu_j^*} \right) = \frac{1}{\mu_i} + \frac{\epsilon_j}{\mu_j^*} \quad (6)$$

In an equilibrium state of the system an arrival rate to node j, d_j , is :

$$d_j = \sum_{i=1}^N \alpha_{ij} \cdot \mu_i^* \cdot (1 - P_i(0)) \quad (7)$$

Where α_{ij} is the probability that a customer goes from node i to node j , N is the number of nodes, and $P_i(0)$ is the probability that no customer is present at node i . We can define:

$\rho_j = (\lambda_j^* + \lambda_j)\mu_j^*$ where λ_j is the arrival rate to node j and ρ_j corresponds to a utilization factor of node j . The effective service rates satisfy the following relations:

$$\frac{1}{\mu_j^*} = \frac{1}{\mu_j} + \frac{\sum_k \alpha_{jk} \epsilon_k \left[1 + \frac{\sum_{l \neq k} \rho_l \alpha_{lk} \mu_l^*}{\mu_l^* + \mu_j} \right]}{\mu_k^*} \quad (8)$$

The probability that n customer are present at node j is $p_j(n) = \rho_j^n p_j(0)$ and $p_j(0) = (1 - \rho_j)/(1 - \rho_j^{S_j+1})$ where S_j is the maximum queue length of node j . The blocking probability of node j is $\epsilon_j = p_j(S_j)$. By solving these simultaneous equations for each node, characteristic quantities of the system can be obtained. The amount of calculation in this method relates to the number of nodes in the network linearly. So, for the large scale networks, the computational difficulties are not big issues. However, this method is useful, when for a specific node, only the effect from the first and the second succeeding nodes need to be considered.

2.4.2. Heuristic Method by Bretthauer et al.

Bretthauer et al.[28] introduce a new heuristic method that overcomes some limitations of the approximation method presented by Takahashi et al. and extended by Koizumi et al. Let $F_{i,j}$ denote the steady state customer flow rate from stage i to stage j in the tandem network with n stages in which nodes are connected to each other in series. With respect to constraints on the customer flow, we can write the following relations:

$$F_{i-1,i} = F_{i,i+1} \quad \text{for } i = 1 \dots n \quad (9)$$

$$F_{0,1} = \lambda(1 - \pi_1) \quad (10)$$

Where λ is the arrival rate, π_1 is the estimated probably that stage 1 is blocked, and $F_{0,1}$ denotes the steady-state external arrival rate to stage 1. The first constraint imposes flow balance at each stage and the second shows that the fraction of external arrivals turned away must equal the steady-state probability that all servers at the first stage are busy. To derive an approximate number of customers waiting for service at a stage, we assume that this potential waiting room is infinite. Based on Gross and Harris [31] the steady-state number of customers waiting for service at a stage is calculated as:

$$L(\lambda, \mu, s) = \frac{s^s \left(\frac{\lambda}{s\mu}\right)^{s+1}}{s! \left(1 - \frac{\lambda}{s\mu}\right)^2} \left[\sum_{n=0}^{s-1} \frac{\left(\frac{\lambda}{\mu}\right)^n}{n!} + \frac{\left(\frac{\lambda}{\mu}\right)^s}{s! \left(1 - \frac{\lambda}{s\mu}\right)} \right]^{-1} \quad (11)$$

where λ , μ , s are the arrival rate, service rate and number of servers, respectively. By using the above equation, we are able to estimate the number of servers blocked at stage i to be the number of customers waiting for service at stage $i+1$. Based on this estimate, we can adjust the characteristics of stage i by reducing the number of servers and calculating the number of effective servers as:

$$s_i^* = [s_i - L_{i+1}]^+ \quad (12)$$

and estimate the effective service rate at stage i as:

$$\mu_i^* = \left[\frac{s_i^*}{s_i} \frac{1}{\mu_i} + \frac{s_i - s_i^*}{s_i} \frac{1}{s_{i+1} \mu_{i+1}} \right]^{-1} \quad (13)$$

The algorithm for finding the blocking probability for each stage at a tandem network is as follows:

- a) Initialize $m = 0, \pi_1^0 = 0, \mu_i^0 = \mu_i, s_i^0 = s_i, \text{ for } i = 1, \dots, n.$

- Repeat (b) to (f) until $|\pi_1^m - \pi_1^{m-1}| < \delta$
- b) Increment m by one
 - c) Derive $F^m = \lambda(1 - \pi_1^{m-1})$
 - d) For $i = 1, \dots, n$ find the number of effective servers: $s_i^m = [s_i - L(F^m, \mu_{i+1}^{m-1}, s_{i+1}^{m-1})]^+$
 - e) For $i = 1, \dots, n$ find the effective service rates: $\mu_i^m = \left[\frac{s_i^m}{s_i} \frac{1}{\mu_i} + \frac{s_i - s_i^m}{s_i} \frac{1}{s_{i+1} \mu_{i+1}} \right]^{-1}$
 - f) Update the blocking probability as : $\pi_1^m = \pi^c(F^m, \mu_1^m, s_1^m)$ where

$$\pi^c(\lambda, \mu, s) = \frac{(\lambda/\mu)^s/s!}{\sum_{j=0}^s (\lambda/\mu)^j/j!} \quad (14)$$

This algorithm is extended to obtain the blocking probability for the general form of a queuing network with arbitrary routing (Bretthauer et al.[28]). In the general form of the network, customers can move from each node to any other node. Using the previous notation, let $v_{i,j}$ denote the fraction (probability) of customers at stage i that proceed to stage j after completion of service. Label “0” is used for anything outside of the system, such that $v_{0,i}$ shows the fraction of external arrivals to node i and $v_{i,0}$ denotes the fraction of customers that leave the network after completion of their services at node i . Without loss of generality and just for simplicity, we assume that $F_{i,i} = 0$; steady state flow rate from each node to itself is zero. For capturing the customers’ potential flows between nodes, arrivals from outside to each node and exit from any stage, equations (9) and (10) are adjusted as follows:

$$\sum_{k=0}^n F_{k,i} = \sum_{k=0}^n F_{i,k} \quad \text{for } i = 0, \dots, n \quad (15)$$

$$F_{0,i} = \lambda v_{0,i}(1 - \pi_i) \quad \text{for } i = 1, \dots, n \quad (16)$$

In the system with general routing, the ratios of internal flow rates are used to estimate the ratios of the steady state corresponding routing probabilities (Bretthauer et al.[28]):

$$\frac{F_{i,k}}{F_{i,j}} = \frac{v_{i,k}}{v_{i,j}} \quad \text{for } i = 1, \dots, n ; j, k = 0, \dots, n \quad (17)$$

Equation (17) is used for estimating the steady state flow rates between any two nodes after a customer is admitted to the system. However equation (17) is only used for interior stages as external arrivals might be turned away. In other words, when the first stages are full, the external arrivals cannot enter into the network and since there is no waiting room for them to stay, they have to be turned away. Equations (15), (16) and (17) solve the steady state flow

rates including stage “0”. For capturing the effects of blocking, the concept of effective system capacity is used by adjusting equations (11), (12) and (13) for the general routing system. As before, equation (11) estimates the average number of customers waiting for service at each stage. As a result, the number of customers waiting at stage i for service at stage j , $L_{i,j}$ is estimated as:

$$L_{i,j}(F, \mu, s) = \frac{F_{i,j}}{\sum_{k=0}^n F_{k,j}} L\left(\sum_{k=0}^n F_{k,j}, \mu_j, s_j\right) \quad (18)$$

Based on the estimated blocked servers at each stage, the effective number of servers at stage i and corresponding effective service rate is estimated by adjusting equations (12) and (13) as follows (Bretthauer et al.[28]):

$$s_i^*(F, \mu, s) = \left[s_i - \sum_{j=1}^n L_{i,j}(F, \mu, s) \right]^+ \quad (19)$$

$$\mu_i^*(F, \mu, s) = \left[\frac{s_i^*(F, \mu, s)}{s_i} \frac{1}{\mu} + \sum_{j=1}^n \left(\frac{L_{i,j}(F, \mu, s)}{s_i} \frac{1}{(F_{i,j}/\sum_{k=1}^n F_{k,j}) s_j \mu_j} \right) \right]^{-1} \quad (20)$$

The above estimated effective number of servers and effective service rates allow us to estimate the blocking probabilities at each stage. The blocking probability for stage i is estimated by (Bretthauer et al.[28]):

$$\pi_i(\lambda, \mu, s) = \frac{F_{0,i}}{\sum_{k=0}^n F_{k,i}} \pi^c(\lambda, \mu, s) + \frac{\sum_{k=1}^n F_{k,i}}{\sum_{k=0}^n F_{k,i}} \pi^{uc}(\lambda, \mu, s) \quad (21)$$

Where:

$$\pi^{uc}(\lambda, \mu, s) = \frac{\frac{s^s (\frac{\lambda}{s\mu})^s}{s! (\frac{\lambda}{s\mu})}}{1 - \frac{\lambda}{s\mu}} \left[\sum_{n=0}^{s-1} \frac{\left(\frac{\lambda}{\mu}\right)^n}{n!} + \frac{\frac{s^s (\frac{\lambda}{s\mu})^s}{s! (\frac{\lambda}{s\mu})}}{1 - \frac{\lambda}{s\mu}} \right]^{-1} \quad (22)$$

Where $\pi^c(\lambda, \mu, s)$ is estimated from equation (14). Using the new estimates for the blocking probabilities, a new solution for the flow-rate equations is derived. This algorithm can be repeated iteratively until convergence. Therefore, the new algorithm for finding the blocking probabilities in a general routing system is suggested as follows (Bretthauer et al.[28]):

- a) Initialize $m = 0, \pi_i^0 = 0, \mu_i^0 = \mu_i, s_i^0 = s_i, \text{ for } i = 1, \dots, n.$
Repeat (b) to (f) until $\sum_{i=1}^n |\pi_i^m - \pi_i^{m-1}| < \delta$
- b) Increment m by one
- c) Using π^{m-1} , solving equations (15), (16) and (17) to derive F^m

- d) For $i = 1, \dots, n$, using equation (19) to find the number of effective servers: $s_i^m = s_i^*(\mathbf{F}^m, \boldsymbol{\mu}^{m-1}, \mathbf{s}^{m-1})$
- e) For $i = 1, \dots, n$, using equation (20) to find the effective service rates: $\mu_i^m = \mu_i^*(\mathbf{F}^m, \boldsymbol{\mu}^{m-1}, \mathbf{s}^{m-1})$
- f) Update the blocking probabilities: $\pi_i^m = \pi_i(\mathbf{F}^m, \boldsymbol{\mu}^m, \mathbf{s}^m)$

In the next chapter the modified algorithm for modeling the CCN is introduced.

3. Research Methodology

3.1. The CCN as an Open Queuing Network

The CCN can be modelled with an open queuing network with blocking in which each node represents one possible destination facility for ALC patients. The structure of the network is defined through the current ALC discharge system in our simplified model. This structure will be modified in the future in order to more accurately model the real system.

In order to obtain the parameters of the model, two different approaches are considered. The first approach uses approximation methods to analyze the queuing network and to obtain the steady state performance measures. The main challenge with the approximation is to limit the error in the solution. Therefore, based on the literature, the accuracy of the approximation methods is tested through a simulation model (our second approach) to determine the conditions under which the algorithm yields a good approximation or at least a solution that mirrors the simulation results. Policies obtained from the final queuing network model should be evaluated in a system simulation model. Finally, the best policies are selected based on some performance metrics such as ALC waiting time, appropriate level of care, and cost.

3.2. Queuing Network Model

In order to address the issue of ALC patient access to the appropriate level of care, the approximation method for solving the queuing network model explained in the literature review is used. Considering that the CCN is an open queuing network with general routing, the suggested algorithm for estimating the blocking probabilities in a general system, introduced by Bretthauer et al.[28], is appropriate. However, for applying the aforementioned approximation method to the ALC network, some modifications are required. In the ALC network the large number of servers (beds) in each node results in computational challenges when attempting to implement the algorithm presented in Bretthauer et al.[28] algorithm. For instance, in using equation (11) for estimating the number of blocked customers at each stage, it is required to calculate s^s and $s!$ where s is the number of servers at each stage. Considering that in the CCN, there may be some nodes with a capacity of more than 1000 beds, it is almost impossible to calculate such terms with regular computers. To resolve this

issue, the equations used in the algorithm are converted into recursive functions thereby avoiding computationally intractable powers and factorials.

