

Markov Decision Process Model for Optimisation of Patient Flow

A. Clissold, J. Filar, S. Qin and D. Ward ^a

^a*School of Computer Science, Engineering and Mathematics, Flinders University, Australia
Email: jerzy.filar@flinders.edu.au*

Abstract: Australian hospitals are facing the challenge of increased demand in ED and inpatient services. Congestion and poor quality of care for some patients may occur as a consequence of such challenging situations. Modelling patient flow is a way to capture a small number of salient characteristics of a hospital operation permitting us to investigate the causes and effects of congestion, and identify optimal improvement policies.

The data used in this project are drawn from an extensive Patient Journey Database of a major South Australia hospital, Flinders Medical Centre (FMC). This database tracks journeys of patients through the hospital system, from arrival to discharge. This allows for very detailed analysis and modelling of these journeys.

Analysis of historical bed occupancy data indicates that there are remarkably regular patterns in patient occupancy, on both daily and weekly time scales. Based on this observation, we develop a Markov decision process (MDP) model over a weekly time horizon to analyse FMC's bed occupancy patterns with a goal of proposing policies that lead to reduced incidence of congestion.

A simple, four state, approach is used in categorising the occupancy level, and daily and weekly probability transition matrices are derived from midnight census data over a three year period. We then introduce a reward function that attempts to balance the benefit of maintaining sufficient buffer in the occupancy of wards versus the cost of diverting patients elsewhere. An optimal policy from the MDP model was found using the dynamic programming approach.

Employing an MDP approach in these problems will also, potentially, allow hospital staff the ability to forecast possible onset of very high occupancy levels, based on the hospital's current situation. Such a "warning signal" may enable hospital managers to be proactive in their strategies to reduce the magnitude and impacts of congestion episodes. Although, to managers, the underlying mathematics may appear relatively complex, it is possible to embed an MDP model inside a software package so that the model is available in a more user-friendly form, allowing it to be easily run on-the-fly so that staff can react to changes in the hospital in real time.

The issue of validation of "optimal" policies suggested by an MDP model is addressed in three, separate, ways. Firstly, the steady state probabilities resulting from such policies are examined to see if they reduce the long-run frequency of highest occupancy levels. Secondly, consultation with FMC experts, an integral part of this study, has provided a degree of qualitative validation. Thirdly, synthetic data generated by a discrete event simulation system developed in another component of the larger project is used to provide surrogate statistical testing environment for the performance of MDP's policies. Simulated experiments also serve the purpose of convincing hospital's management before piloting policies in real setting.

Future work includes refinement of both the MDP and the simulation model and establishment of a formal framework for cross-validating MDP and simulation modelling results.

Keywords: *Markov decision process, hospital patient flow, process optimisation*

1 INTRODUCTION

Australia's growing ageing population and a rise in chronic diseases are causing an increased demand for emergency and inpatient services in the country's hospitals FitzGerald *et al.* (2012). Currently, many hospitals are often operating at or close to full capacity. Coupled with the stochastic nature of bed availability, patient arrivals and the duration of treatments and care, episodes of congestion often occur, which lead to wide-ranging effects such as decreased quality of care and fatigued staff, among others.

However, despite uncertainties, it has been discovered through analysis of historical data that there are remarkably regular patterns in patient occupancy, on both day-to-day and week-to-week time scales. These patterns can suitably be modelled mathematically as Markov processes.

Markov processes and their extensions have been used in many applications in the health care setting including admission control Helm *et al.* (2011), resource availability Garg *et al.* (2012), inventory management Broyles *et al.* (2010) and bed occupancy Christodoulou and Taylor (2001).

In this study, we develop a Markov decision process (MDP) model over a weekly time horizon. We use a given hospital's occupancy data to construct probability transitions consistent with these data. Also, we introduce a reward function that attempts to balance the benefit of maintaining adequate (but small) buffer in the occupancy of wards versus the cost of diverting patients elsewhere. The weekly time scale is chosen due to the natural operation cycle of most hospitals.

2 THE MDP MODEL

The MDP model was formulated using midnight occupancy data provided from the Flinders Medical Centre (FMC), a metropolitan hospital in the southern Adelaide area. The data were extracted from FMC's Patient Journey Database, for the periods of January 2009 to June 2013. Only patients within the General Medical and Surgical streams were considered. These represent approximately 65% of inpatient bed allocations at the hospital. According to expert advice, these streams of patients constitute the main source of variability in occupancy levels at FMC. Further, the period between December the 25th and 31st (inclusive) was omitted from the analysis due to the Hospital's policy of minimising bed occupancy over this period.

