Vol. 00, No. 0, Xxxxx 0000, pp. 000–000
ISSN 0000-0000 | EISSN 0000-0000 | 00 | 0000 | 0001

INFORMS
DOI 10.1287/xxxx.0000.0000

Dynamic Capacity Allocation for Long-Term Care

The allocation of long-term care capacity (beds) is a complex problem mainly because of the non-homogeneity of long-term care facilities and the presence of client choice. Additionally, there is the need to keep wait times for community demand short while also controlling the census of patients waiting in the hospitals. We describe a Markov decision process (MDP) model whose approximate solution provides reasonable capacity allocation policies for long-term care placements for both hospital and community demand.

Key words: Long-term care; Alternative level of care; Bed blockers; Capacity planning; Markov decision processes; Approximate dynamic programming; Linear Programming
History: Last review July 22, 2019

1. Introduction

In every hospital, bed management issues in units like the emergency department (ED) or the Intensive Care Unit (ICU) are very critical if they want to take care of the patients without problems and the soonest possible. And is not that uncommon that the capacity of these units is limited, which worsens the criticality of the situation. In fact, the overcrowding of these units has come to the attention of many researchers. As ED tend to be saturated, it is common that patients had to wait for their attention in the health sector.

Media coverage of hospital congestion most often focuses on the wait times experienced by patients in the ED. However, if one talks to those working in the emergency department, most will say that the reason for congestion in the ED is not internal inefficiencies but the inability to move admitted patients out of the ED and into the wards. If one investigates further one often hears the complaint on the wards that a major contributing factor to ward congestion is the inability to move those patients that are ready to be discharged, but must remain at the hospital until a bed in a geriatric institution or in a long-term care (LTC) facility becomes available. Such patients are designated as “bed-blockers” or “alternative level of care” (ALC) patients. Particularly in Canada, they are so mainly because of lack of beds and long wait lists for alternate care locations (Rubin and Davies, 1975; Styrborn and Thorslund, 1993; Rose et al., 2015).

In general, bed-blockers are mainly a group of elderly patients with varying medical, functional and psychosocial requirements(Styrborn and Thorslund, 1993). Growing population, increasing morbidity, and scientific and technological advances prolonging life, with the growing burden of chronic disease makes that more and more patients are requiring for prolonged mechanical ventilation, making ICUs less capable of delivering an adequate level of health (Rose et al., 2015). Many developed countries are starting to explore new options that help them to reorganize their health systems from acute care toward increased chronic care provision (Mac Arthur and Hendry, 2017). Even more, there are hospitals that maintain thresholds policies that state that if an ICU patient falls below the threshold that justify intensive care, they could be justifiably removed from the ICU. However, not many patients do so, making the scarcity of ICU beds continues (Fleck and Murphy, 2018). This is why in this paper we refer to LTC patients in general, and not only geriatric patients, because age is not the only reason that contributes to a prolonged stay in a hospital (Maguire et al., 1986).

They can often take up to 15% of a hospital’s bed capacity thus constituting a significant factor in hospital congestion. Ask Davood for stats about this. Patients who potentially could be transferred to an alternate care location occupied 11% of ventilatorcapable acute care beds in Canada (Rose et al., 2015). There are studies that have even shown that, for some hospitals, half of bed-days could have been potentially avoided (Zhou et al., 2017). Hence, research about the bed-blocking problem can potentially reduce the mounting costs associated with it (Cochran and Bharti, 2006).

In Canada, there is an additional level of complexity that comes with accommodating ALC patients in LTC centres, that is related with social sustainability. One of the major challenges to modelling LTC access is the presence of patient choice. There are a number of papers that have dealt with this challenging issue in other settings. So and Tang (1996) introduce patient choice into the management of operating capacity in two different classes of single-stage service systems and Gupta and Wang (2008) discuss the impact of patient choice on a primary care clinic's optimal profit. In this case, every patient that comes from some hospital or from the community, could have a preference for his/her stay in a particular LTC. This preference usually comes related with proximity to their close relatives or because they want to stay in the same facility as their spouse Patrick et al. (2015). Every patient could list up to five LTC facilities in the region according to their preference, as well as the type of accommodation they are seeking (private beds or ward beds). Hence, if a patient is allocated in a bed that is not his/her first preference, s/he could stay in the wait list until a bed becomes available, or move to another LTC facility as soon as a bed of the appropriate accommodation is available, as it is better that they continue waiting in a LTC home rather than in the hospital (McArdle et al., 1975).

From the hospital perspective, the objective is to keep the number of ALC patients to a manageable level. Thus, a reasonable goal is to determine an "optimal" policy for hospital placements in order to keep the number of ALC patients below a pre-determined threshold while maintaining as much capacity as possible for community demand. And research has enlightened saying that proactive discharge planning and a better access to LTC services are necessary to reduce avoidable bed days (Zhou et al., 2017).

This article seeks to determine policies that help to determine how patients should move from the community and hospitals to the LTC facilities, having many aspects in consideration. The first one is their stay preference, which makes the amount of possible destinations really big. The second one is that patients that are in the community cannot stay for too much time in the wait list, because they can abandon it and go to a costlier private LTC facility. Finally, it seeks to avoid, whenever possible, that the threshold of amount of patients in the hospitals' wait lists is surpassed, for the efficiency of care in ED and ICUs not be affected.

To do this, the authors present a discounted infinite-horizon MDP that addresses this problem with all its considerations. Given the huge size of the state and action vectors, we propose an affine function approximation of the value function of the MDP. The resulting instance uses data on past client preferences to determine a realistic demand stream for the 28 different LTC homes in the region, and the actual demand for LTC homes from both the Community and all the hospitals from the region.

2. Literature review

This literature review requires an update. I propose putting papers regarding blocking patients in general, different methodologies used to address this blocking-bed problems, and how ADP has been used to address healthcare problems but not this one and all it considers.

