Inventory control for point-of-use locations in hospitals

M Bijvank[1][*][†] and IFA Vis[2]

[1]*Université de Montréal, Montréal, Québec, Canada; and* [2]*University of Groningen, Groningen, The Netherlands*

Most inventory management systems at hospital departments are characterised by lost sales, periodic reviews with short lead times, and limited storage capacity. We develop two types of exact models that deal with all these characteristics. In a capacity model, the service level is maximised subject to a capacity restriction, and in a service model the required capacity is minimised subject to a service level restriction. We also formulate approximation models applicable for any lost-sales inventory system (cost objective, no lead time restrictions etc). For the capacity model, we develop a simple inventory rule to set the reorder levels and order quantities. Numerical results for this inventory rule show an average deviation of 1% from the optimal service levels. We also embed the single-item models in a multi-item system. Furthermore, we compare the performance of fixed order size replenishment policies and (R, s, S) policies.

1. Introduction

The main objective of a hospital is to provide high-quality health care. Sufficient medical items need to be on hand to enable hospital staff to perform their daily work. Typically, medical supplies are stored at many locations in a hospital and in large quantities to prevent stock outs as much as possible. However, hospitals have lack of available storage space (Lapierre and Ruiz, 2007; Little and Coughlan, 2008) and millions of dollars are tied up in inventories that consume on average 20% of net patients revenues and represent the second largest expense after labour (Moon, 2004). Therefore, the available storage capacity should be used efficiently, and it is important to find a balance of the service quality and the desired inventory levels between the different items. The goal of this paper is to develop inventory models which consider the service level as well as the capacity limitations at hospitals. Our main contribution is the development of new solution techniques to support hospital management decisions with regard to inventory control.

Two types of models are proposed to deal with the capacity limitations and the service requirements: a capacity model and a service model. In a *capacity model*, the objective is to maximise the service level when the storage capacity is limited. This model can be used in a multi-item inventory control system with a storage capacity constraint. In a *service model*, the required capacity is minimised while satisfying a service constraint instead of having a cost objective. The service model results in the lowest inventory levels within a service level restriction. Consequently, the model will minimise inventory holding costs. We will show that our approximation procedures can be used to minimise costs in Section 8. For a more detailed discussion on models with a service constraint and models with a cost objective we refer to van Houtum and Zijm (2000).

In Section 2 we derive the specific characteristics of the inventory system under study. In Section 3 we state the added value of this research by providing an overview on the literature for hospital and general inventory management. We develop a model for hospital inventory systems in Section 4. Moreover, we present an approach to use this single-item capacity model in a multi-item inventory control system to assign items to the total capacity available at hospitals. In Section 5 we demonstrate and compare the performance of different replenishment policies based on data from a practical case study. Simple and effective closed-form expressions are derived in Section 6 for several performance measures to set the

[†]Research performed at Department of Mathematics, Faculty of Exact Sciences and Department of Information Systems and Logistics, Faculty of Economics and Business Administration, VU University Amsterdam, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands.
[*]*Correspondence: M Bijvank, Department of Computer Science and Operations Research (DIRO), Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal, Québec, H3C 3J7, Canada.*
E-mail: bijvankm@iro.umontreal.ca

inventory control variables of the replenishment policy. Furthermore, we develop a new heuristic approach in Section 7 to derive near-optimal solutions for the capacity model indicating the reorder level and order quantity. Numerical experiments are performed in Section 8 to show the general applicability and to test the performance of the different procedures and replenishment policies for a wide range of settings. Section 9 presents our conclusions.

2. POU inventory systems in hospitals

We roughly distinguish three types of inventories in hospitals, namely perishable items including medicines and blood (see, eg, Prastacos, 1984; Katsaliaki and Brailsford, 2007), non-disposables (eg, instruments) and disposables (eg, gloves, needles, sutures). Our main focus in this research is on disposable items since this type of products is stored in almost all locations in hospitals and, therefore, more difficult to control.

Traditionally, health-care supply chains are characterised by a multitude of different suppliers, products and patient care units that arbitrarily order multiple items (Rivard-Royer *et al*, 2002). A hospital storage room receives these items and distributes them to the right lower-level *point-of-use* (POU) locations, such as nursing units and operating rooms (see Figure 1a). Another option is to outsource the replenishment activities and make suppliers deliver directly to the POU locations (see Figure 1b). Such stockless or just-in-time (JIT) inventory systems are described by, for example, Nathan and Trinkaus (1996) and Danas *et al* (2002). The cost and service level differences between these two systems are quantified by Nicholson *et al* (2004).

In the literature two concepts are proposed to improve the performance of health-care supply chains: product standardisation and selecting prime vendors (see, eg, Johnston, 1992). The former concept reduces the number of different items that have to be stocked. Consequently, it reduces the volatility of the demand. The latter concept can reduce lead times and, therefore, the safety stock as

well. Besides these tactical and strategical decisions, it is also essential to decide on the time instant to place a new replenishment order (ie, the reorder level) and the corresponding order size for each item on stock at a single POU location. Such operational decisions are the focus in this paper.