2.1 Daily Probability Transition Matrices

After discussions with FMC staff a simple, four state, model was designed. It was felt that these four states, while coarse, achieved adequate balance between interpretability and parametric complexity on the one hand, and the degree of detail on the other hand. Too many states are difficult to interpret for health care professionals (and require estimates of too many parameters), whereas too few make the model too coarse. Based on consultation with FMC staff, the four occupancy states were defined on the basis of deviations from the mean occupancy. The boundaries between the four states are given below.

Definition 1. *The states $\{1, 2, 3, 4\}$ will correspond intuitively to Low, Medium, High and Very High occupancies as indicated below:*

$$\left\{ \begin{array}{ll} \text{Low}(S = 1) : & s \leq \mu - \frac{\sigma}{2} & [0, 279) \\ \text{Medium}(S = 2) : & \mu - \frac{\sigma}{2} < s \leq \mu + \frac{3\sigma}{4} & [279, 307) \\ \text{High}(S = 3) : & \mu + \frac{3\sigma}{4} < s \leq \mu + \frac{3\sigma}{2} & [307, 322) \\ \text{Very High}(S = 4) : & s > \mu + \frac{3\sigma}{2} & [322, \infty) \end{array} \right. ,$$

where the aggregated states denoted by S actually correspond to collections of micro-states denoted by s that are the daily occupancy levels. The parameters, μ and σ denote the overall mean and standard deviation of occupancy levels in the calibration portion of the data set used. Also shown above are the occupancy intervals for the respective aggregated states calculated from the historical data.

Each day within the data is classified as Low, Medium, High or Very High occupancy using Definition 1, and the transition from weekday to weekday ($P_i, i = 0, \dots, 6$) is classified as one of the 16 state-to-state pairs, e.g., Low-Low, Medium-High, etc. An example of the resulting, estimated, daily probability transitions matrix for Monday to Tuesday P_0 is provided below

$$P_0 = \begin{bmatrix} 0.8205 & 0.1795 & 0 & 0 \\ 0.1538 & 0.7949 & 0.0513 & 0 \\ 0 & 0.4694 & 0.4286 & 0.102 \\ 0 & 0.04 & 0.64 & 0.32 \end{bmatrix}. \quad (1)$$

It can be seen from P_0 that the first two diagonal entries are dominant. Thus the probability of remaining in a Low or Medium state on Tuesday when the system was in such a state on Monday is clearly the highest. However, for High and Very High states, there are substantial probabilities of either moving to or remaining in the Very High state. Other interesting patterns can be detected in the remaining six transition matrices, but are not presented here due to page limitations.

2.2 Weekly Probability Transition Matrices

The weekly probability transition matrices $Q_i, i = 1, \dots, 6$ are given by products of the daily probability transition matrices, namely

$$Q_0 = P_0 P_1 \dots P_5 P_6; \quad Q_1 = P_1 P_2 \dots P_6 P_0; \quad \dots \quad Q_6 = P_6 P_0 \dots P_4 P_5, \quad (2)$$

where Q_0 contains weekly probability transitions from a Monday to the following Monday, Q_1 contains similar transitions from a Tuesday to the following Tuesday, and so on. It has been verified, empirically, that these matrices approximately satisfy the time homogeneity property, thus defining seven homogeneous Markov chains. Matrix Q_0 , calculated from the product of daily probability transition matrices for Monday-to-Monday, is given here as an example:

$$Q_0 = \begin{bmatrix} 0.2028 & 0.5485 & 0.1826 & 0.066 \\ 0.178 & 0.5246 & 0.2041 & 0.0933 \\ 0.1491 & 0.4841 & 0.2313 & 0.1355 \\ 0.124 & 0.4436 & 0.2558 & 0.1766 \end{bmatrix}. \quad (3)$$

An important observation to be made when comparing daily and weekly probability transition matrices is that the latter contain no zero entries and hence are automatically irreducible. From discussions with hospital staff we inferred that in terms of management policies, it is the long-term behaviour induced by the weekly probability transitions that could influence these policies rather than any daily transitions.