Research about the bed-blocking problem (e.g., Rubin and Davies 1975, Namdaran et al. 1992, El-Darzi et al. 1998, Koizumi et al. 2005, Cochran and Bharti 2006, Travers et al. 2008, Osorio and Bierlaire 2009, Shi et al. 2015) is important since it can potentially improve the quality of patient care and reduce the mounting costs associated with bed-blocking (Cochran and Bharti 2006). For example, the estimated cost of bed-blocking in the UK exceeds 1.2 billion dollars per year (BBC News 2016). ——————————

A quick scan of major Operations Research journals demonstrates that an impressive amount of effort has gone into improving the day-to-day management of emergency departments. In contrast, there is next to nothing in the Operations Research literature regarding the question of LTC planning. This is somewhat surprising as it is quite clear that improving the efficiency of an upstream process is of limited use if there is a significant backlog downstream causing congestion. Two exceptions are a paper by Weiss and McClain (1987) that uses a queueing analytic approach to describe the process from being labeled as ALC to being placed in a LTC facility and a paper by Koizumi et al. (2005) that models internal movement in a hospital using a queuing approach with blocking to consider downstream congestion. Both papers provide some interesting analytical results of the impact of ALC on hospital congestion but neither seek to determine an optimal policy for LTC placement.

A separate paper by one of the authors (Patrick, 2011) solves the preference problem as a Markov decision process. As might be expected, the resulting policy does not wait for the threshold to be reached before placing hospital patients but rather preemptively insures that the threshold is not reached by ramping up placement as the ALC census approaches the threshold. Though such a model is useful from a planning perspective for the hospital, it fails to take into account the impact of the policy on the wait times of community clients. Thus, a more comprehensive model must also take into account demand from the community.

Zychlinski et al. (2019) considers an analytical approach that minimizes operational costs while addressing a long-term geriatric bed allocation problem. In their model, they consider the community, hospitals, nursing homes and geriatric institutions. However, they do not consider in anyway patients' preferences in their model. Since many works had stated the importance of this factor (Patrick et al., 2015; Hallal, 2015; Rowen et al., 2018; Milte et al., 2018), further work has to be done about this in order to consider it. To our knowledge, this is the first paper that attempts to analyse scheduling policies for LTC placement for both hospital and community demand, while also considering patients' preferences for allocation.

3. The Problem

3.1. The Generalized Problem

The generalized version of the problem is as follows. Demand arrives from multiple wait lists. Each wait list has associated a demand origin and a homogeneous group of clients. There are multiple servers with each client listing a subset of servers where s/he is willing to be served as well as a preferred server. Though clients can be assigned to one of a subset of the servers, there is a cost associated with serving a client at a non-preferred server. Demand origins have their own performance targets regarding length of wait and/or length of queue.

3.2. The Specific Capacity Allocation Problem

Those charged with placing clients in community services as they become available are faced with the difficult task of balancing competing priorities. On the one hand, there is the clear need to keep the number of ALC patients from impinging on hospital function. On the other hand, LTC demand also arises directly from the community. Long wait times in the community result in societal costs both in servicing more complex clients in the non-ideal setting of the home and in the strain such a scenario places on the families involved.

For the purposes of this article, community services refers to both LTC and assisted living or supportive housing (SH). SH acts as middle layer of service between care in the home and LTC. LTC capacity is differentiated both by facility (any given region will have multiple LTC facilities) and by bed type (private, semi-private and ward). Both private and semi-private beds require an additional payment on the part of the client while the ward beds are completely subsidized. Each client stipulates which type of bed and up to three facilities as acceptable placement options. We define the set of destinations as SH or a bed type and facility pairing. Clients who are placed in a non-preferred destination have the choice to remain on the wait list for their preferred destination.

We refer to the set of destinations where a client is willing and/or able to be placed as a preference class. Thus, demand is differentiated by preference class and demand origin. Demand origin refers to whether the client is an urgent community client, a regular community client or a hospital client. Urgent clients receive priority access. For many years, all hospital clients were treated as equivalent to regular community clients on the grounds that hospitals patients are receiving adequate, if inappropriate, care. In April 2006, as a concession to the hospitals, the Community Care Access Center (CCAC), the governing organization that determines placements in the province of Ontario, instituted a policy that for two days of the week, hospital patients were to be classified as urgent. This has improved hospital congestion but remains a rather ad-hoc solution that lacks any quantitative basis. Our goal is to situate the placement of clients on a firmer empirical foundation. To that end, we present an MDP model that intelligently allocates available capacity to the various demand classes.

In principle, such a capacity allocation problem is solvable as an MDP model. However, as the complexities of multiple priority classes, client preferences and a non-homogeneous bed supply are added, the model quickly becomes intractable due to the "curse of dimensionality".

4. A Markov Decision Process Model

4.1. Decision Epochs and State Space

Decisions are assumed to be made once a day. Demand for placement comes in two forms - external demand (i.e., clients waiting for placement) and internal demand (i.e., clients that have already been placed but not in their preferred destination). In our example, external demand comes from three origins - urgent and non-urgent community demand as well as hospital demand. We let I represent the number of external demand classes (preference class and demand origin pairings) and J represent the number of destinations. For ease of notation, we let $[\mathrm{A}] = \{1,\ldots,\mathrm{A}\}$. In addition, we let $J(i)$ represent the set of possible destinations for a client in demand class $i \in [I]$.

We let the state of the system be represented by the vector $\vec{s} = (\vec{u}, \vec{v})$, where u_{in} is the number of clients of demand class $i \in [I]$ who have been waiting $n \in [N]$ periods for service and v_{ijl} is the number of clients of demand class $i \in [I]$ who have been in destination $j \in [J]$ for $l \in [L]$ periods. The state variables u_{iN} and v_{ijL} represent the number of clients of demand class i who have waited at least N periods and the number of clients of demand class i who have been in destination j at least L periods, respectively.

4.2. Actions

The central action is to assign available capacity to waiting demand. This can be accomplished either through external placements (from the wait lists) or internal transfers (from one destination to another). We represent actions by the vector $\vec{a} = (\vec{x}, \vec{y})$, where x_{inj} represents the number of clients from demand class $i \in [I]$ who have been waiting $n \in [N]$ periods to place in destination $j \in [J]$ and y_{ijlk} corresponds to the number of clients from demand class $i \in [I]$ who have been at destination $j \in J(i)$ for $l \in [L]$ periods to transfer to destination $k \in J(i)$, $k \neq j$. We assume that within a demand class, clients are served on a first-in-first-out basis.