Starting from equation (14), the blocking probability function is rewritten as:

$$\frac{1}{\pi^c(\lambda, \mu, s)} = \frac{A}{B} \quad (23)$$

Where:

$$B = \frac{(\lambda/\mu)^s}{s!} \quad \text{and} \quad A = \sum_{j=0}^s (\lambda/\mu)^j / j! \quad (24)$$

Therefore:

$$\frac{1}{\pi^c(\lambda, \mu, s+1)} = \frac{A + \frac{(\lambda/\mu)^{s+1}}{(s+1)!}}{B \frac{\lambda}{\mu(s+1)}} = \frac{A}{B} \frac{\mu(s+1)}{\lambda} + \frac{(\lambda/\mu)^s}{s!} = \frac{A}{B} \frac{\mu(s+1)}{\lambda} + 1 \quad (25)$$

and we can obtain $\pi^c(\lambda, \mu, s+1)$ as a function of $\pi^c(\lambda, \mu, s)$ as follows:

$$\pi^c(\lambda, \mu, s+1) = \frac{\pi^c(\lambda, \mu, s)}{\pi^c(\lambda, \mu, s) + \frac{\mu(s+1)}{\lambda}} \quad (26)$$

The same procedure can be used for equation (22):

$$\frac{1}{\pi^{uc}(\lambda, \mu, s)} = KC + 1 \quad (27)$$

Where

$$K = \frac{1 - \frac{\lambda}{s\mu}}{\frac{s^s (\frac{\lambda}{s\mu})^s}{s!}} \quad \text{and} \quad C = \sum_{n=0}^{s-1} \frac{(\frac{\lambda}{\mu})^n}{n!} \quad (28)$$

Therefore:

$$\frac{1}{\pi^{uc}(\lambda, \mu, s+1)} = K' C + K' \frac{(\frac{\lambda}{\mu})^s}{s!} + 1 \quad (29)$$

Where

$$K' = \frac{1 - \frac{\lambda}{(s+1)\mu}}{\frac{(s+1)^{(s+1)} (\frac{\lambda}{(s+1)\mu})^{(s+1)}}{(s+1)!}} \quad (30)$$

Thus:

$$\frac{1}{\pi^{uc}(\lambda, \mu, s+1)} = \frac{K'}{K} \frac{1}{\pi^{uc}(\lambda, \mu, s)} - \frac{K'}{K} + 1 + K' \frac{\left(\frac{\lambda}{\mu}\right)^s}{s!} \quad (31)$$

Using equations (28) and (30) we obtain:

$$\frac{K'}{K} = \frac{s\mu}{\lambda} \left(1 + \frac{\mu}{s\mu - \lambda}\right) \quad \text{and} \quad K' \frac{\left(\frac{\lambda}{\mu}\right)^s}{s!} = \frac{s\mu + \mu - \lambda}{\lambda} \quad (32)$$

By putting equation (32) in (31) we obtain:

$$\frac{1}{\pi^{uc}(\lambda, \mu, s+1)} = \frac{s\mu}{\lambda} \left(1 + \frac{\mu}{s\mu - \lambda}\right) \left(\frac{1}{\pi^{uc}(\lambda, \mu, s)} - 1\right) + \frac{(s+1)\mu}{\lambda} \quad (33)$$

Finally, for converting equation (11) to a recursive formula, we rewrite it as:

$$\frac{1}{L(\lambda, \mu, s)} = WD + E \quad (34)$$

Where:

$$W = \frac{s! \left(1 - \frac{\lambda}{s\mu}\right)^2}{s^s \left(\frac{\lambda}{s\mu}\right)^{s+1}}, \quad D = \sum_{n=0}^{s-1} \frac{\left(\frac{\lambda}{\mu}\right)^n}{n!} \quad \text{and} \quad E = \frac{s\mu - \lambda}{\lambda} \quad (35)$$

Therefore:

$$\frac{1}{L(\lambda, \mu, s+1)} = W' \left(D + \frac{\left(\frac{\lambda}{\mu}\right)^s}{s!} \right) + \frac{(s+1)\mu - \lambda}{\lambda} \quad (36)$$

Where:

$$W' = \frac{(s+1)! \left(1 - \frac{\lambda}{(s+1)\mu}\right)^2}{(s+1)^{(s+1)} \left(\frac{\lambda}{(s+1)\mu}\right)^{s+2}} \quad (37)$$

Thus:

$$\frac{1}{L(\lambda, \mu, s+1)} = \frac{W'}{W} \frac{1}{L(\lambda, \mu, s)} - \frac{W' s\mu - \lambda}{W} + W' \frac{\left(\frac{\lambda}{\mu}\right)^s}{s!} + \frac{(s+1)\mu - \lambda}{\lambda} \quad (38)$$

Using equations (35), (37) and (38), we obtain:

$$\frac{1}{L(\lambda, \mu, s+1)} = \frac{s\mu}{\lambda} \left(\frac{(s+1)\mu - \lambda}{s\mu - \lambda}\right)^2 \frac{1}{L(\lambda, \mu, s)} - \frac{\mu}{\lambda} \left(\frac{(s+1)\mu - \lambda}{s\mu - \lambda}\right) \quad (39)$$

Equations (26), (33) and (39) are the computable recursive forms of equations (14), (22) and (11) respectively, and are used in the suggested algorithm for estimating the blocking probability of the different stages in the CCN.

3.3. Simulation

Simulation is an alternative approach for comparing the performance of the CCN under different scenarios. It is a general technique capable of incorporating many complexities of the systems and enabling us to see how the system works and how we can improve the performance measures by changing the systems' characteristics. In our study, the results of the simulation model and network queuing method are compared in order to check the accuracy of the approximation method.

4. Results

4.1. Data

The CCN parameters are obtained from three different datasets that have been provided by Champlain Community Care Access Centre (CCAC) and Champlain LHIN separately. Those datasets are analyzed and merged in order to find the required parameters for the queuing model. These parameters include average patients' flow rates among different facilities, mean length of stay at each node, total patients' arrival rate to the network and finally the capacity of each node. Since there is no unique identifier among the datasets, CCN's parameters have been derived based on some assumptions and simplifications. In the following, descriptive statistics of the available data as well as their trends in the time horizon of the study are provided.

The first dataset from the Champlain LHIN contains detailed information on ALC and LTC patients' routings. Specifically, the dataset contains 12,175 records for 10,374 ALC patients and 13,979 records for 9,267 LTC patients. After cleaning and linking ALC and LTC information, we were left with 16,940 ALC and LTC patients. Data include ALC patients who had been admitted and discharged from acute care in the period of April 2009 to April 2012. LTC patients were admitted to LTC facilities in the same period. 5,756 out of 9,267 patients were discharged (or exited) before April 2012. In table 2 summary of the final data on ALC and LTC patients are presented.

	# of patients	# of records
ALC	10,023	11,508
LTC-admissions	9,267	13,979
LTC-discharge	5,756	8,813

Table 2-Summary of available data on ALC and LTC patients

Table 3 presents the discharge rate of ALC patients from acute care to other facilities in each year.

	Apr-09 to Dec-09	Jan-10 to Dec-10	Jan-2011 to Mar-12
Acute care	2.9%	3.2%	2.6%
Ambulatory care	0.0%	0.1%	0.1%
Chronic care	9.6%	11.8%	12.5%
Rehab	16.2%	15.7%	20.6%
Home care	18.0%	14.7%	18.8%
Home for the aged	20.3%	18.8%	14.5%
Nursing home	14.1%	17.4%	13.6%
Psychiatric	0.3%	0.3%	0.4%
Unclassified/Other	0.5%	0.7%	1.3%
Exit (Deceased, Home without services,...)	18.0%	17.4%	15.6%
Total	100.0%	100.0%	100.0%
All # of discharged ALC patients	2206	3921	5381

Table 3- Discharge destinations of ALC patients- Apr. 2009- Mar 2012

The home care destination in table 3 includes assisted living and home care services. In addition, LTC facilities consisted of nursing home and home for the aged (table 2). According to table 2, the number of admitted ALC patients has grown dramatically from 2009 to 2012. The growth rate from 2009 to 2010 is 33.3% (by converting values for the year 2010 to nine month basis) and from 2010 to 2011 is 9.8% (by converting values for the year 2012 to one year basis). Although the time horizon is too short for a confident conclusion, it seems that the growth rate is decreasing. The main destination of ALC patients is LTC composing, on average, about 33% of ALC transfers from Apr 2009 to Mar 2012. About 3% of ALC patients moved to acute care either by being readmitted to the same or another hospital. In our study, we assume that there is no feedback from the same node; therefore, the patients who moved from acute care to acute care were eliminated from the dataset. In addition, the percentage of ALC patients transferred to ambulatory and psychiatric facilities are not substantial. Therefore, they are omitted from our dataset.

Figure 7 illustrates the acute care arrival probability distribution in the three year period from April 2009 to March 2012. The distribution has been fitted by the Poisson probability function very well (rank 1 among other possible probability functions), that shows the condition for the queuing model is met. Within this period, on average, ALC patients

admission to acute care per day is 10.549 ($\lambda = 10.549$). It worthwhile to note that the obtained arrival rate just shows ALC and LTC patient admissions to acute care from all other nodes (including external arrivals from outside of the network) and does not represent other types of patients' arrivals.

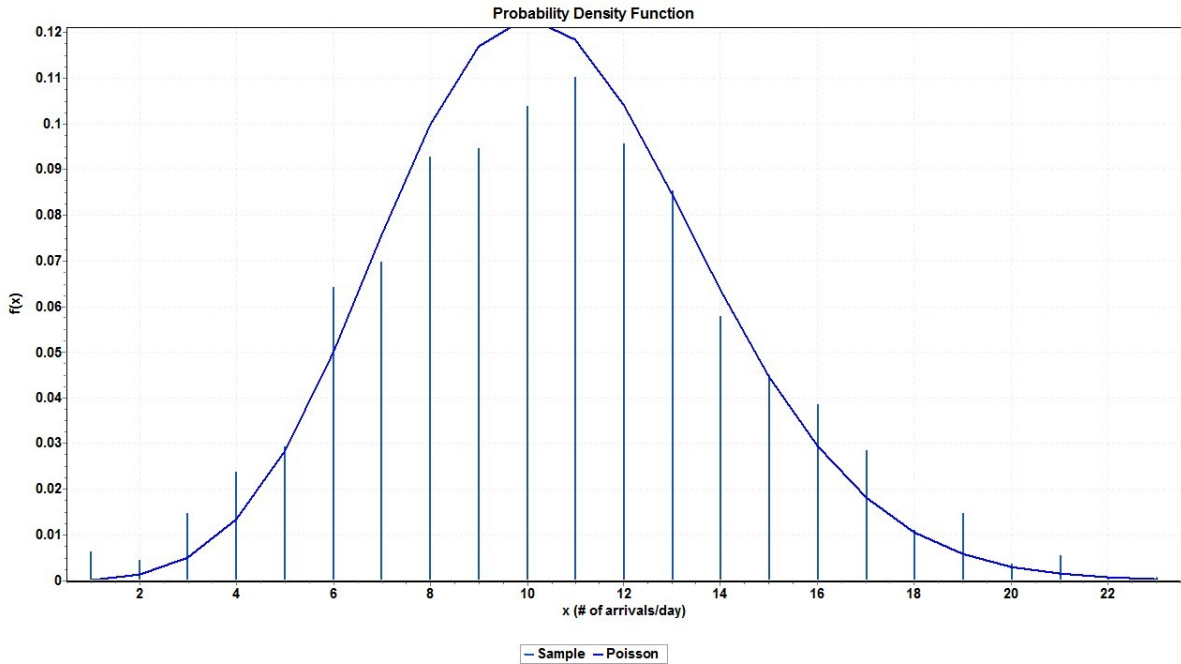


Figure 7– Acute care arrival's probability distribution

In addition, data shows that the average ALC days was 24.79 days with a standard deviation of 34.97 days and a maximum of 743 days. The average hospitalization days were 16.16 days with a standard deviation of 18.03 days and a maximum of 547 days. The probability distribution of length of stay in acute care for ALC patients is depicted in figure 8. The graph is fitted by exponential density function after removing the outliers from the original dataset. The well fitted exponential function satisfies the condition of the service time for the queuing model.