2.3 Steady State Distributions

In a finite Markov chain, two states *communicate* if it is possible to reach one from the other in a finite number of transitions, and vice versa. A set of states that communicate with one another but not with any other states is called a *communicating class*. A Markov chain is *irreducible* if it contains a single, exhaustive, communicating class. The weekly transition matrices Q_0, \dots, Q_6 corresponds to an irreducible Markov chain as none of them contain any zero entries, for instance, see Q_0 in (3). It is well known that irreducibility guarantees the existence of a unique steady state distribution for the given Markov chain. Figure 1 shows the trends of these weekly steady state probabilities for each occupancy level.

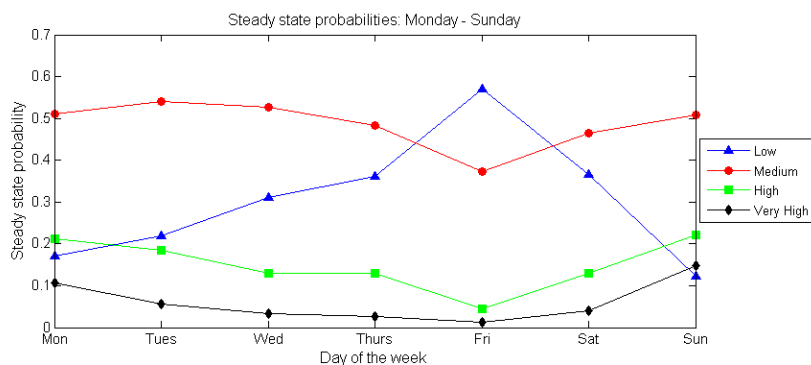


Figure 1. Comparison of steady state probabilities for each day of the week

Figure 1 shows that across the working week, the steady-state probabilities of being in a higher occupancy state decreases until the minimum is reached on Friday (see the plots labelled with squares and diamonds). This coincides with the discharge of patients for the approaching weekend. On weekends, the probability of

being in a higher state begins to increase again. Naturally, the Friday minimum for High, Very High (and also Medium) states is accompanied by a peak for the Low occupancy state. The relative stability, across the week, of the steady-state probability of being in the Medium occupancy state is also natural and reassuring in terms of FMC's mode of operations. However, the fact that on Sundays and Mondays the steady state probabilities of being in a High or Very High state exceed 0.3 (30% long-run frequency) could be seen as a cause for concern.

2.4 MDP Parameters and Formulation

The goal of congestion relief is cast in the framework of a finite horizon Markov Decision Process. Towards that goal, we define probability transitions that depend on decision makers actions and the immediate costs resulting from a choice of a given action in a given state. We adopt the notation of Filar and Vrieze (1997).

Definition 2. *In a finite Markov Decision Process, the system is in state $s \in S = \{1, 2, 3, 4\}$ at time t , and the decision maker is obliged to choose an action $a \in A(s) = \{1, 2, \dots, m_s\}$. As a result of this choice, the following probabilistic transition occurs: $p_t(s, s', a) = P(S_{t+1} = s' | S_t = s, A_t = a)$, and an immediate cost $c_t(s, a)$ is accrued.*

Note that, in some states, there could be only a single action, which often corresponds to a decision that "no action is needed". Indeed, this is the case in the Low and Medium occupancy states introduced above. To fully describe the MDP model we need to specify the set S of states, the sets $A(s)$ of actions available in each state, the time horizon T and the above probability transitions and rewards. These will be called the parameters of the MDP.

We shall use the four states, ordered with increasing levels of occupancy, given in Definition 1. In our four state occupancy model, actions will require physical interpretation that corresponds to various levels of intervention designed to lower the occupancy. It should be noted that a selection of an action may result in altered probabilities of the system being in any given state at the next stage. Of course, the external factors responsible for the stochasticity in the model ensure that the decision maker cannot guarantee what state the system will be in at the next stage. Nonetheless, the ability to influence the probability transitions can be taken advantage of, so as to increase the likelihood of passing to states with improved occupancy levels.

2.5 Actions

Many possible actions could be taken to mitigate congestion episodes in a hospital. They include discharging patients early, cancelling elective treatments or transferring some patients to other hospitals. All of these actions involve releasing some number of patients from the system, so each possible action will be defined as releasing a multiple of four patients at a time from the system up to some practical upper bound. Of course, actions could have been more finely defined had we taken multiples of 1, 2, or 3 patients. However, releasing a single patient may not have a discernible impact on occupancy levels. In order to consider only actions that have a discernible level of impact, it was decided that multiples of four will be appropriate. In this study we have decided that the set of all possible actions will be: 0: No action; 1: Release four patients; 2: Release eight patients; 3: Release 12 patients; 4: Release 16 patients; 5: Release 20 patients.