Actions are constrained by a number of factors. First, the number of clients assigned to a given destination must not exceed the available capacity at that destination,

$$\sum_{i\in I(j)}\sum_{n\in[N]} x_{inj} + \sum_{i\in I(j)}\sum_{k\in J(i)}\sum_{l\in[L]} y_{iklj} - \sum_{i\in I(j)}\sum_{k\in J(i)}\sum_{l\in[L]} y_{ijlk} \leq C_j - \sum_{i\in I(j)}\sum_{l\in[L]} v_{ijl} \qquad \forall j \in [J] \qquad (1)$$

where C_j is the total capacity (number of beds) at destination j and $I(j)$ represents the set of possible demand classes at destination j. Second, clients can only be assigned to destinations that meet their preferences,

$$x_{inj} = 0 \qquad \forall (i,n,j) \in [I] \times [N] \times [J] \text{ and } j \notin J(i) \tag{2}$$
$$y_{ijlk} = 0 \qquad \forall (i,j,l,k) \in [I] \times [J] \times [L] \times [J] \text{ and } (j \notin J(i) \text{ or } k \notin J(i) \text{ or } j = k) \tag{3}$$

where $J(i)$ is the set of admissible destinations for clients from demand class i. Third, actions cannot exceed waiting demand,

$$\sum_{j \in J(i)} x_{inj} \leq u_{in} \qquad \forall (i,n) \in [I] \times [N] \tag{4}$$
$$\sum_{k \in J(i)} y_{ijlk} \leq v_{ijl} \qquad \forall (i,j,l) \in [I] \times J(i) \times [L] \tag{5}$$

Fourth, to ensure that clients in the same demand class are treated in a first-come-first-served fashion, we impose the following conditions,

$$\sum_{j \in J(i)} x_{inj} \leq M\delta_{in} \qquad \forall (i,n) \in [I] \times [N-1] \tag{6}$$
$$\sum_{m=n+1}^{N} \left(u_{im} - \sum_{j \in J(i)} x_{imj} \right) \leq M(1-\delta_{in}) \qquad \forall (i,n) \in [I] \times [N-1] \tag{7}$$
$$\sum_{k \in J(i)} y_{ijlk} \leq M\delta_{ijl} \qquad \forall (i,j,l) \in [I] \times J(i) \times [L-1] \tag{8}$$
$$\sum_{m=l+1}^{L} \left(v_{ijm} - \sum_{k \in J(i)} y_{ijmk} \right) \leq M(1-\delta_{ijl}) \qquad \forall (i,j,l) \in [I] \times J(i) \times [L-1] \tag{9}$$

where δ_{in} and δ_{ijl} are binary variables and M is a sufficiently large number. The first two equations force $\sum_{j \in J(i)} x_{inj}$ to be zero unless $u_{im} - \sum_{j \in J(i)} x_{imj} = 0 \quad \forall m > n$. The last two equations force $\sum_{k \in J(i)} y_{ijlk}$ to be zero unless $v_{ijm} - \sum_{k \in J(i)} y_{ijmk} = 0 \quad \forall m > l$. To ensure that the state space remains finite even if capacity is insufficient, we introduce a further action z_i that removes demand entirely from the system. This could be a last resort measure that pays for clients to be placed in private retirement homes. This action satisfies the following two conditions:

$$z_i \geq u_{iN} - \sum_{j \in J(i)} x_{iNj} - W_i \qquad \forall i \in [I] \tag{10}$$
$$z_i \leq u_{iN} - \sum_{j \in J(i)} x_{iNj} \qquad \forall i \in [I] \tag{11}$$

Equation (10) guarantees that the number of clients of demand class i who have been waiting for at least N periods never exceeds W_i. We will place a high cost to any positive z_i thus ensuring that this will be a last resort action. Finally, all actions must be positive and integer,

$$x_{inj} \in \mathbb{Z}^+ \qquad \forall (i,n,j) \in [I] \times [N] \times J(i) \tag{12}$$
$$y_{ijlk} \in \mathbb{Z}^+ \qquad \forall (i,j,l,k) \in [I] \times J(i) \times [L] \times J(i) \backslash \{j\} \tag{13}$$
$$z_i \in \mathbb{Z}^+ \qquad \forall i \in [I] \tag{14}$$

We define the set $A(\vec{s})$ as the set of actions satisfying equations (1) to (14) for a given state $\vec{s}$.

4.3. Transition Probabilities

New demand can arrive to any of the wait lists (demand classes) while exits from the system can occur from any wait list or destination (since clients can die while on a wait list). Some movement between destinations is also possible as clients may be transferred. Transitions are written as:

$$\begin{aligned}\vec{s}=(\vec{u},\vec{v}) \to \Bigg(\Bigg\{& d_i, u_{i1} - \sum_{j\in J(i)} x_{i1j} - e_{i1}, \ldots, u_{i(N-2)} - \sum_{j\in J(i)} x_{i(N-2)j} - e_{i(N-2)}, \\ & \sum_{n=N-1}^{N}\Bigg(u_{in} - \sum_{j\in J(i)} x_{inj} - e_{in}\Bigg) - z_i\Bigg\}_{i\in[I]}, \\ & \Bigg\{\sum_{n\in[N]} x_{inj} - e_{ij0}, v_{ij1} + \sum_{k\in J(i)} y_{ik1j} - \sum_{k\in J(i)} y_{ij1k} - e_{ij1}, \ldots, v_{ij(L-2)} + \sum_{k\in J(i)} y_{ik(L-2)j} \\ & - \sum_{k\in J(i)} y_{ij(L-2)k} - e_{ij(L-2)}, \sum_{l=L-1}^{L}\Bigg(v_{ijl} + \sum_{k\in J(i)} y_{iklj} - \sum_{k\in J(i)} y_{ijlk} - e_{ijl}\Bigg)\Bigg\}_{i\in[I], j\in J(i)}\Bigg)\end{aligned}$$

where d_i corresponds to new demand, e_{in} represents exits from the wait lists after n periods and e_{ijl} are departures from the destinations after l periods.