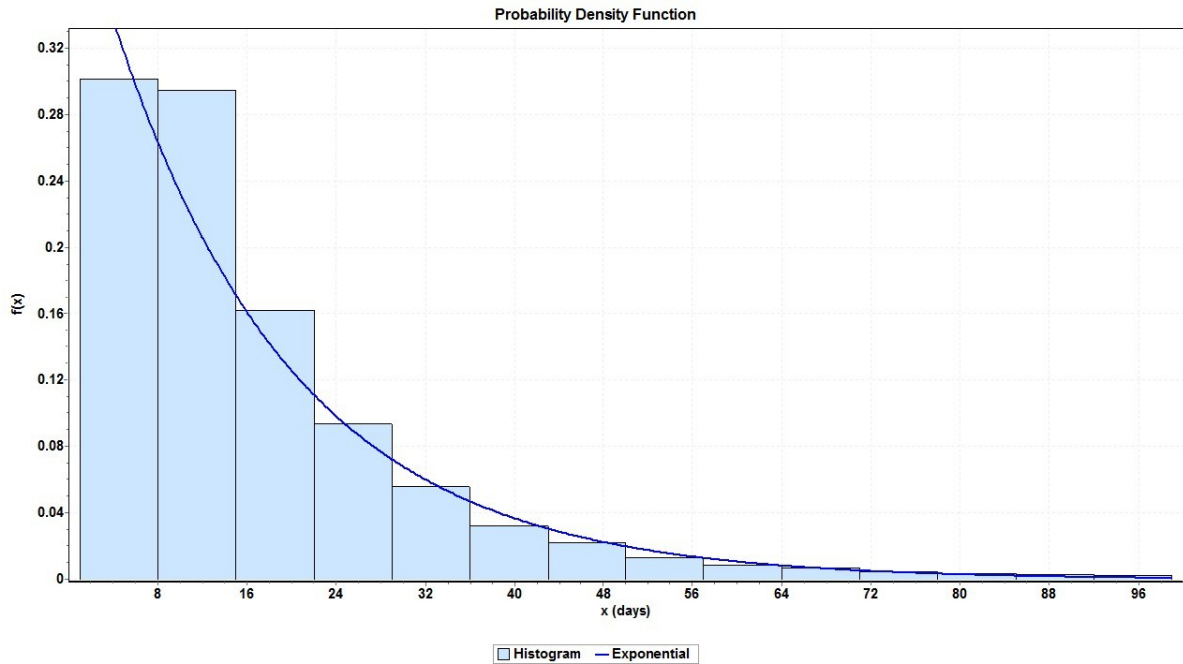


Figure 8 – Acute care length of stay’s probability distribution

The discharge rate of LTC patients to other facilities in each year are presented in table 4.

	Apr-09 to Dec-09	Jan-10 to Dec-10	Jan-2011 to Mar-12
Ambulatory	0.4%	0.5%	0.2%
Home care	4.2%	3.0%	3.2%
Acute care	37.9%	38.1%	42.7%
Continuing care	1.0%	0.8%	1.2%
Psychiatric	2.8%	2.2%	2.2%
Rehab	2.9%	4.3%	3.5%
Unclassified/Other	0.4%	0.6%	0.1%
Assisted living	8.5%	8.8%	6.9%
Exit	41.9%	41.8%	40.0%
Total	100.0%	100.0%	100.0%
All # of discharged LTC patients	1640	3240	3933

Table 4- Discharge destinations of LTC patients- Apr. 2009- Mar 2012

Data shows that there is a high fluctuation in the number of LTC patients during the years 2009 to 2012. The growth rate for year 2009 to 2010 is 48.2% (by converting values for the year 2010 to nine month basis) and for the year after is -2.9% (by converting values for the year 2012 to one year basis). By assuming that the number of LTC beds was kept constant during the mentioned period, the decreasing number of discharged patients would result in more congestion in LTC facilities. According to table 3, one of the main destinations of the LTC patients is acute care services. It shows that LTC patients move to hospitals and return to LTC during their stays in long term care with about a 40% probability.

The values reported in table 3 do not capture the effect of patients whose LOS was greater than 3 years. Based on information in table 1, 37% of LTC patients stayed more than 3 years in the LTC facilities which may affect the discharges probabilities in table 3 substantially but only if those patients with longer LOS exhibit different transfer distributions.

In figure 9, LTC patients' arrival is illustrated. As it is shown, arrivals are well fitted by the Poisson distribution function which satisfies our queuing model condition. The average arrival rate is 17.05 patients per day.

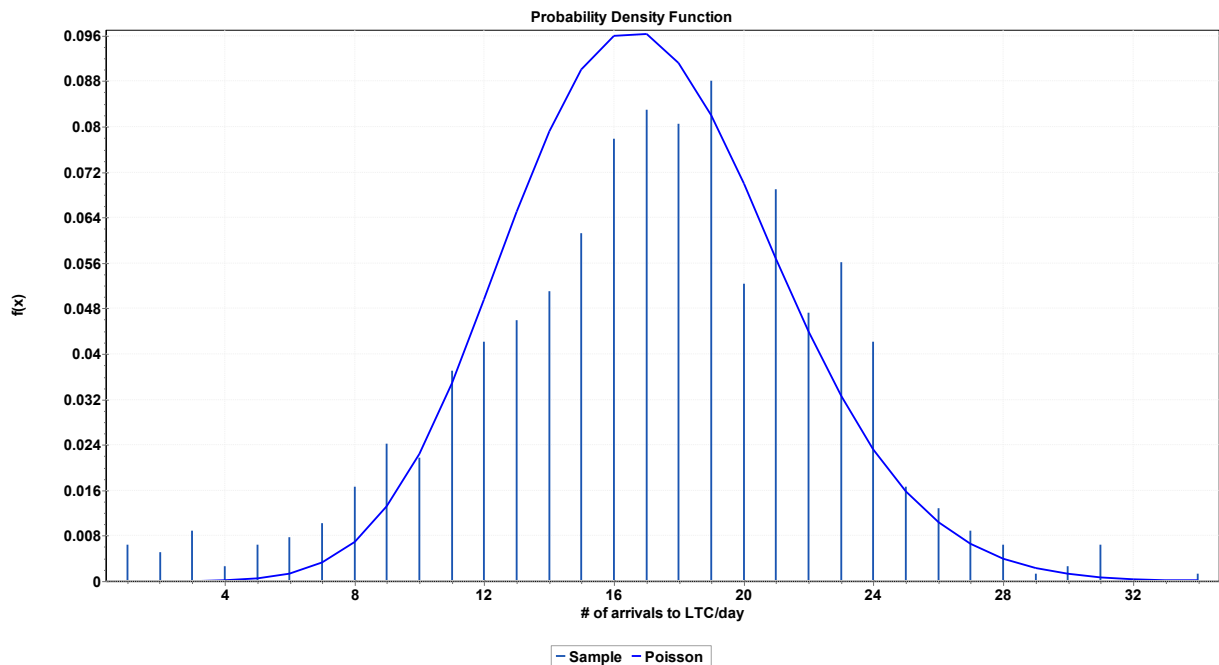


Figure 9- LTC arrivals

The other two datasets, provided by the Champlain Community Care Access Centre (CCAC), contain detailed information on CCAC patients' routing among the same aforementioned facilities. For consistency, the period of April 2009 to March 2012 has been selected as the research horizon. One dataset includes patient discharge dates to each facility and the other one contains information on LTC demand for the same patients. The former dataset contains 202,979 records for 74,586 CCAC patients and the later one contains information on 10,221 CCAC patients who were eligible for LTC at the time of their application. The later dataset also shows that 9,753 of the eligible patients were entered in the waiting list and 4,096 of them were finally discharged to LTC before October 2013. The others were withdrawn from the LTC waiting list for other reasons (death, ineligibility, client choice, ...) or still waiting for an available bed. The average waiting time for those discharged to LTC is 346 days. The histogram and the fitted exponential density function for the waiting days of CCAC patients before discharging to LTC is depicted in figure 10.

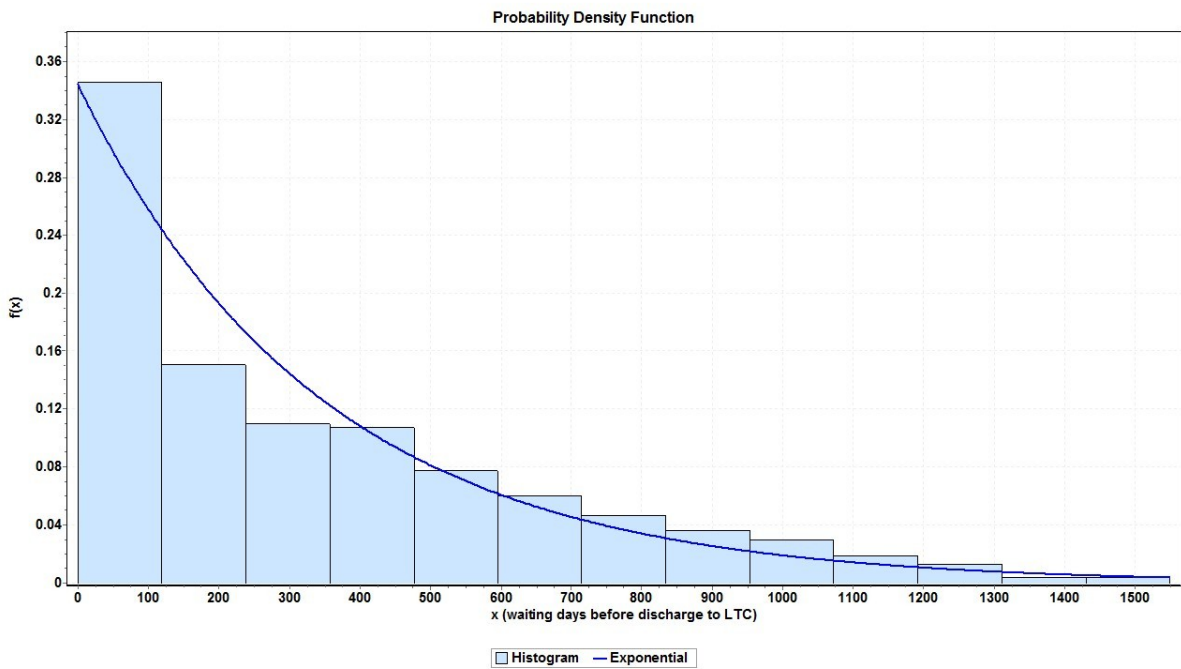


Figure 10–Probability distribution of CCAC patients' waiting days before discharging to LTC

Figure 11 shows the monthly number of additional CCAC patients who became eligible for LTC services. The average number of patients is 271 patients per month and it shows a slight decrease after June 2010, but it is almost steady during April 2009 till March 2012.