Twenty was chosen as a cap for a release of patients, as it seemed that removing any more patients than that was impractical when considered in the context of the current policies. Consultation with senior FMC staff affirmed that this cap was reasonable.

We have also decided that $A(1) = A(2) = \{0\}$, $A(3) = A(4) = \{1, 2, \dots, 5\}$. Thus, the occupancy lowering actions can be invoked only in the High and Very High states. This is reasonable because in states 1 and 2, the hospital is already operating at a manageable level.

2.6 Cost Functions

Next, the cost functions $c(s, a)$ need to be modelled. The challenge here is to capture both the benefit of reducing occupancy when it is too high and the cost resulting from inconveniencing patients and the associated harm to the reputation of the hospital. Furthermore, there is a cost involved in having a hospital that is operating too far below capacity, which must be included in the calculation. Even though these are qualitative benefits and costs, we require quantitative expressions that balance them.

Since our states are labelled 1,2,3,4 in the order of increasing occupancy - and the actions are also labelled in the order of increasing numbers of patients being released - it is natural to assume that the relief provided by

a release of a certain number of patients is more significant in the Very High state than in the High state. The cost, on the other hand, is assumed to increase with the number of patients released from the system.

In the health care industry, around 85 percent of bed occupancy is often said to be a "golden level of occupancy", which we shall denote by j . FMC currently has a total of 362 beds dedicated to General Medicine and Surgery, so the golden level of occupancy, j , would be 307. Arguably, the hospital operates efficiently when there are 307 admitted patients which, coincidentally, is the threshold between the Medium and High occupancy states 2 and 3. In our model, we decided that using a single golden level of occupancy, $j = 307$, is too restrictive. For instance, on Saturday and Sunday a lower level seems appropriate as we expect a greater influx of patients over the weekend, who need to be ready to undergo treatments starting on Monday. This, coupled with lower weekend discharge rates, increases the risk of congestion episodes in the following days.

The cost function is adjusted for the day of the week, with different cost values applying on busier or quieter days. This is achieved by modifying the value of $j = j_t$ to penalise higher and lower occupancies differently. The value for j_t is moved by approximately eight patients either side of the 85% capacity value of 307. This represents a change of approximately 2 percent of the hospital's capacity. Saturday and Sunday are classified as busy days, where fewer patients are discharged though admissions remain close to usual rates. The quiet days of the week are Thursday and Friday, where patients are discharged early to free up beds for the ensuing weekend. The actual cost functions used are given in (4) as follows:

$$c_t(s, a) = \begin{cases} \frac{9}{10} \max(0, s - 307 - a) + \frac{1}{10} \max(0, 307 - s), & \text{Mon-Wed} \\ \frac{9}{10} \max(0, s - 315 - a) + \frac{1}{10} \max(0, 315 - s), & \text{Thu-Fri} \\ \frac{9}{10} \max(0, s - 300 - a) + \frac{1}{10} \max(0, 300 - s). & \text{Sat-Sun} \end{cases} \quad (4)$$

2.7 Action-dependent Probability Transitions

The actions defined in a Markov Decision Process alter the probabilities of a transition from one state to another. There needs to be a simple way to estimate these probabilities. One method of doing this is by splitting the aggregated states into micro-states for each number of beds occupied. Let $n_{s,s'}$ be the number of observed transitions from s beds occupied on one day to s' beds occupied on the following day, and N_s be the total number of days that s beds were occupied, thus: $\hat{P}(X_{t+1} = s' | X_t = s) = \frac{n_{s,s'}}{N_s}$. This is valid without considering the effects of any actions on these probabilities. When an action a is taken, the transition probability effectively becomes a transition from $s - 4a$ beds occupied to s' beds occupied, which is estimated by: $\hat{P}(X_{t+1} = s' | X_t = s, A_t = a) = \frac{n_{s-4a,s'}}{N_{s-4a}}$, recalling that each action corresponds to a multiple of four patients being released. Using this equation it is now possible to estimate the new transition probabilities of moving between the aggregated states under a given action. By the rules of calculus of probabilities, these were found to be:

$$P(X_{t+1} \in S' | X_t \in S, A_t = a) = \sum_{s' \in S'} \sum_{s \in S} P(X_{t+1} = s' | X_t = s, A_t = a) P(X_t = s | X_t \in S). \quad (5)$$

The first factor in each of the terms of the double summation in (5) is equal to the probability of moving from each micro-state in aggregated state at time t to each micro-state in the aggregated state at time $t + 1$, conditional on the micro-state at time t being s and the action chosen being a . The second factor is the probability of the micro-state at time t being s , conditional on the macro-state being S .