4.4. Costs

There are costs associated with clients waiting too long, with the size of the wait list at each facility, with clients residing in non-preferred destinations, with transferences of clients and of course with the last resort action of removing demand from the wait lists. We write the cost as:

$$\begin{aligned}c(\vec{s},\vec{a}) = \sum_{f\in[F]} c_f^W \left[\sum_{i\in I(f)}\sum_{n\in[N]} u_{in} - T_f\right]^+ + \sum_{i\in[I]}\sum_{n\in[N]} c_{in}^W u_{in} + \sum_{i\in[I]}\sum_{j\in J(i)}\sum_{l\in[L]} c_{ijl}^P v_{ijl} \\ \sum_{i\in[I]}\sum_{j\in J(i)}\sum_{l\in[L]}\sum_{k\in J(i)} b_{ijlk} y_{ijlk} + \sum_{i\in[I]}\sum_{n\in[N]}\sum_{j\in J(i)} a_{inj} x_{inj} + \sum_{i\in[I]} c_i^D z_i \qquad (15)\end{aligned}$$

where $I(f)$ represents the set of demand classes associated with demand origin $f \in [F]$ and T_f is the maximum recommended size of the corresponding wait list.

4.5. The Bellman Equation

The value function $v(\vec{s})$ of the MDP specifies the minimum discounted cost over the infinite horizon for each state and satisfies the following optimality equations:

$$\begin{aligned}v(\vec{s}) = \min_{\vec{a}\in A(\vec{s})} \Bigg\{& c(\vec{s},\vec{a}) + \gamma \sum_{\substack{(\vec{d},\vec{e})\in \\ D\times E}} \Pr(\vec{D}=\vec{d})\Pr(\vec{E}=\vec{e}|\vec{s},\vec{a}) \times \qquad (16) \\ & v\Bigg(\Bigg\{d_i, u_{i1} - \sum_{j\in J(i)} x_{i1j} - e_{i1}, \ldots, u_{i(N-2)} - \sum_{j\in J(i)} x_{i(N-2)j} - e_{i(N-2)}, \\ & \sum_{n=N-1}^{N}\Bigg(u_{in} - \sum_{j\in J(i)} x_{inj} - e_{in}\Bigg) - z_i\Bigg\}_{i\in[I]},\end{aligned}$$

$$\left\{\sum_{n\in[N]} x_{inj} - e_{ij0}, v_{ij1} + \sum_{k\in J(i)} y_{ik1j} - \sum_{k\in J(i)} y_{ij1k} - e_{ij1}, \ldots, v_{ij(L-2)} + \sum_{k\in J(i)} y_{ik(L-2)j} \right.$$
$$\left. - \sum_{k\in J(i)} y_{ij(L-2)k} - e_{ij(L-2)}, \sum_{l=L-1}^{L}\left(v_{ijl} + \sum_{k\in J(i)} y_{iklj} - \sum_{k\in J(i)} y_{ijlk} - e_{ijl}\right)\right\}_{i\in[I], j\in J(i)}\Bigg)$$

$$\forall \vec{s} \in S$$

5. Linear Programming Approach

Converting the Bellman equation into the equivalent linear program yields the following LP:

$$\max_{v\in\mathbb{R}} \sum_{\vec{s}\in S} \alpha(\vec{s}) v(\vec{s}) \tag{17}$$

subject to

$$c(\vec{s},\vec{a}) + \gamma \sum_{\substack{(\vec{d},\vec{e})\in \\ D\times E}} \Pr(\vec{D}=\vec{d}) \Pr(\vec{E}=\vec{e}|\vec{s},\vec{a}) \times$$
$$v\left(\left\{d_i, u_{i1} - \sum_{j\in J(i)} x_{i1j} - e_{i1}, \ldots, u_{i(N-2)} - \sum_{j\in J(i)} x_{i(N-2)j} - e_{i(N-2)}, \right.\right.$$
$$\left.\sum_{n=N-1}^{N}\left(u_{in} - \sum_{j\in J(i)} x_{inj} - e_{in}\right) - z_i\right\}_{i\in[I]},$$
$$\left\{\sum_{n\in[N]} x_{inj} - e_{ij0}, v_{ij1} + \sum_{k\in J(i)} y_{ik1j} - \sum_{k\in J(i)} y_{ij1k} - e_{ij1}, \ldots, v_{ij(L-2)} + \sum_{k\in J(i)} y_{ik(L-2)j} \right.$$
$$\left. - \sum_{k\in J(i)} y_{ij(L-2)k} - e_{ij(L-2)}, \sum_{l=L-1}^{L}\left(v_{ijl} + \sum_{k\in J(i)} y_{iklj} - \sum_{k\in J(i)} y_{ijlk} - e_{ijl}\right)\right\}_{i\in[I], j\in J(i)}\Bigg)$$
$$\geq v(\vec{s}) \qquad \forall (\vec{s},\vec{a}) \in S \times A(\vec{s})$$

Neither the Bellman equations given in Equation (16) nor the linear program given in (17) are tractable due to the sheer size of the state space. We thus seek to solve the MDP using the linear programming approach to approximate dynamic programming (ADP). In this methodology, the value function is assumed to have a certain parametric form with a reasonable number of parameter. This approximate form is inserted into the linear program in equation (17) and the "optimal" value function approximation is determined. We use the following approximation, that was first proposed by Schweitzer and Seidmann (1985) and has given succesful results by several works (Adelman, 2004, 2007; de Farias and Roy, 2004, 2006; Sauré et al., 2012; Patrick et al., 2008; Sauré and Puterman, 2017; González et al., 2018; Marquinez et al., 2019) More references:

$$v(\vec{s}) = U_0 + \sum_{i\in[I]}\sum_{n\in[N]} U_{in} u_{in} + \sum_{i\in[I]}\sum_{j\in J(i)}\sum_{l\in[L]} V_{ijl} v_{ijl} \tag{18}$$
$$U_{in} \geq 0, \quad \forall (i,n) \in [I]\times[N], \qquad V_{ijl} \geq 0, \quad \forall (i,j,l) \in [I]\times J(i)\times[L]$$