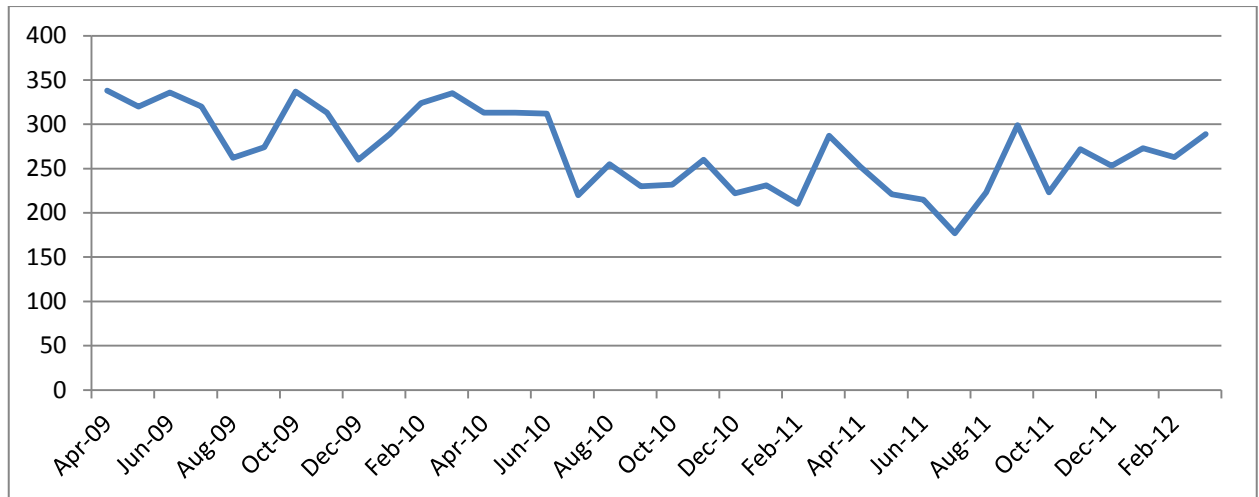


Figure 11- Number of patients added the LTC waiting list

Probability routing to each of the discharged destinations are calculated and the results are depicted in table 5. The table represents the percentage of patients who move from row *i* to column *j*. Since the percentage of departure from each of the stages to psychiatric care and rehabilitation are very low (less than 0.5%), both nodes have been eliminated from the CCAC network.

	Exit	Acute care	LTC	Chronic care	Home care	Assisted living	total
Enter	0%	0%	0%	0%	100%	0%	100%
Acute care	54.3%	0%	5.9%	2.5%	35.2%	1.6%	100%
LTC	26.2%	2.8%	0%	0.9%	69.9%	0.1%	100%
Chronic care	0.0%	1.4%	23.1%	0%	74.3%	1.2%	100%
Home care	77.9%	19.2%	8.0%	0.6%	0%	0.6%	100%
Assisted living	24.9%	30.3%	7.5%	1.7%	35.2%	0%	100%

Table 5- CCAC patients routing probabilities

The first row of table 5 represents the percentage of total arrivals from outside of the network to each of the other facilities. The only entrance of CCAC patients to the network is from home care, according to table 4. The column “exit” represents the percentage of patients who exited the network from each node.

Similar to acute care, we have found the probability distribution of CCAC patient arrivals to the network. Figure 12 shows the probability density function of patients who were admitted for CCAC regular services in the period of April 2009 to March 2012.

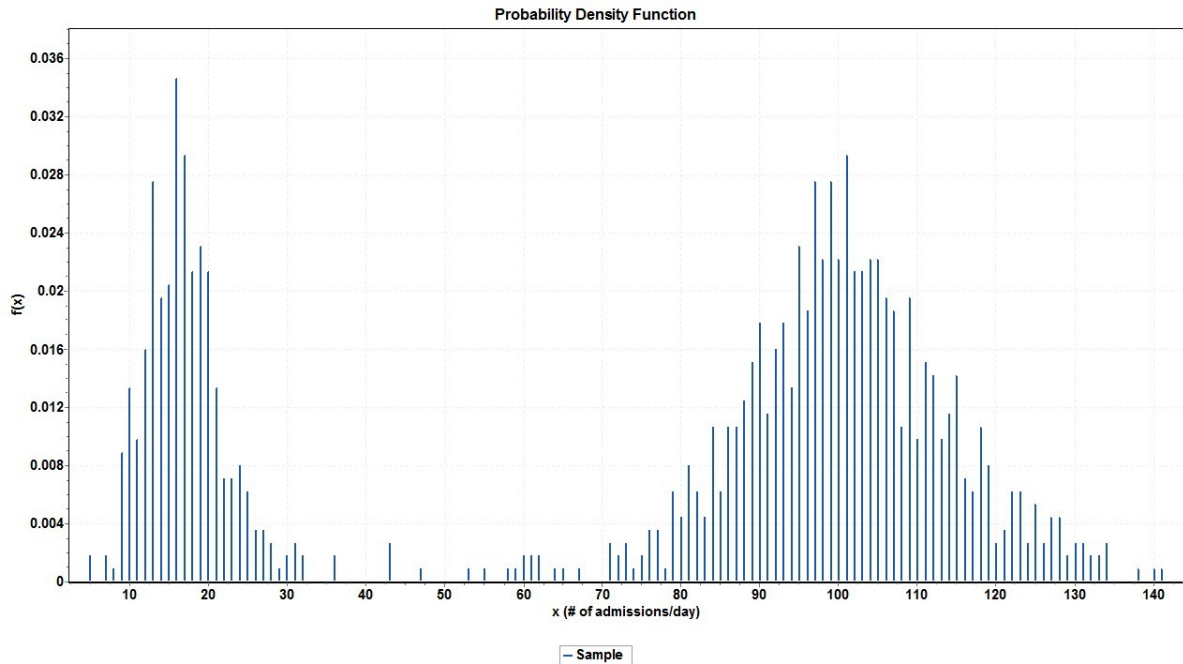


Figure 12– CCAC regular service admissions probability distribution

At the first glance, two separate distributions can easily be seen. With more investigation, we have figured out that there is a significant difference between the number of admissions on weekends and weekdays for CCAC regular service. This difference has been captured in figure 12. Therefore, arrivals are categorized based on the day of the week and two separate probability distribution graphs have been made for weekend or weekday arrivals, presented in figures 13 and 14, respectively.

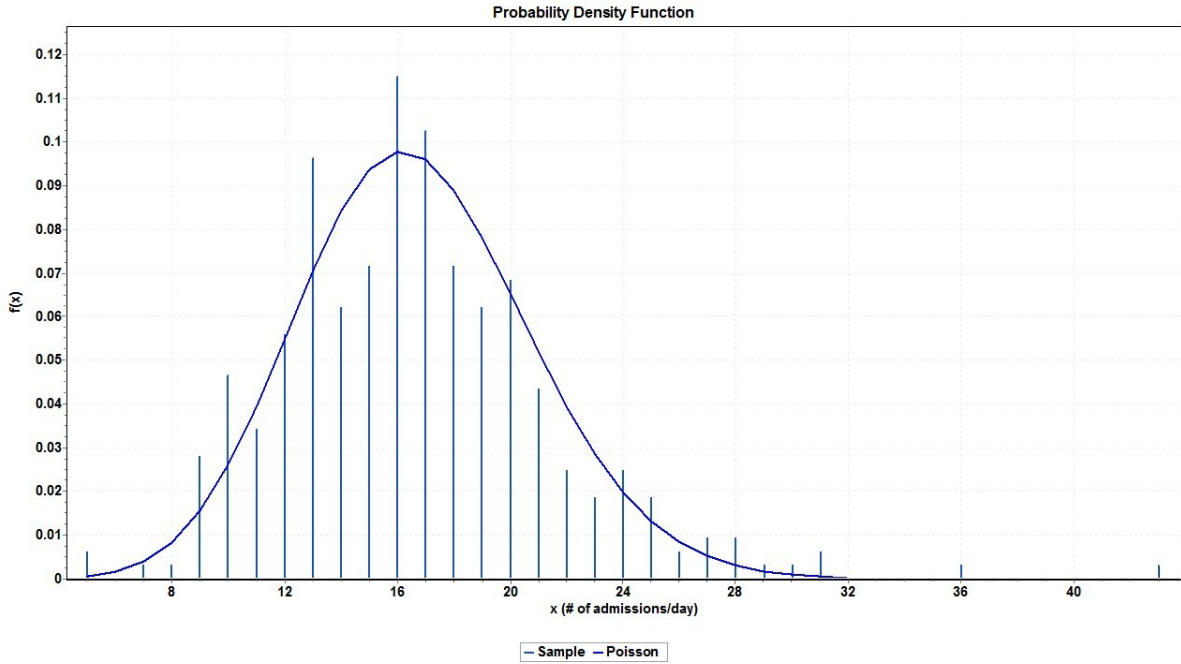


Figure 13– Weekends CCAC regular service admissions probability distribution

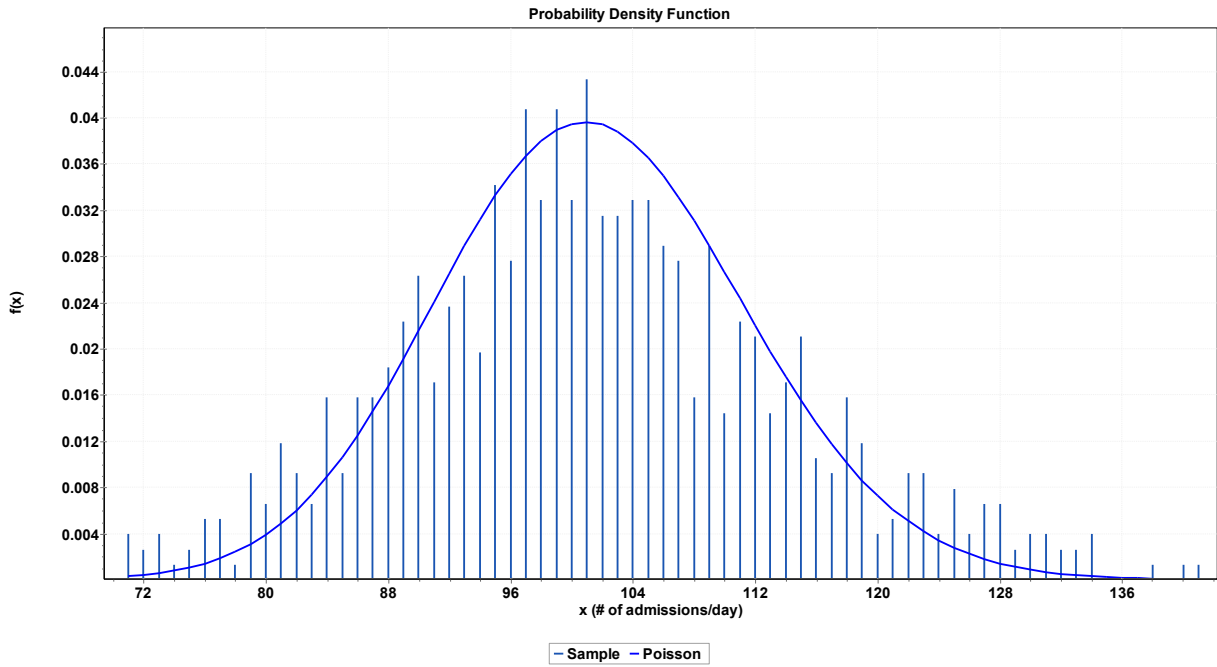


Figure 14– Weekdays CCAC regular service admissions probability distribution

Figures 13 and 14 are fitted by Poisson density functions properly and as a result the queuing model condition for arrivals is met.

In order to derive the parameters of the CCN, we had to merge three datasets. As already mentioned, there is no common identifier in the datasets and as a result, merging datasets was based on some assumptions. Consequently, the obtained parameters of the network are approximate. Since home care facility contains CCAC patients as well as other ALC and LTC patients, we used this node as the mutual node among the datasets for merging and deriving the CCN parameters. We also assumed that home care service and CCAC regular services are the same. Our data for ALC and LTC patients does not contain information on patients who were admitted directly to other nodes rather than acute care. As a result, the estimated demands for those facilities are underestimated. Even for ALC and LTC patients, because of the limitation of data, we could not track patients who moved among facilities rather than LTC and acute care. Therefore, the estimated probabilities have some deviations from the real values. We also assumed that when a patient moves to another facility, his bed is vacated, while in reality it is possible the bed is kept reserved for that patient. For example, when patients move to acute care for short term, their beds remain as occupied in their former facilities, but in our analysis the bed is considered as available for next client. Despite these assumptions and difficulties, experts at both the CCAC and the Champlain LHIN were comfortable with the usability of the model.

Table 6 and figure 15 show the routing probabilities of the CCN. As it shows, the two entrances of patients to the network are “Acute care” and “Home care” with the proportion of total demand arriving at each set at 15% and 85%, respectively. There are on average 50 patient arrivals per day. After entering the network, patients will move among the network’s nodes until they exit the network for any reason such as death or discharging to their homes without any services.