The use of actions in the above formula presents a technical challenge, as the raw data matrix of observed transitions ($n_{s,s'}$) is quite sparse. This means that some rows that are needed in the computation of the probabilities are empty. To alleviate this problem, the empty rows of the matrix are replaced with estimates. We choose to replace each empty row with the closest non-empty row. All entries are rounded up to the nearest integer, as a fraction of a transition count is inappropriate. Once all empty rows are filled in, the calculations for each of the new transition probabilities are well defined. We refer the reader to Clissold (2014) for details of this calculation.

3 FINDING AN OPTIMAL POLICY FOR THE MDP

With a fully defined MDP, an optimal policy for the given process can be found. Strictly speaking, the preceding only defines a deterministic Markov policy, but these are sufficient to find optimal policies even in larger policy spaces that are not of interest here (e.g. see Derman (1970)).

In a finite time horizon setting, the standard method for calculating optimal policies in an MDP is the Dynamic Programming algorithm. This can be very time consuming due to the phenomenon known as "the curse of

dimensionality”. Fortunately, the MDP in our application is sufficiently small that the computational burden is reasonable. A time horizon of seven days is used in this analysis to replicate the week-long planning patterns at Flinders Medical Centre.

3.1 An Optimal Solution to the MDP

Running the Backward Recursion of Dynamic Programming algorithm on the Markov Decision Process defined earlier yields an optimal solution for each state and time point as follows, where the columns correspond to the seven days of the week starting on Monday:

$$f^* = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 2 & 2 & 0 & 4 & 4 & 2 \\ 5 & 5 & 5 & 5 & 5 & 5 & 5 \end{bmatrix}. \tag{6}$$

The first two states, Low and Medium, only have one possible action, hence the repetition of action zero in the first two rows. The final two rows contain the optimal actions to be taken on each day if the Markov process is in either the High or Very High states. If the MDP is in the High state on Monday through Wednesday, eight patients should be released, 16 on Friday and Saturday, and eight on Sunday. On Thursdays, no action should be taken. This is an ideal outcome, as it does not require any patients to be inconvenienced but allows for the natural operational patterns of the hospital to reduce the number of occupied beds.

If the MDP is in a state of Very High occupancy on any day of the week, then 20 patients should be released under the policy f^* . These days occur rarely, approximately once every two weeks, so the model recommends that staff act quickly to mitigate the congestion when these days do occur. The model also recommends that staff take the most drastic action on days of Very High occupancy, allowing the hospital to minimise the flow-on effects of these Very High occupancy days.

4 DISCUSSION AND CONCLUSIONS

Now that an optimal policy has been found, the improvement between the current situation and this optimal policy can be quantified. As the time horizon considered was a calendar week, the weekly steady state distributions of Q_0, \dots, Q_6 under the optimal policy f^* will be compared with the corresponding distributions under the existing mode of operations. The comparison of steady state probabilities is shown in Figure 2 below. In the figure, each of the four panels corresponds to the four aggregate states Low, Medium, High and Very High. The dot (red) curve plots the steady-state probabilities of being in the fixed aggregate state on each day of the week under the optimal policy and the square (blue) curve plots the corresponding probabilities under the current mode of operations.

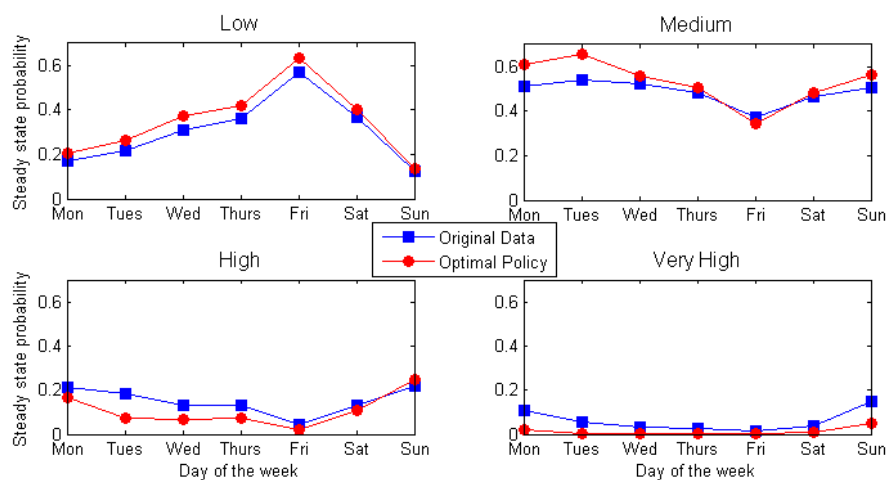


Figure 2. Comparison of steady state distributions for weekly probability transition matrices under current conditions and the optimal policy