Substituting (18) into (17) yields the following approximate linear program (ALP):

$$\max_{\vec{U},\vec{V}\geq 0}\left\{U_0+\sum_{i\in[I]}\sum_{n\in[N]}E_\alpha[u_{in}]U_{in}+\sum_{i\in[I]}\sum_{j\in J(i)}\sum_{l\in[L]}E_\alpha[v_{ijl}]V_{ijl}\right\} \tag{19}$$

subject to

$$\begin{aligned}
&(1-\gamma)U_0+\sum_{i\in[I]}(u_{i1}-\gamma\lambda_i)U_{i1}+\sum_{i\in[I]}\sum_{n=2}^{N-1}\left[u_{in}-\gamma(1-\rho_{i(n-1)})\left(u_{i(n-1)}-\sum_{j\in J(i)}x_{i(n-1)j}\right)\right]U_{in}\\
&+\sum_{i\in[I]}\left\{u_{iN}-\left(\sum_{n=N-1}^{N}\left(\gamma(1-\rho_{in})\left[u_{in}-\sum_{j\in J(i)}x_{inj}\right]\right)\right)+\gamma(1-\rho_{iN})z_i\right\}U_{iN}\\
&+\sum_{i\in[I]}\sum_{j\in J(i)}\left[v_{ij1}-\gamma(1-\rho_{ij0})\left(\sum_{n\in[N]}x_{inj}\right)\right]V_{ij1}\\
&+\sum_{i\in[I]}\sum_{j\in J(i)}\sum_{l=2}^{L-1}\left[v_{ijl}-\gamma(1-\rho_{ij(l-1)})\left(v_{ij(l-1)}+\sum_{k\in J(i)}y_{ik(l-1)j}-\sum_{k\in J(i)}y_{ij(l-1)k}\right)\right]V_{ijl}\\
&+\sum_{i\in[I]}\sum_{j\in J(i)}\left\{v_{ijL}-\sum_{l=L-1}^{L}\left(\gamma(1-\rho_{ijl})\left[v_{ijl}+\sum_{k\in J(i)}y_{iklj}-\sum_{k\in J(i)}y_{ijlk}\right]\right)\right\}V_{ijL}\\
&\leq c(\vec{s},\vec{a}) \qquad \forall(\vec{s},\vec{a})\in S\times A(\vec{s})
\end{aligned}$$

where $\vec{\lambda}$ represents the mean arrival rates and $\vec{\rho}$ the probabilities of individuals departing the system. The ALP results in a linear program with a reasonable number of variables but an intractable number of constraints so we look to solve the dual through column generation. The dual of (19) is:

$$\min_{\vec{X}\geq 0}\left\{\sum_{(\vec{s},\vec{a})\in S\times A(\vec{s})}c(\vec{s},\vec{a})X(\vec{s},\vec{a})\right\} \tag{20}$$

subject to

$$(1-\gamma)\sum_{\substack{(\vec{s},\vec{a})\in\\ S\times A(\vec{s})}}X(\vec{s},\vec{a})=1$$

$$\sum_{\substack{(\vec{s},\vec{a})\in\\ S\times A(\vec{s})}}(u_{i1}-\gamma\lambda_i)X(\vec{s},\vec{a})\geq E_\alpha[u_{i1}] \quad \forall i\in[I]$$

$$\sum_{\substack{(\vec{s},\vec{a})\in\\ S\times A(\vec{s})}}\left[u_{in}-\gamma(1-\rho_{i(n-1)})\left(u_{i(n-1)}-\sum_{j\in J(i)}x_{i(n-1)j}\right)\right]X(\vec{s},\vec{a})\geq E_\alpha[u_{in}]$$
$$\forall(i,n)\in[I]\times\{2,\ldots,N-1\}$$

$$\sum_{\substack{(\vec{s},\vec{a})\in\\ S\times A(\vec{s})}}\left\{u_{iN}-\left(\sum_{n=N-1}^{N}\left(\gamma(1-\rho_{in})\left[u_{in}-\sum_{j\in J(i)}x_{inj}\right]\right)\right)+\gamma(1-\rho_{iN})z_i\right\}X(\vec{s},\vec{a})\geq E_\alpha[u_{iN}] \quad \forall i\in[I]$$

$$\sum_{\substack{(\vec{s},\vec{a})\in\\ S\times A(\vec{s})}} \left[v_{ij1} - \gamma(1-\rho_{ij0})\left(\sum_{n\in[N]} x_{inj}\right)\right] X(\vec{s},\vec{a}) \geq E_\alpha[v_{ij1}] \quad \forall (i,j)\in[I]\times J(i)$$

$$\sum_{\substack{(\vec{s},\vec{a})\in\\ S\times A(\vec{s})}} \left[v_{ijl} - \gamma(1-\rho_{ij(l-1)})\left(v_{ij(l-1)} + \sum_{k\in J(i)} y_{ik(l-1)j} - \sum_{k\in J(i)} y_{ij(l-1)k}\right)\right] X(\vec{s},\vec{a}) \geq E_\alpha[v_{ijl}] \quad \forall (i,j,l)\in[I]\times J(i)\times\{2,\dots,L-1\}$$

$$\sum_{\substack{(\vec{s},\vec{a})\in\\ S\times A(\vec{s})}} \left\{v_{ijL} - \sum_{l=L-1}^{L}\left(\gamma(1-\rho_{ijl})\left[v_{ijl} + \sum_{k\in J(i)} y_{iklj} - \sum_{k\in J(i)} y_{ijlk}\right]\right)\right\} X(\vec{s},\vec{a}) \geq E_\alpha[v_{ijL}] \quad \forall (i,j)\in[I]\times J(i)$$