	Exit	Acute care	LTC	Chronic care	Rehab	Home care	Assisted living	total
Enter	0.0%	15.0%	0.0%	0.0%	0.0%	85.0%	0.0%	100%
Acute care	37.8%	0.0%	22.0%	5.0%	14.0%	14.0%	7.2%	100%
LTC	46.2%	45.0%	0.0%	2.0%	2.5%	3.3%	1.0%	100%
Chronic care	54.0%	23.0%	5.0%	0.0%	0.0%	0.0%	18.0%	100%
Rehab	5.0%	13.0%	4.0%	0.0%	0.0%	78.0%	0.0%	100%
Home care	87.0%	4.0%	8.0%	0.0%	0.0%	0.0%	1.0%	100%
Assisted living	15.0%	20.0%	65.0%	0.0%	0.0%	0.0%	0.0%	100%

Table 6- CCN patients routing probabilities

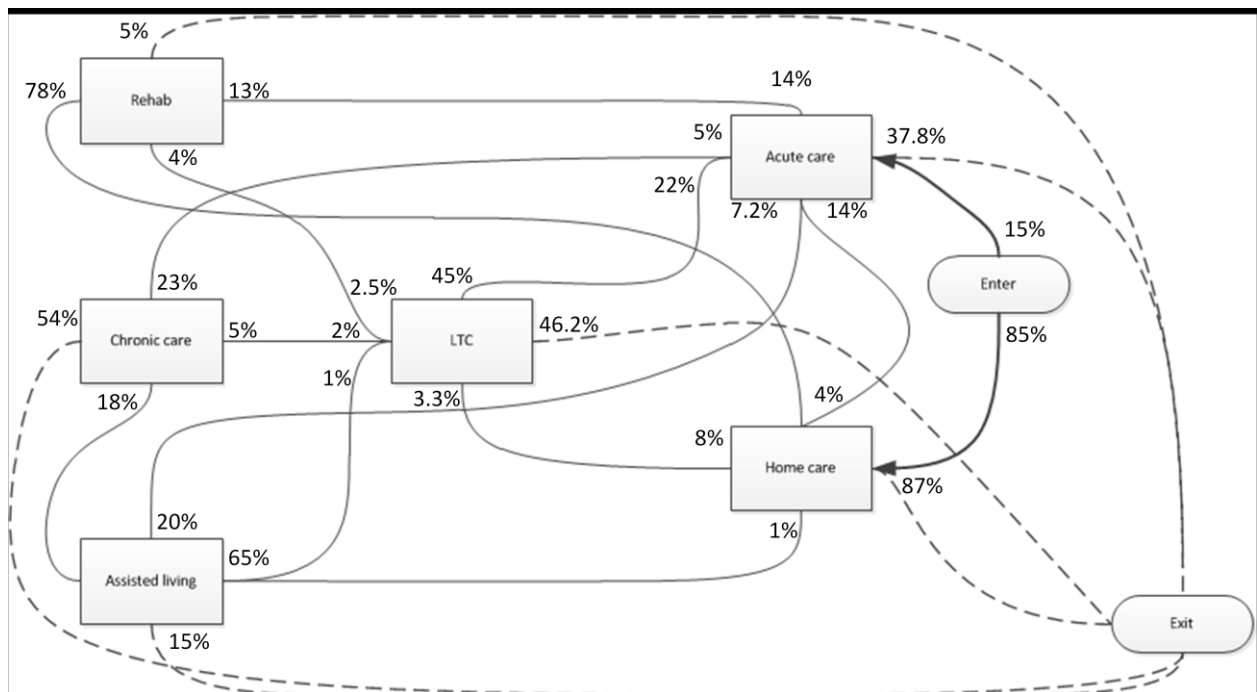


Figure 15– CCN and routing probabilities

In addition to the routing probabilities, the capacity of each facility and the average length of stay (LOS) at each node are the required parameters for the queuing model. In table 7, those parameters are presented. It should be considered that, some of the information on the LOS and capacity numbers was elicited from knowledge and experience of experts at the CCAC and the Champlain LHIN. More specifically, all LOS and capacity information was provided by specialists, except LOS for acute care, chronic care, and rehab that were derived from available data.

	LOS (day)	Capacity (bed)
Acute care	16.16	1500
LTC	1131.5	7570
Chronic care	77	471
Rehab	24.5	185
Home care	170	20000
Assisted living	300	540

Table 7- LOS and Capacity of the CCN nodes

4.2. Queuing Model Results

Using the modified heuristic approximation method, introduced in “Research Methodology”, the results of the queuing network model for the CCN are derived. The algorithm introduced in section 3.2 has been implemented in the statistical software R (64 bit- version 3.0.1). To use the algorithm it is required to satisfy the stability condition of each node. In other words, the arrival rate at each node should be less than $s\mu$. In order to make sure that this condition is satisfied, the arrival rate is truncated and replaced by $\lambda' = \text{Min}(\lambda, s\mu - \varepsilon)$. We set the $\varepsilon = 10^{-4}$ in the model. For the highly congested system, the number of patients estimated in equation 39 (patients blocking at stage i and waiting for stage j) may exceed the capacity of stage i which does not make sense in reality. To prevent such cases, in the algorithm, the total number of patients waiting at stage i is set to be at most equal to capacity of stage i . To speed up convergence, in the first 20 iterations we use an average of the past two blocking probabilities to update the current blocking probabilities in step f of the algorithm[28]. In other words, we set $\pi_i^m = (\pi_i^{m-1} + \pi_i(\mathbf{F}^m, \boldsymbol{\mu}^m, \mathbf{s}^m))/2$. After the first 20 iterations, the weights are reduced and the updated blocking probabilities are estimated with $\pi_i^m = ((m - 1)\pi_i^{m-1} + \pi_i(\mathbf{F}^m, \boldsymbol{\mu}^m, \mathbf{s}^m))/m$. Convergence criterion (δ) is set to 10^{-5} . We developed four scenarios that all converge after 10-30 iterations. In order to validate the developed queuing model, we implemented the parameters of the Bretthauer et al. [28] model in our model. The results of both models are presented in table 8. The comparisons show that the modified approximation method performs very well and the difference between the two models is negligible.

	Modified Heuristic Model				Bretthauer et al model			
Case 1	14.4	2.8	6.7	0.0	14.6	2.8	6.8	0.0
Case 2	14.4	2.8	6.7	0.4	14.6	2.8	6.8	0.3
Case 3	15.4	3.1	7.0	52.0	15.5	3.1	7.1	53.2

Table 8- Blocking probabilities (%) in different scenarios introduced by Bretthauer et al model and modified heuristic model

Numbers in table 8 are the blocking probabilities (%) of different cases introduced by Bretthauer et al. [28].

In the next subsection, the analysis of the different scenarios will be explained. The purpose of analysing different scenarios is twofold. On the one hand, we will be able to test the

robustness of our results with respect to the initial parameters. On the other hand, we consider different scenarios that might be used for further policy recommendations in the future.

4.2.1. Base Case Scenario:

Using the CCN parameters presented in tables 7 and 8, blocking probabilities and steady state queue sizes at each stage before moving to their downstream facility are estimated. The results are illustrated in tables 9 and 10.

Acute Care	Home Care	LTC	Rehab	Assisted Living	Chronic Care
0.0%	0.0%	99.8%	0.0%	31.4%	0.0%

Table 9- Blocking probabilities for CCN facilities- Queuing model –Base case scenario

	Acute Care	LTC	Chronic Care	Rehab	Home Care	Assisted Living	Total
Acute Care	0	349.4	0	0	0	3.0	352.4
LTC	0	0	0	0	0	0.2	0.2
Chronic Care	0	376.2	0	0	0	0.5	376.7
Rehab	0	7.8	0	0	0	0.5	8.3
Home Care	0	1694.1	0	0	0	1.4	1695.5
Assisted Living	0	371.6	0	0	0	0	371.6
Total	0	2799.1	0	0	0	5.6	2804.7

Table 10- Steady states number of blocked patients at each stage - Queuing model – Base case scenario

Table 9 shows that LTC is almost always full in the time horizon of the study causing long queues to form at other nodes as patients await placement. Another busy node is assisted living facility. Since LTC is almost full and a large proportion of assisted living patients discharge to LTC, the result for the blocking probability in assisted living node is reasonable (more than 30 percent). Based on the results, we see that patients can always find available beds in other nodes. However, it should be considered that the demand for rehab and chronic care is not fully captured in our model. So, the blocking probabilities for those nodes are underestimated.

Table 10 illustrates the steady state number of waiting patients at each node. The total number of patients that are blocked in the acute care shows the total number of ALC patients at steady state; 352 patients in average. Moreover, each column of the table shows the number of patients who are waiting for the services in the associated column facility. Around 2,800 patients are waiting to move to LTC. Most of them are receiving home care service while they wait. CCAC data verifies the obtained results about waiting lists.

For designing other scenarios, we focus on two parameters of interests: LTC capacity and patient arrival rate. By varying these parameters, while keeping all other parameters as in the base case scenario, performance metrics are obtained. Finally, a mixed scenario is developed. In this scenario capacity and demand for assisted living will be increased, resulting in a decrease in the average LOS of LTC patients.

4.2.2.Scenario 1: Increasing LTC capacity

As LTC has the most congested node among the other facilities in the network, one possible option for reducing the number of ALC patients in the CCN is increasing the capacity of LTC. Therefore, two different cases are considered to see the effects of increasing the capacity of LTC on the different metrics of the CCN. The first case is adding 1000 beds to LTC and increasing its capacity to 8570 beds. In the second case the capacity of LTC is set to 9070 beds which is 1500 beds more than base case parameter value. In tables 11 to 14, we illustrate the impact of these two capacity increases on the blocking probabilities and the number of patients waiting for different stages.

Case I:

Acute Care	Home Care	LTC	Rehab	Assisted Living	Chronic Care
0.0%	0.0%	99.8%	0.0%	30.9%	0.0%

Table 11- Blocking probabilities for CCN facilities- Queuing model -Scenario 1-Case I

	Acute Care	LTC	Chronic Care	Rehab	Home Care	Assisted Living	Total
Acute Care	0	348.3	0	0	0	3.0	351.3
LTC	0	0	0	0	0	0.2	0.2
Chronic Care	0	376	0	0	0	0.5	376.5
Rehab	0	7.8	0	0	0	0	7.8
Home Care	0	1688.9	0	0	0	1.4	1690.3
Assisted Living	0	370.5	0	0	0	0	370.5
Total	0	2791.5	0	0	0	5.1	2796.6

Table 12- Steady states number of blocked patients at each stage - Queuing model –Scenario 1 case I

Case II:

Acute Care	Home Care	LTC	Rehab	Assisted Living	Chronic Care
0.0%	0.0%	33.7%	0.0%	8.4%	0.0%

Table 13- Blocking probabilities for CCN facilities- Queuing model -Scenario 1-Case II

	Acute Care	LTC	Chronic Care	Rehab	Home Care	Assisted Living	Total
Acute Care	0	17	0	0	0	0.8	17.8
LTC	0	0	0	0	0	0.1	0.1
Chronic Care	0	0.2	0	0	0	0.1	0.3
Rehab	0	0.5	0	0	0	0	0.5
Home Care	0	20.9	0	0	0	0.4	21.3
Assisted Living	0	6.2	0	0	0	0	6.2
Total	0	44.8	0	0	0	1.4	46.2

Table 14- Steady states number of blocked patients at each stage - Queuing model –Scenario 1 case II

Not surprisingly, increasing capacity of LTC, the most congested node, can reduce the number of ALC patients as well as that of other waiting patients in the CCN. However, the effects of adding 1000 beds in LTC’s capacity to the number of blocked patients and blocking probabilities are very small (tables 11 and 12). In other words, it indicates that for reducing number of ALC patients significantly, there is a minimum number of additional beds need to be added to LTC capacity. What is perhaps more surprising is the dramatic impact of 1500 extra beds. Table 13 shows that adding 1500 beds to the current LTC

capacity, can resolve the ALC challenge by reducing number of ALC patients less than 20 patients. It should be mentioned that although this policy can be one of the possible solutions for the ALC challenge, it is not the optimal one. For finding the optimal policy, an optimization model should be developed. However, the results indicate that for solving the ALC challenge in the CCN, at most 1500 beds are required, assuming that all other parameter values are kept as in the base case. (One large concern is that the introduction of additional beds may cause demand to increase as wait times are known to act as a deterrent to demand). Although table 10 suggests that 2500 beds are necessary to admit all waiting patients, our analysis shows that at most 1500 beds can solve the problem due to the removal of additional time caused by blocking. The dramatic improvement with 1500 additional beds is due to the exponential nature of the relationship between wait times and utilization so that small decreases in average utilization can in fact have huge impacts on wait times. This fact might be considered in the future policy recommendations.