From Figure 2 it can be seen that the optimal policy makes a significant difference in the steady state probabilities for each day of the week. There are increases in the probabilities of being in the two lower occupancy states (Low and Medium) for nearly every day of the week, demonstrated by the fact that the dot curve lies above the square curve. There are also analogous decreases of the probabilities of being in the upper capacity states (High and Very High) for each day of the week, demonstrated by the dot curve lying below the square curve in the bottom two panels. Note the significant improvement in the proportion of days that are considered to have Very High occupancy; it drops to one percent or below for each day between Tuesday and Saturday inclusive. Sunday and Monday are much larger than one percent, but are still lower compared to the current situation by approximately one-third and two-thirds respectively. The decrease in most of the High steady state probabilities is also a desirable outcome.

Figure 2 also indicates that there is a slight increase in the probability of High state on Sunday when compared to the current policy and a slight decrease in that of Medium state on Friday. This seems counterintuitive, but those anomalies are offset by corresponding changes in the probabilities of being in the adjacent states. For instance, the decrease in the steady-state probability of the process being in the Medium state on Friday is offset by an increase of being in the Low state. On Sundays, the steady-state probability of being in the High state increases slightly but the probability of being in the Very High state drops by approximately one-third. This is an advantageous outcome as it is the days of Very High occupancy that are the worst in terms of contributing to congestion episodes.

The validation of an MDP model such as the one presented here can be addressed in multiple ways. Consultation with industry experts has already provided a degree of qualitative validation. However, short of carrying out a live experiment in implementing a new discharge policy, the most practical validation strategy is to use synthetic data from a simulation model. Indeed, for FMC a hospital event simulation model: arrival to discharge (HESMAD) was developed in parallel with the MDP model. The latter confirmed that the optimal policy from the MDP model generates occupancy data that are in concordance with forecasts such as those described in Figure 2. Finally, the MDP methodology outlined here is generic in the sense that it could be adapted to policies of another hospital. Of course, the parameters would have to be re-estimated and re-calibrated based on the data of that hospital.

ACKNOWLEDGEMENT

This work was supported by the ARC linkage grant LP130100323, Congestion Recovery and Optimisation of Patient Flows, jointly awarded to Flinders University, the Southern Adelaide Health Service (Flinders Medical Centre) and the Central Adelaide Local Health Network (Royal Adelaide Hospital). The authors also acknowledge useful discussions with Dr M. Mackay.

REFERENCES

- Broyles, J. R., J. K. Cochran, and D. C. Montgomery (2010). A statistical markov chain approximation of transient hospital inpatient inventory. *Artificial Intelligence in Medicine* 207(3), 1645–1657.
- Christodoulou, G. and G. J. Taylor (2001). Using a continuous time markov process, with covariates, to model bed occupancy of people over 65 years of age. *Health Care Management Science* 4, 21–24.
- Clissold, A. (2014). *Modelling of Patient Flows by Markov Decision Processes (Hons thesis)*. Flinders University.
- Derman, C. (1970). *Finite State Markovian Decision Processes*. Orlando, FL, USA: Academic Press, Inc.
- Filar, J. and K. Vrieze (1997). *Competitive Markov Decision Processes*. New York: Springer.
- FitzGerald, G., S. Toloo, J. Rego, J. Ting, P. Aitken, and V. Tippett (2012). Demand for public hospital emergency department services in australia: 2000–2001 to 2009–2010. *Emergency Medicine Australasia* 24(1), 72–78.
- Garg, L., S. I. McClean, M. Barton, B. J. Meenan, and K. Fullerton (2012). Intelligent patient management and resource planning for complex, heterogeneous, and stochastic healthcare systems. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 42(6), 1332–1345.
- Helm, J. E., S. AhmadBeygi, and M. P. V. Oyen (2011). Design and analysis of hospital admission control for operational effectiveness. *Productions and Operations Management* 20(3), 359–374.