Here we have assumed that U_{in} and V_{ijl} are positive. This makes intuitive sense as one would expect a positive cost for each additional client waiting to be placed and each additional bed currently occupied. Solving the dual has the advantage of a reasonable number of constraints but at the expense of creating an intractable number of variables - one for each state-action pair. Column generation solves this problem by starting with a small set of feasible state-action pairs to the dual and then (using the dual prices as estimates for $\vec{U}$ and $\vec{V}$) finding one or more violated constraints in the primal. It then adds the state-action pair(s) associated with these violated constraints into the current set of columns before re-solving the dual. The process iterates until either no primal constraint is violated or one is "close enough" to optimality to quit. The integer program required to find the most violated constraint is:

$$\begin{aligned}
\min_{\substack{(\vec{s},\vec{a})\in\\ S\times A(\vec{s})}} \quad & \Bigg\{\sum_f c_f^W \left[\sum_{i\in I(f)}\sum_{n\in[N]} u_{in} - T_f\right]^+ \\
&+ \sum_{i\in[I]}\sum_{n=1}^{N-1}\left[c_{in}^W + \gamma(1-\rho_{in})U_{i(n+1)} - U_{in}\right] u_{in} \\
&+ \sum_{i\in[I]}\left[c_{iN}^W + \gamma(1-\rho_{iN})U_{iN} - U_{iN}\right] u_{iN} \\
&+ \sum_{i\in[I]}\sum_{j\in J(i)}\sum_{l=1}^{L-1}\left[c_{ijl}^P + \gamma(1-\rho_{ijl})V_{ij(l+1)} - V_{ijl}\right] v_{ijl} \\
&+ \sum_{i\in[I]}\sum_{j\in J(i)}\left[c_{ijL}^P + \gamma(1-\rho_{ijL})V_{ijL} - V_{ijL}\right] v_{ijL} \\
&+ \sum_{i\in[I]}\sum_{n=1}^{N-1}\sum_{j\in J(i)}\left[a_{inj} + \gamma(1-\rho_{ij0})V_{ij1} - \gamma(1-\rho_{in})U_{i(n+1)}\right] x_{inj} \\
&+ \sum_{i\in[I]}\sum_{j\in J(i)}\left[a_{iNj} + \gamma(1-\rho_{ij0})V_{ij1} - \gamma(1-\rho_{iN})U_{iN}\right] x_{iNj} \\
&+ \sum_{i\in[I]}\sum_{j\in J(i)}\sum_{l=1}^{L-1}\sum_{k\in J(i)}\left[b_{ijlk} + \gamma(1-\rho_{ikl})V_{ik(l+1)} - \gamma(1-\rho_{ijl})V_{ij(l+1)}\right] y_{ijlk} \\
&+ \sum_{i\in[I]}\sum_{j\in J(i)}\sum_{k\in J(i)}\left[b_{ijLk} + \gamma(1-\rho_{ikL})V_{ikL} - \gamma(1-\rho_{ijL})V_{ijL}\right] y_{ijLk} \\
&+ \sum_{i\in[I]}\left[c_i^D - \gamma(1-\rho_{iN})U_{iN}\right] z_i
\end{aligned}$$

$$+\sum_{i\in[I]}\gamma\lambda_i U_{i1}-(1-\gamma)U_0\Bigg\}$$

5.1. Approximate Optimal Policy model

When the process finishes, we use the values $\vec{U}^*$ and $\vec{V}^*$ to identify the approximate optimal policy. Since computing and storing approximate optimal actions for each state is very expensive, we can compute them as needed. When the value function defined by $\vec{U}^*$ and $\vec{V}^*$ is inserted in the right hand side of (16), the approximate optimal policy $d^*(\vec{s})$ is obtained solving the obtained IP model:

$$\begin{aligned} d^*(\vec{s}) \in \arg\min_{\vec{a}\in A_{\vec{s}}} \Bigg\{ & \sum_{i\in[I]}\sum_{n=1}^{N-1}\sum_{j\in J(i)}\left[a_{inj}+\gamma(1-\rho_{ij0})V^*_{ij1}-\gamma(1-\rho_{in})U^*_{i(n+1)}\right]x_{inj} \\ & +\sum_{i\in[I]}\sum_{j\in J(i)}\left[a_{iNj}+\gamma(1-\rho_{ij0})V^*_{ij1}-\gamma(1-\rho_{iN})U^*_{iN}\right]x_{iNj} \\ & +\sum_{i\in[I]}\sum_{j\in J(i)}\sum_{l=1}^{L-1}\sum_{k\in J(i)}\left[b_{ijlk}+\gamma(1-\rho_{ikl})V^*_{ik(l+1)}-\gamma(1-\rho_{ijl})V^*_{ij(l+1)}\right]y_{ijlk} \\ & +\sum_{i\in[I]}\sum_{j\in J(i)}\sum_{k\in J(i)}\left[b_{ijLk}+\gamma(1-\rho_{ikL})V^*_{ikL}-\gamma(1-\rho_{ijL})V^*_{ijL}\right]y_{ijLk} \\ & +\sum_{i\in[I]}\left[c^D_i-\gamma(1-\rho_{iN})U^*_{iN}\right]z_i\Bigg\}+\text{constant}. \end{aligned} \tag{21}$$

6. Results and analysis

6.1. Data

Ir order to solve this model with a real instance, the CCAC has provided us with data. This data set, that includes all clients on the wait list in a 3-year period, was used to estimate the parameters. From this realistic demand stream for the 28 different LTC homes in the region, we determined the distribution of demand, the rate at which LTC beds become available, the relative preferences, and the rate at which clientes leave wait lists. This data suggests that both hospital and community arrival rates follow Poisson distributions; as does the number of beds that become available (see Figure 1 for example).