4.2.3.Scenario 2: Increasing demand

Another parameter of interest in the CCN is the patient arrival rate. As mentioned in the introduction, it is reasonable to assume that population aging impacts on patient flow. Therefore, one of the probable future challenges is patient flow growth through the health care system. As a result, by increasing the patient arrival rate in our model, blocking probabilities as well as the number of patients waiting for different facilities are obtained, keeping all other parameter values as in the base case. Tables15 and 16 illustrate the results of the model where the patient arrival rate has been increased by 5%.

Acute Care	Home Care	LTC	Rehab	Assisted Living	Chronic Care
0.0%	0.0%	99.8%	0	82.8%	0.0%

Table 15- Blocking probabilities for CCN facilities- Queuing model –Scenario 2

	Acute Care	LTC	Chronic Care	Rehab	Home Care	Assisted Living	Total
Acute Care	0	349.4	0	0	0	45.7	395.1
LTC	0	0	0	0	0	3.7	3.7
Chronic Care	0	376.2	0	0	0	7	383.2
Rehab	0	7.8	0	0	0	0	7.8
Home Care	0	1694.1	0	0	0	21.4	1715.5
Assisted Living	0	371.6	0	0	0	0	371.6
Total	0	2800	0	0	0	77.8	2876.9

Table 16- Steady states number of blocked patients at each stage- Queuing model –Scenario 2

The results show that increasing the patients’ arrival rate will affect the number of ALC patients as well as blocking probabilities in each node. Since the external arrivals will be turned away if the first stages (acute care and home care in our model) are full, the effect of increasing demand on performance of the network might be more influential in real cases. The main effect of the arrival rate growth is an increase in the assisted living blocking probability. Consequently, the number of ALC patients in the network increases to 395.1 patients on average (12.1% more than the base case scenario) and the total number of patients in the CCN who stay in their stages because of the lack of a vacant bed downstream increases to around 3000 patients on average (2.6% more than the base case scenario). The results show that CCN parameters are highly sensitive to the patients’ arrival rate. Considering that in the coming years patient flow is expected to increase, developing the appropriate strategies is critical.

4.2.4.Scenario 3: Mixed policy

The last developed scenario shows the effect of changing more than one parameter on the CCN performance. One possible scenario that can help to reduce the blocking probability in the LTC facilities is to switch the LTC demand to assisted living node, temporarily. After providing service for these patients (equal to average service rate of the SH), they will move to LTC. The direct result of this policy is a reduction in the demand for LTC from acute care and home care. The patients’ shorter stay at LTC (after being served in assisted living and then moving to LTC) causes decreased LOS in LTC. To implement such a policy, increasing the capacity of assisted living service is required. Since the average cost of assisted living

service per patient is less than a LTC bed, this policy might save a substantial amount of money for health care system.

Therefore, in this scenario the assisted living capacity is increased by 500 beds, and the discharge probability of acute care and home care services to LTC decreases by 10% and 0.5%, respectively and this demand rerouted to assisted living node results in increasing the discharged probabilities from acute care and home care to this node. Using the weighted average from each source of demand for LTC, these changes in the discharge probabilities will reduce the average LOS of LTC from 3.1 years to 2.75 years. Tables 17 and 18 report the obtained measures of the above scenario.

Acute Care	Home Care	LTC	Rehab	Assisted Living	Chronic Care
0.0%	0.0%	6.3%	0	6.65%	0.0%

Table 17- Blocking probabilities for CCN facilities- Queuing model –Scenario 3

	Acute Care	LTC	Chronic Care	Rehab	Home Care	Assisted Living	Total
Acute Care	0	0.8	0	0	0	1.0	1.8
LTC	0	0	0	0	0	0	0
Chronic Care	0	0	0	0	0	0.1	0.1
Rehab	0	0	0	0	0	0	0
Home Care	0	1.7	0	0	0	0.3	2
Assisted Living	0	1.0	0	0	0	0	1
Total	0	3.5	0	0	0	1.4	4.9

Table 18- Steady states number of blocked patients at each stage- Queuing model –Scenario 3

The results show that using this scenario might resolve the ALC challenge. By reducing the LOS of LTC patients, the blocking probability of LTC decreases dramatically which eliminates the long waiting patients for LTC services in other facilities. In other words, the patients’ movements among the different nodes are smoother in comparison with the base case scenario: more efficiency in the system in terms of waiting times and providing appropriate services to the patients.

4.3. Simulation Results

Since the results derived from the queuing network model rely on an approximation method, there might be significant approximation error. Therefore, we test the accuracy of the queuing model through a simulation model. The behaviours of the system in different scenarios are investigated in the simulation model. We used ARENA (version 13.9) to develop the simulation model. To obtain the simulation results, 100 replications for each scenario are generated. In each replication, the warm up time and run time are set to 600 days and 6000 days, respectively. To speed up the simulation process, all of the nodes are prefilled to a certain portion of the nodes' capacity.

4.3.1. Base Case Scenario:

Using the same parameter values in the queuing model, the blocking probabilities and number of patients who are blocked in each node are derived. Table 19 and 20 demonstrates the base case scenario results.

Acute Care	Home Care	LTC	Rehab	Assisted Living	Chronic Care
0	0	99.8%	0	1.7%	0

Table 19- Blocking probabilities for CCN facilities- Simulation model – Base case scenario

	Acute Care	LTC	Chronic Care	Rehab	Home Care	Assisted Living	Total
Acute Care	0	373.6	0	0	0	0.2	373.8
LTC	0	0	0	0	0	0	0
Chronic Care	0	0	0	0	0	0	0
Rehab	0	0	0	0	0	0	0
Home Care	0	2558.2	0	0	0	0	2558.2
Assisted Living	0	0.5	0	0	0	0	0.5
Total	0	2932.3	0	0	0	0.2	2932.5

Table 20- Average number of blocked patients at each stage - Simulation model- Base case scenario

The simulation model results show that the most congested node in the CCN is LTC which is almost always full, consistent with the queuing model results. The number of ALC patient

estimated in the simulation model is about 374 patients which is only 7% more than the queuing model result. But the number of patients who are blocked at home care, before moving to LTC, is around 2,560 patients which is 51% more than the result of the queuing model. It seems that the queuing model could capture all the blocking probabilities more precisely than the number of blocked patients at other nodes. One possible reason is that in the queuing model there is a restriction on the maximum number of blocked patients at each node that may cause increasing the approximation error of the model when network is highly congested.

In addition to the aforementioned performance metrics, waiting time at each node and utilization of each facility may give us a more comprehensive picture of the CCN behavior. Fortunately, the simulation model can provide us with such important information. Based on the results, waiting time for the blocked patients in acute care, before discharge to LTC, is 129.5 days. Similarly, waiting time for the ones who are blocked in home care is 693 days. Compared with the available datasets, the simulation model overestimates both waiting times. Using Little’s law in the queuing model shows that in the steady states, the average waiting time for the patients in the acute care is 25.7 days which means queuing model estimates this parameter more accurately than the simulation model (comparing with the calculated value from the dataset equal to 41 days). However, queuing model underestimates the waiting time for the ones who blocked in the home care (37 days compare to 346 days). Figure 16 shows the average utilization at each node.

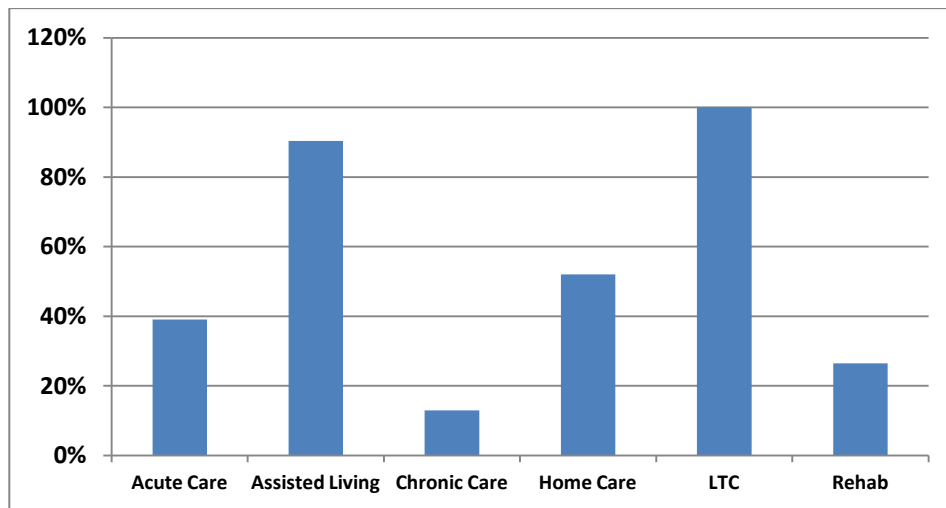


Figure 16- Average utilization of the CCN facilities- Base case scenario

The next busiest facility, after LTC, is assisted living with the average utilization around 90%. Assisted living is the facility that can play a buffer role between acute care and LTC. In other words, patients who need to discharge from acute care to LTC can first move to assisted living, when there is not an available bed in LTC, and afterwards proceed to LTC. Since the next destination of patients after assisted living stage is LTC and because LTC is always full in the current situation, assisted living utilization is also high.

4.3.2.Scenario 1: Increasing LTC capacity

As with the queuing model, the impact of increasing LTC capacity on CCN metrics is analyzed in two different cases. In the first case, the capacity of LTC is set to 8570 which is 1000 more than the base case scenario and in the second case it is set to 9070. Tables 21 to 24 demonstrate the blocking probabilities and the average number of blocked patients in each node for the two cases. Figures 17 and 18 also depict the average utilization of each facility in both cases.

Case I:

Acute Care	Home Care	LTC	Rehab	Assisted Living	Chronic Care
0.0%	0.0%	99.8%	0.0%	8.2%	0.0%

Table 21- Blocking probabilities for CCN facilities- Simulation model -Scenario1-case I

	Acute Care	LTC	Chronic Care	Rehab	Home Care	Assisted Living	Total
Acute Care	0	11.1	0	0	0	1.6	12.7
LTC	0	0	0	0	0	0	0
Chronic Care	0	0	0	0	0	0	0
Rehab	0	0	0	0	0	0	0
Home Care	0	1040	0	0	0	0.1	1040.1
Assisted Living	0	0.4	0	0	0	0	0.4
Total	0	1051.5	0	0	0	1.7	1053.2

Table 22- Average number of blocked patients at each stage - Simulation model- Scenario1-caseI

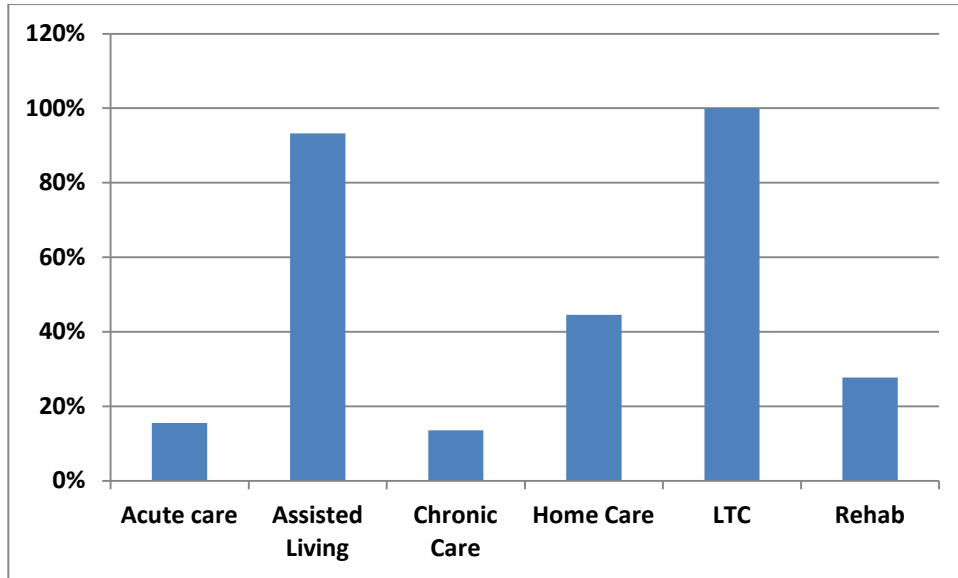


Figure 17- Average utilization of the CCN facilities- Scenario 1- Case I

Case II:

Acute Care	Home Care	LTC	Rehab	Assisted Living	Chronic Care
0.0%	0.0%	49.6%	0.0%	9.7%	0.0%

Table 23- Blocking probabilities for CCN facilities- Simulation model- Scenario 1 case II

	Acute Care	LTC	Chronic Care	Rehab	Home Care	Assisted Living	Total
Acute Care	0	3.1	0	0	0	1.4	4.5
LTC	0	0	0	0	0	0	0
Chronic Care	0	0	0	0	0	0	0
Rehab	0	0	0	0	0	0	0
Home Care	0	67.3	0	0	0	0.1	67.4
Assisted Living	0	0	0	0	0	0	0
Total	0	70.4	0	0	0	1.5	71.9

Table 24- Average number of blocked patients at each stage- Simulation model Scenario1-case II

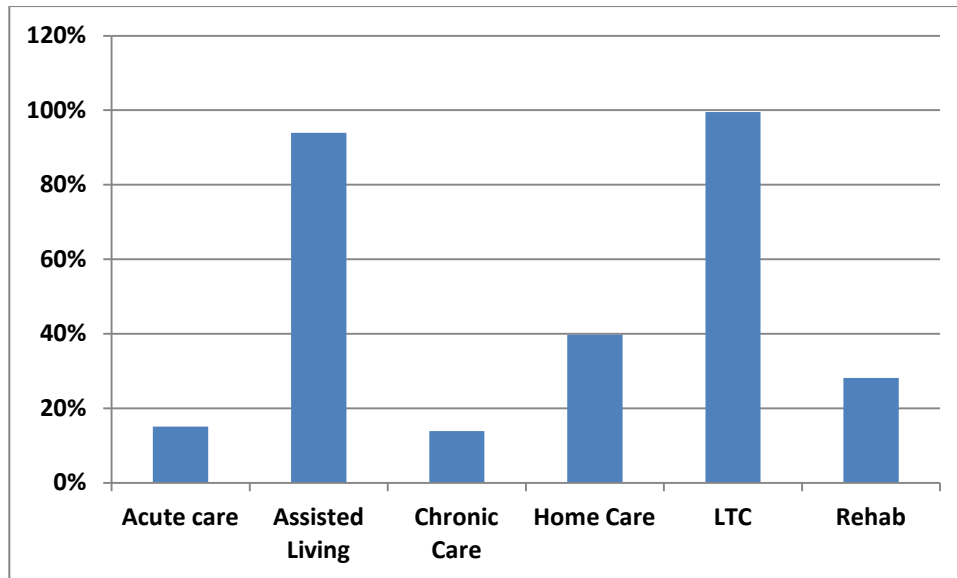


Figure 18- Average utilization of the CCN facilities- Scenario 1- Case II

Results show that increasing LTC capacity will reduce the number of ALC patients as well as the number of blocked patients at other nodes, dramatically. By adding 1000 and 1500 beds to LTC, the total number of blocked patients in the CCN will be reduced from 2933 patients in the base case scenario to 1053 and 72 patients respectively. As with the queuing model's results, the necessary increase in LTC beds is less than the current backlog might suggest.

In addition, waiting times for the patients in congested nodes are reduced in both cases of scenario 1 in comparison with the base case scenario. Patients in the acute care will wait only 3.7 days on average before discharging to LTC in case I and one day in case II. For the home care blocked patients, this waiting time is 281 days and 18 days in case I and case II, respectively.

The results show that there is not a substantial difference between two cases in terms of node utilization. In both cases, LTC and assisted living have the highest utilization percentage among all the facilities. Moreover, by comparing the results of the two cases with the base case scenario, it can be concluded that increasing the LTC capacity might reduce the acute care utilization significantly while it does not have a strong influence on the other nodes' utilization.

Comparison of the queuing model results with the simulation for case two also shows that the queuing model might capture the behaviour of the CCN in terms of blocking probabilities

and number of blocked patients quite well. In other words, it seems that while the network is not highly congested, the queuing model performs more properly than when the network contains nodes that are almost full.

4.3.3.Scenario 2: Increasing demand

In this scenario, the effect of increasing demand on different performance metrics of the CCN is analyzed. For consistency with the queuing model, scenario 2 simulates the CCN while the arrival rate is increased by 5% compared to the base case scenario. Obtained blocking probabilities, number of blocked patients at each node and utilization of facilities are illustrated in tables 25 and 26 and figure 19

Acute Care	Home Care	LTC	Rehab	Assisted Living	Chronic Care
1.5%	0.0%	99.9%	0.0%	9.6%	0.0%

Table 25- Blocking probabilities for CCN facilities- Simulation model- Scenario 2

	Acute Care	LTC	Chronic Care	Rehab	Home Care	Assisted Living	Total
Acute Care	0	713.5	0	0	0	1.7	715.2
LTC	0	0	0	0	0	0	0
Chronic Care	0	0	0	0	0	0	0
Rehab	0	0	0	0	0	0	0
Home Care	0	3122	0	0	0	0.1	3122.1
Assisted Living	0	0.6	0	0	0	0	0.6
Total	0	3836.1	0	0	0	1.8	3837.9

Table 26- Average number of blocked patients at each stage - Simulation model- Scenario 2

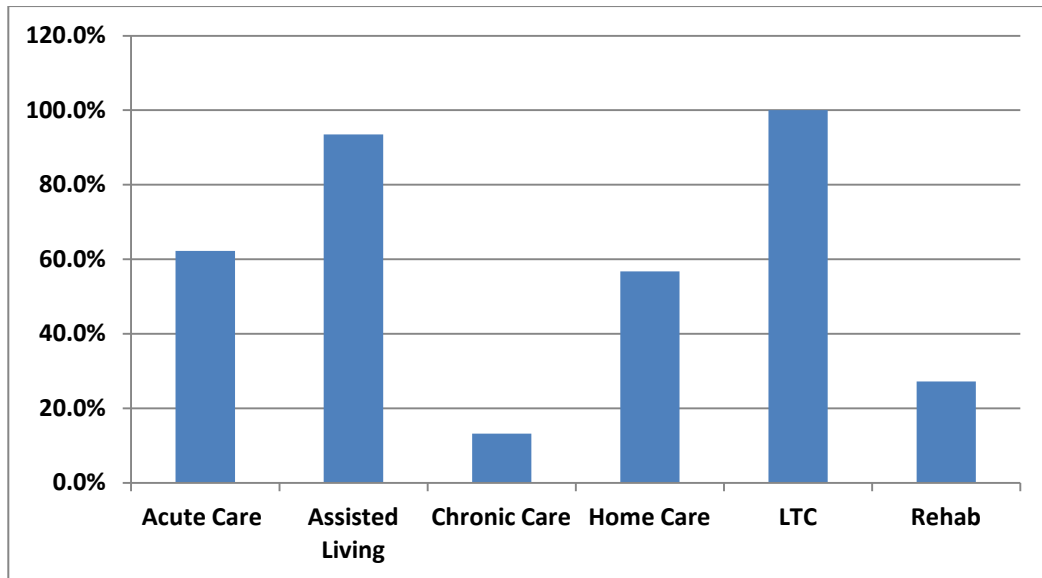


Figure 19-Average utilization of the CCN facilities- Scenario 2

Results show that increasing demand increases the blocking probability of the CCN facilities, results in increased number of ALC patients and utilization rates. The number of ALC patients is increased by 90% compared to the base case scenario. Since the demand of moving to LTC is high for home care patients, increasing the total arrival rate will result in a higher number of blocked patients in the home care services. Another impact of the increasing arrival rate is an increase in the utilization, specifically in acute care. The average utilization rate for acute care is estimated to be slightly more than 60% which is 59% more than the base case scenario. It should be noted that in our study ALC and LTC demand represent only a partial demand for acute care. Waiting time for acute care and home care patients before reaching to LTC are estimated 241 and 807 days, respectively.

4.3.4.Scenario 2: Mixed policy

As with the queuing model, this scenario increases the demand for assisted living while decreasing the demand for LTC. Using the same parameters as the queuing model and keeping the other values the same as base case scenario, the blocking probabilities and the number of patients waiting at each node as well as the utilization of each facility are presented in tables 27 and 28 and figure 20.

Acute Care	Home Care	LTC	Rehab	Assisted Living	Chronic Care
0.0%	0.0%	7.0%	0.0%	9.7%	0.0%

Table 27- Blocking probabilities for CCN facilities- Simulation model- Scenario 2

	Acute Care	LTC	Chronic Care	Rehab	Home Care	Assisted Living	Total
Acute Care	0	0.1	0	0	0	2.5	2.6
LTC	0	0	0	0	0	0	0
Chronic Care	0	0	0	0	0	0	0
Rehab	0	0	0	0	0	0	0
Home Care	0	3.2	0	0	0	0	3.2
Assisted Living	0	0.1	0	0	0	0	0.1
Total	0	3.4	0	0	0	2.5	5.9

Table 28- Average number of blocked patients at each stage - Simulation model- Scenario 2

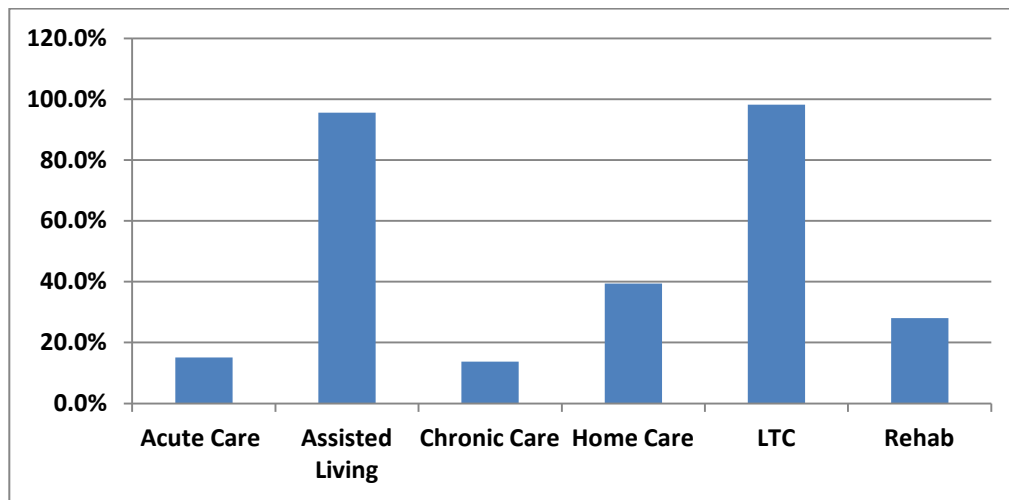


Figure 20-Average utilization of the CCN facilities- Scenario 3

The simulation results suggest that applying this scenario might overcome the ALC challenge as well as the long waiting times and waiting list to access LTC. The blocking probabilities for LTC and assisted living in this scenario are less than 10% percent, similar to the results of the queuing model (table 17). The queuing model works properly in the situations where the network is not highly congested. In addition, the number of patients

waiting at each facility because of a lack of capacity at other nodes is less than 10 patients. The queuing model is also giving us almost the same results. According to figure 20, utilization in LTC and assisted living services are the highest among all the facilities which is consistent in all the scenarios. The results also reveal that utilization in acute care service in the mixed scenario is less than base case scenario.