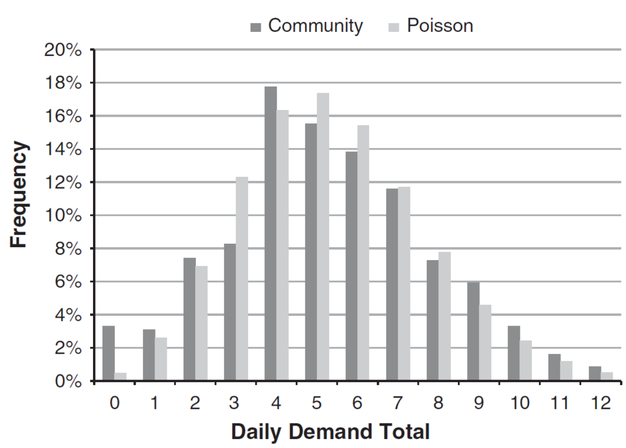

Figure 1 **Distribution of daily demand for community clients.**

In order to make the model tractable, we decided to divide the Ottawa region into five sectors, and chose to work only with the east region, consistent in 8 of those 28 LTC homes. Each of these 8 centres offer both private and ward beds in them. These centres offers 1,332 beds, 29.40% of the total beds for the whole Ottawa region. With the same argument, and because the time between discharges and abandons follow an exponential distribution, we consider that $L = 3$. This is a valid assumption, since the probability of discharge of abandon is the same for each period. However, we set $N = 13$ since we want to start sending clients in the wait list to private LTC centres only when patients have been waiting for 13 or more weeks without being placed in any facility.

On the other side, we considered the five main hospitals in the region and the community as the six possible patients' source. Table 1 shows the daily arrival rate for each source, broken down by accommodation type. Each patient can choose between a ward or private bed, but can also state a willingness to take either option. As can be seen, almost 9 patients comes to a wait list everyday.

	Private bed	Ward bed	Either bed	All
Community	2.92	1.76	0.79	5.47
Hospital #1 (TOH)	0.45	0.76	0.70	1.91
Hospital #2 (MON)	0.23	0.15	0.08	0.46
Hospital #3 (QCH)	0.29	0.23	0.14	0.65
Hospital #4 (EB)	0.11	0.07	0.04	0.23
Hospital #5 (OTH)	0.07	0.17	0.00	0.24
All demand	4.06	3.15	1.75	8.96

Table 1 Average daily demand for the wait list from the community and each of the hospitals considered.

As for the preferences, Figure 2 depict that not all the LTC homes are equally preferred. Since clients are encouraged to take the first available bed till they can be transferred to their preferred facility, some centres act as holding bays. Since not all facilities are equally preferred, there is a significant variation in wait times, given that patients will be finally be directed to their top preferred facility. These preferences were based on historical preferences of past clients, and we considered —again for tractability— only the first preference, since there is the final destination of each patient. These percentages were signifactivelly different for those in the hospitals and those in the community, and each one of them was weighted with the corresponding type of accommodation preference and demand at each source of patients. This we did to capture client preferences into the model, for us to keep up with the social aspect of sustainability.

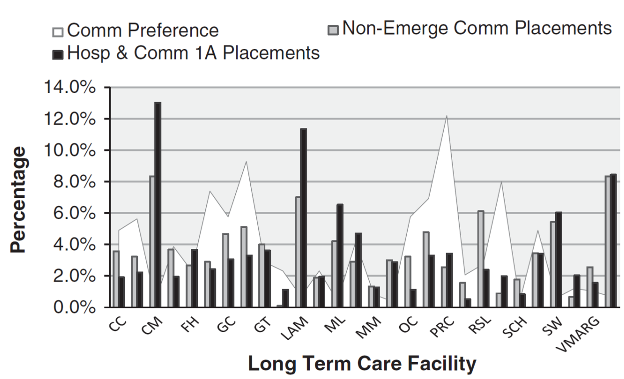

Figure 2 Difference between preferences and initial placements. This bars consider only first placement location for all the clients considered, and the white shaded region represents client facility preferences.

6.2. Results

7. Conclusions

8. Citas a incluir

The higher the proportion of institutional beds per 1000 inhabitants over the age of 80 in a county, the fewer bed-blockers in acute wards. Similar results have been mentioned from Canada (Roos, 1989). (Styrborn and Thorslund, 1993)

Differences in health and long-term care spending emerge across OECD countries partly reflecting differing demographic trends as well as initial levels of income and informal long-term care supply. Korea, Chile, Turkey and Mexico, for example, are projected to experience above average increases in public health expenditures. (Maisonneuve and Martins, 2013)

Overall, for hip fracture patients we find evidence consistent with the 'bed-blocking' hypothesis that availability of long-term care affects the length of stay of patients who no longer need to be in acute hospital and are ready to be discharged. Caring for such patients in hospital is more costly than long-term care. Our results suggest that for hip fracture patients an expansion of the long-term care sector can reduce hospital length of stay and reduce the total cost of caring for these patients. (Gaughan et al., 2017)

Using a point-prevalence survey, we found that PMV patients occupied 11% of Canadian acute care ventilator bed capacity. Limited discharge options and significant discharge barriers require consideration in health policy decision making and health care resource allocation. (Rose et al., 2015)

Home Health-Care (HHC) is a concept slowly expanding over time, introduced to reduce pressure on inpatient hospital beds by providing care to patients at home. (Rodriguez-Verjan et al., 2017)

Bed Blocking in Hospitals due to Scarce Capacity in Geriatric Institutions – Cost Minimization via Fluid Models - Noa Zychlinski, Avishai Mandelbaum, Petar Momcilovic, Izack Cohen. HICIERON LO MISMO PERO SIN CONSIDERAR PREFERENCIAS DE PACIENTES. Zychlinski et al. (2019)

Acknowledgments

Acknowledge Conicyt and the CCAC.

References

Adelman, Daniel. 2004. A Price-Directed Approach to Stochastic Inventory/Routing. *Operations Research* **52**(4) 499–514.

Adelman, Daniel. 2007. Dynamic Bid Prices in Revenue Management. *Operations Research* **55**(4) 647–661.

Cochran, Jeffery K., Aseem Bharti. 2006. Stochastic bed balancing of an obstetrics hospital. *Health Care Management Science* **9**(1) 31–45.

de Farias, Daniela Pucci, Benjamin Van Roy. 2004. On Constraint Sampling in the Linear Programming Approach to Approximate Dynamic Programming. *Mathematics of Operations Research* **29**(3) 462–478.

de Farias, Daniela Pucci, Benjamin Van Roy. 2006. A Cost-Shaping Linear Program for Average-Cost Approximate Dynamic Programming with Performance Guarantees. *Mathematics of Operations Research* **31**(3) 597–620.