5. Limitations of the model

One should bear in mind that there are some limitations to the proposed method and therefore the results should be analyzed with caution. Some of the limitations are:

- We assumed that patient inter-arrival and service times follow exponential distributions. Although by using the available data we show that this assumption is valid for some facilities, it might be violated for other nodes.
- We assumed that the arrival rate and discharge probabilities remained constant during the time horizon of the study. Clearly, due to population aging this assumption is not true.
- We also assumed that the capacity of each facility is constant. Any development in the health care system might change this parameter of the model.
- There is no distinction among patients in our model. The service time for newly admitted patients are treated the same as the ones who have already got admission and returned for any other reason.
- It might be possible that a patient, who moves from one facility to another, still keeps the previous bed as reserved. For instance, patients in the LTC who require temporary hospitalization do not lose their LTC bed unless their stay in the hospital exceeds 30 days. In our model, if a patient leaves the facility, his bed is vacated.
- Obtained parameters of the network in our model are based on some simplifications and assumptions. As a result, accessing higher quality data, if possible, might improve our parameters estimates.
- In our model, we assumed that patients who enter the network must follow the predetermined discharge probabilities before exiting the system. In reality, a patient may choose to wait in the queue, or may leave, if the wait time gets too long (reneging).

6. Conclusions and Further Extensions

The increasing number of ALC patient mainly due to the aging phenomenon will be a serious challenge in the next few years. Nevertheless, this issue is understudied and there is a lack of information to inform policy makers regarding future strategies. The present research aims to contribute to filling this gap by proposing a queuing network and simulation model for modelling patient flow through the community care network and finding capacity allocation policies for the different facilities that resolves the ALC challenge using scenario analysis.

To address this issue, we first presented the heuristic approximation method for estimating the blocking probabilities in the queuing network model. This method is mainly based on the previous study done by Bretthauer et al.[28]. We applied the method to the CCN and found the blocking probabilities at each facility as well as the steady states number of patients who were blocked at each node. Then, we used a simulation model to test the accuracy of the queuing model. Finally, by developing different scenarios we checked the robustness of our model and determined a capacity allocation policy that might resolve the ALC challenge.

Moreover, comparison of the results from the queuing and simulation models reveals that the heuristic approximation method performs quite well, especially when there is not a highly congested node in the network. Our results also indicate that increasing total demand to the CCN may result in a large increase in the number of ALC patients in hospitals.

Scenario analyses in this research indicate that there might be different ways to improve the efficiency of the system. One obvious way is to increase the capacity of LTC facilities in order to reduce the blocking probability at this node and also to decrease the waiting time before moving to this stage. Two analyzed cases in this scenario show that adding 1500 beds to the LTC node will resolve the ALC issue in the network. Moreover, for the reduction of the number of blocked patients under the predetermined threshold, it is not necessary to increase the capacities by the exact number of patients currently on the wait list.

Another solution for resolving the ALC challenge is to use the assisted living service as an intermediate stage for LTC patients in order to reduce the LOS of LTC patients. Since the cost of assisted living service is less than LTC, this solution is more cost effective than the previous one.

Following this study several research topics seem to be worth pursuing. Obtaining the optimal capacity allocations for resolving the ALC challenges in the health care system

needs more applicable and realistic scenarios in addition to sensitivity analysis and optimization programming that will be suggested in further studies.

The network queuing model can be extended by adding a renegeing option to it. For that reason, the renegeing function should be added to the model to adjust the number of waiting patient in each node.

Total patient arrivals to the network can be predicted using demographic data as well as health care information. This may give us better picture of the future demands.

Appendix I: ALC Discharge Destinations' Description³

Complex Continuing Care (CCC) Bed - A designated bed providing specialized care to patients who are medically complex, require hospital stays, regular onsite physician care and assessment, and active management over extended periods of time.

Convalescent Care Bed - Provision of care to support the gradual recovery of health and strength after illness or surgery. Convalescent Care programs provide 24-hour care to people who require specific medical and therapeutic services in supportive environments for defined periods of time.

Home First- Private residence where a patient will live in the community upon discharge from hospital. Provision of an array of services that enables clients to live at home, often with the effect of preventing, delaying, or substituting for long-term care or acute care alternatives.

Long Term Care (LTC) Bed - A designated bed providing care to meet both the medical and non-medical needs of people with chronic illnesses or disabilities who require care that is not available in the community.

Mental Health Bed - A designated bed providing therapeutic services to patients with addictions, psychological, behavioural or emotional illnesses.

Palliative Care Bed - Provision of medical or comfort care to support end-of-life planning to reduce the severity of a disease or slow its progress. The focus is on quality of life measures rather than providing a cure.

Rehabilitation Bed - A designated bed providing care aimed at maximizing patients' overall physical, sensory, intellectual, psychological and social functions. This may include the acquisition of special equipment or other resources.

Supervised or Assisted Living - Provision of care for patients (e.g., the elderly or people with physical disabilities) who are able to mobilize independently but who may require assistance with activities of daily living.

³ - Descriptions adopted from cancer care Ontario's Data book 2012-2013, Appendix 2C.14

Appendix II: Abbreviations

AC : Acute Care

ALC : Alternative Level of Care

ALS-HRS: Assisted living services for high risk seniors

CCAC : Community Care Access Center

CCN : Community Care Network

DES : Discrete Event Simulation

FCFS: First-Come-First-Served

HCC : Home and Community Care

LOS : Length of Stay

LTC : Long Term Care

MDP : Markov Decision Process

MIP : Mixed Integer Programming

PAC : Post Acute Care

SH : Supportive Housing

References

- [1] Canadian Institute For Health Information, “Health Care in Canada, 2011, A Focus on Seniors and Aging,” 2011.
- [2] Ontario Association of Community Care Access Centres (OACCAC), Ontario Association of Non-Profit Homes and Services for Seniors, Ontario Hospital Association, Ontario Long-Term Care Association, “Alternate Level of Care -Challenges and Opportunities: A Collaborative Position Paper,” Toronto: Ontario Hospital Association, May 2006.
- [3] Ministry of Health and Long-Term Care, “Discharge of Hospital Patients,” Chapter 3, Section 3.02.
- [4] Canadian Institute For Health Information, “Analysis in Brief: Alternate Level of Care in Canada,” Jan. 2009.
- [5] D. Walker, “Caring For Our Aging Population and Addressing Alternate Level of Care,” Ministry of Health and Long-Term Care, Jun. 2011.
- [6] A. P. Costa and J. P. Hirdes, “Clinical Characteristics and Service Needs of Alternate-Level-of-Care Patients Waiting for Long-Term Care in Ontario Hospitals,” *Health care policy*, vol. 6, no. 1, pp. 32–46, Aug. 2010.
- [7] Ministry of Health and Long-Term Care, “Alternate Level of Care Patients,” Ministry of Health and Long-Term Care, Fact Sheet, May 2008.
- [8] J. B. Jun, S. H. Jacobson, and J. R. Swisher, “Application of discrete-event simulation in health care clinics: A survey,” *Journal of the Operational Research Society*, vol. 50, no. 2, pp. 109–123, Feb. 1999.
- [9] S. H. Jacobson, S. N. Hall, and J. R. Swisher, “Discrete-Event Simulation of Health Care Systems,” in *Patient Flow: Reducing Delay in Healthcare Delivery*, vol. 91, R. W. Hall, Ed. Boston, MA: Springer US, pp. 211–252.
- [10] J. Beliën and E. Demeulemeester, “Building cyclic master surgery schedules with leveled resulting bed occupancy,” *European Journal of Operational Research*, vol. 176, no. 2, pp. 1185–1204, Jan. 2007.
- [11] V. S. Chow, M. L. Puterman, N. Salehirad, W. Huang, and D. Atkins, “Reducing Surgical Ward Congestion Through Improved Surgical Scheduling and Uncapacitated

- Simulation,” *Production and Operations Management*, vol. 20, no. 3, pp. 418–430, May 2011.
- [12] S. Thompson, M. Nunez, R. Garfinkel, and M. D. Dean, “OR Practice--Efficient Short-Term Allocation and Reallocation of Patients to Floors of a Hospital During Demand Surges,” *Operations Research*, vol. 57, no. 2, pp. 261–273, Feb. 2009.
- [13] W. L. Hare, A. Alimadad, H. Dodd, R. Ferguson, and A. Rutherford, “A deterministic model of home and community care client counts in British Columbia,” *Health Care Management Science*, vol. 12, no. 1, pp. 80–98, Sep. 2008.
- [14] P. Intrevado, V. Verter, and L. Tremblay, “A Patient-Centric Model for the Design of Long-Term Care Networks,” Nov. 2013.
- [15] T. Cardoso, M. D. Oliveira, A. Barbosa-Póvoa, and S. Nickel, “Modeling the demand for long-term care services under uncertain information,” *Health Care Management Science*, vol. 15, no. 4, pp. 385–412, Jul. 2012.
- [16] F. Lin, N. Kong, and M. Lawley, “Capacity Planning for Publicly Funded Community Based Long-Term Care Services,” in *Community-Based Operations Research*, vol. 167, M. P. Johnson, Ed. New York, NY: Springer New York, 2012, pp. 297–315.
- [17] Y. Zhang, M. L. Puterman, M. Nelson, and D. Atkins, “A Simulation Optimization Approach to Long-Term Care Capacity Planning,” *Operations Research*, vol. 60, no. 2, pp. 249–261, Apr. 2012.
- [18] Y. Zhang and M. L. Puterman, “Developing an adaptive policy for long-term care capacity planning,” *Health Care Management Science*, vol. 16, no. 3, pp. 271–279, Mar. 2013.
- [19] J. Patrick, “Access to Long-Term Care: The True Cause of Hospital Congestion?,” *Production and Operations Management*, vol. 20, no. 3, pp. 347–358, May 2011.
- [20] H. Chen and D. D. Yao, *Fundamentals of queuing networks : performance, asymptotics, and optimization*. New York: Springer, 2001.
- [21] U. N. Bhat, *An introduction to queuing theory modeling and analysis in applications*. Boston, Mass.; London: Birkhäuser ; Springer [distributor], 2008.
- [22] J. R. Jackson, “Networks of Waiting Lines,” *Operations Research*, vol. 5, no. 4, pp. 518–521, Aug. 1957.

- [23] L. Green and V. Nguyen, "Strategies for cutting hospital beds: The impact on patient service," *Health Services Research*, vol. 36, no. 2, pp. 421–442, 2000.
- [24] E. N. Weiss and J. O. McClain, "Administrative days in acute care facilities: a queueing-analytic approach," *Oper Res*, vol. 35, no. 1, pp. 35–44, Feb. 1987.
- [25] F. S. Hillier and R. W. Boling, "Finite Queues in Series with Exponential or Erlang Service Times--A Numerical Approach," *Operations Research*, vol. 15, no. 2, pp. 286–303, Mar. 1967.
- [26] Y. Takahashi, H. Miyahara, and T. Hasegawa, "An Approximation Method for Open Restricted Queueing Networks," *Operations Research*, vol. 28, no. 3-Part-I, pp. 594–602, May 1980.
- [27] N. Koizumi, E. Kuno, and T. E. Smith, "Modeling Patient Flows Using a Queueing Network with Blocking," *Health Care Management Science*, vol. 8, no. 1, pp. 49–60, Feb. 2005.
- [28] K. M. Bretthauer, H. S. Heese, H. Pun, and E. Coe, "Blocking in Healthcare Operations: A New Heuristic and an Application," *Production and Operations Management*, vol. 20, no. 3, pp. 375–391, May 2011.
- [29] H. G. Perros, *Queueing networks with blocking: exact and approximate solutions*. New York: Oxford University Press, 1994.
- [30] S. Balsamo, V. De Nitto Personé, and R. O. Onvural, *Analysis of queueing networks with blocking*. Boston: Kluwer Academic Publishers, 2001.
- [31] D. Gross and C. M. Harris, *Fundamentals of queueing theory*. New York: Wiley, 1998.