Fleck, Leonard M., Timothy F. Murphy. 2018. First Come, First Served in the Intensive Care Unit: Always? *Cambridge Quarterly of Healthcare Ethics* **27**(01) 52–61.

Gaughan, James, Hugh Gravelle, Rita Santos, Luigi Siciliani. 2017. Long-term care provision, hospital bed blocking, and discharge destination for hip fracture and stroke patients. *International Journal of Health Economics and Management* **17**(3) 311–331.

González, Jaime, Juan-Carlos Ferrer, Alejandro Cataldo, Luis Rojas. 2018. A proactive transfer policy for critical patient flow management. *Health Care Management Science* 1–17.

Gupta, Diwakar, Lei Wang. 2008. Revenue Management for a Primary-Care Clinic in the Presence of Patient Choice. *Operations Research* **56**(3) 576–592.

Hallal, Louisa. 2015. Queuing Models for Long Term Care Wait Time Reduction & Capacity Optimization: A Nova Scotia Study. Mastersthesis, Saint Mary's University.

Koizumi, N., E. Kuno, T. E. Smith. 2005. Modeling Patients Flows Using a Queueing Network with Blocking. *Health Care Management Science* **8** 49–60.

Mac Arthur, Ian, Anne Hendry. 2017. The "Intermediate Care Hospital:" Facility Bed-Based Rehabilitation for Elderly Patients. Tech. rep., Inter-American Development Bank.

Maguire, Pamela A, Ian C Taylor, Robert W Stout. 1986. Elderly patients in acute medical wards: factors predicting length of stay in hospital. *Bmj* **292**(6530) 1251–1253.

Maisonneuve, Christine De La, Joaquim Oliveira Martins. 2013. Public spending on health and long-term care: a new set of projections. *Documents de travail du Département des affaires économiques de l'OCDE* **6**(06) 1–39.

Marquinez, José Tomás, Antoine Sauré, Alejandro Cataldo, Juan Carlos Ferrer. 2019. A framework for dynamic ICU bed management in a public-private multi-hospital system.

McArdle, Christine, J. C. Wiley, W.D. Alexander. 1975. Geriatric patients in an acute medical ward. *British Medical Journal* **4**(5996) 568–569.

Milte, Rachel, Julie Ratcliffe, Gang Chen, Maria Crotty. 2018. What Characteristics of Nursing Homes Are Most Valued by Consumers? A Discrete Choice Experiment with Residents and Family Members. *Value in Health* **21**(7) 843–849.

Patrick, Jonathan. 2011. Access to long-term care: The true cause of hospital congestion? *Production and Operations Management* **20**(3) 347–358.

Patrick, Jonathan, K Nelson, D. Lane. 2015. A simulation model for capacity planning in community care. *Journal of Simulation* **9**(2) 111–120.

Patrick, Jonathan, Martin L. Puterman, Maurice Queyranne. 2008. Dynamic Multipriority Patient Scheduling for a Diagnostic Resource. *Operations Research* **56**(6) 1507–1525.

Rodriguez-Verjan, Carlos, Vincent Augusto, Xiaolan Xie. 2017. Home health-care network design: Location and configuration of home health-care centers. *Operations Research for Health Care* .

Roos, Noralou P. 1989. How a Universal Health Care System Responds to an Aging Population. *Journal of Aging and Health* **1**(4) 411–429.

Rose, Louise, Robert A. Fowler, Eddy Fan, Ian Fraser, David Leasa, Cathy Mawdsley, Cheryl Pedersen, Gordon Rubenfeld. 2015. Prolonged mechanical ventilation in Canadian intensive care units: A national survey. *Journal of Critical Care* **30**(1) 25–31.

Rowen, Donna, Katherine Stevens, Alexander Labeit, Jackie Elliott, Brendan Mulhern, Jill Carlton, Hasan Basarir, Julie Ratcliffe, John Brazier. 2018. Using a Discrete-Choice Experiment Involving Cost to Value a Classification System Measuring the Quality-of-Life Impact of Self-Management for Diabetes. *Value in Health* **21**(1) 69–77.

Rubin, S. G., G. H. Davies. 1975. Bed blocking by elderly patients in General-Hospital Wards. *Age and Ageing* **4**(March) 142–147.

Sauré, Antoine, Jonathan Patrick, Scott Tyldesley, Martin L. Puterman. 2012. Dynamic multi-appointment patient scheduling for radiation therapy. *European Journal of Operational Research* **223**(2) 573–584.

Sauré, Antoine, Martin L. Puterman. 2017. Advance Patient Appointment Scheduling. Richard J. Boucherie, Nico M. van Dijk, eds., *Markov decision processes in practice*, chap. 8. Springer International Publishing, Cham, 245–268.

Schweitzer, Paul J., Abraham Seidmann. 1985. Generalized polynomial approximations in Markovian decision processes. *Journal of Mathematical Analysis and Applications* **110**(2) 568–582.

So, Kut C., Christopher S. Tang. 1996. On managing operating capacity to reduce congestion in service systems. *European Journal of Operational Research* **92**(1) 83–98.

Styrborn, Karin, Mats Thorslund. 1993. 'Bed-blockers': Delayed discharge of hospital patients in a nationwide perspective in Sweden. *Health policy* **26**(2) 155–170.

Weiss, Elliott N, John O McClain. 1987. Administrative Days in Acute Care Facilities: A Queueing-Analytic Approach **35**(1) 35–44.

Zhou, Ke, Arpana R. Vidyarthi, Chek Hooi Wong, David Matchar. 2017. Where to go if not the hospital? Reviewing geriatric bed utilization in an acute care hospital in Singapore. *Geriatrics and Gerontology International* **17**(10) 1575–1583.

Zychlinski, Noa, Avishai Mandelbaum, Petar Momčilović, Izack Cohen. 2019. Bed Blocking in Hospitals Owing to Scarce Capacity in Geriatric Institutions - Cost Minimization via Fluid Models. *Manufacturing & Service Operations Management* .