Inventory management of disposable items at each POU location can be described as a system where all items are stored in bins. Each bin has a total storage capacity of $C(i)$ units for item i that is used to fulfil demand of medical staff whenever required. If s or less units of an item are available in the bin a signal is given (eg, a nurse puts a bar code of the item on an ordering board). This level s is called the *reorder level*. These signals are scanned at prespecified time intervals of length R (ie, *periodic reviews*) that may range from days to weeks. Items are usually ordered in *fixed quantities* of size Q to provide a transparent and easy-to-understand inventory policy for hospital staff. After L time units the ordered items are resupplied from the higher-level stock point to the specific bins (see Figure 1). In a hospital setting the lead time L is known and relatively short due to the high product availability at nearby higher-level stock points. Therefore, another characteristic for the inventory system at hospitals is that the lead time L is shorter than the length of the review period R (see also Duclos, 1993). This is referred to as *fractional lead times*.

If a required item is not available in the right quantity at a specific POU location to provide the required health-care service, a substitute product is used or an emergency delivery is performed (eg, from another POU location). The original demand for the item is considered to be *lost*. Such situations are time costly and should be avoided as much as possible. Therefore, we define the service level as the fraction of demand to be satisfied directly from stock on hand (ie, *item fill rate*). Note that this definition does not include the fraction of demand that is satisfied due to a substitution or an emergency delivery in case of a stock out.

To summarise, the inventory system in most hospitals is characterised by periodic reviews, an (R, s, Q) replenishment policy, short lead times, lost sales, capacity restrictions and a service level objective. Besides this replenishment policy, order-up-to policies are also commonly observed at hospitals (see Table 1). In this policy the inventory position (inventory on hand plus inventory on order minus back-orders) is raised to an order-up-to level S at each review instant. This policy is denoted as an (R, S) policy. More generally, in an (R, s, S) policy the inventory position is increasing to level S at a review instant when the inventory position is at or below reorder level s. When $s = S-1$, this policy corresponds to the order-up-to policy. Both policies are considered in this paper as alternative replenishment policy next to the (R, s, Q) policy. The advantage of this
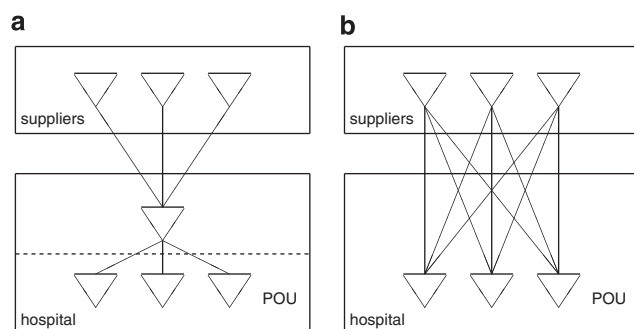

Figure 1 A traditional inventory system *versus* a JIT stockless system: (a) traditional inventory system; (b) JIT stockless system.

Table 1 The characteristics of the relevant literature discussed in this paper: deterministic (det) or stochastic (stoch) demand, excess demand is backordered (back) or lost, lead time is an integral multiple of the review period ($L = nR$) or fractional ($L \leqslant R$), periodic (P) or continuous (C) review, replenishment policy, number of items, capacity limitation, and the objective function

	Demand process	Excess demand	Lead time	Review	Policy	Nr. of items	Capacity	Objective
Inventory management for hospitals								
Vincent and Ranton (1984)	det	back	$L = nR$	P	(R, s, Q)	single	yes	cost
Nicholson *et al* (2004)	stoch	back	—	P	(R, S)	single	no	cost
Lapierre and Ruiz (2007)	det	back	$L = nR$	P	(R, s, Q)	multi	yes	cost
Dellaert and Van de Poel (1996)	stoch	back	$L = 0$	P	(R, s, Q)	multi	no	cost
Little and Coughlan (2008)	stoch	back	$L = 0$	P	(R, S)	multi	yes	service
General inventory management								
Janssen *et al* (1998)	stoch	back	$L = nR$	P	(R, s, Q)	single	no	service
Hadley and Whitin (1963)	stoch	lost	$L = nR$	C	(s, Q)	single	no	cost
Hill (1994)	stoch	lost	$L = nR$	C	(s, Q)	single	no	cost
Johansen and Thorstenson (1996, 2004)	stoch	lost	$L = nR$	C	(s, Q)	single	no	cost
Hill (1999)	stoch	lost	$L = nR$	P	(R, S)	single	no	cost
Bijvank and Johansen (2009)	stoch	lost	—	P	(R, S)	single	no	cost
Downs *et al* (2001)	stoch	lost	$L = nR$	P	(R, S)	multi	yes	cost
Hill and Johansen (2006)	stoch	lost	$L = nR$	P	(R, s, S)	single	no	cost
Johansen and Hill (2000)	stoch	lost	$L = nR$	P	(R, s, Q)	single	no	cost
Chiang (2006)	stoch	lost	$L \leqslant R$	P	(R, S)	single	no	cost
Chiang (2007)	stoch	lost	$L \leqslant R$	P	(R, s, S)	single	no	cost
van Donselaar *et al* (1996)	stoch	lost	$L = nR$	P	(R, S)	single	no	service
Bijvank and Vis (2010)	stoch	lost	$L = nR$	P	(R, s, S)	single	no	service
Tijms and Groenevelt (1984)	stoch	lost	—	P	(R, s, S)	single	no	service
Kapalka *et al* (1999)	stoch	lost	$L \leqslant R$	P	(R, s, S)	single	no	service
Sezen (2006)	stoch	lost	$L \leqslant R$	P	(R, S)	single	no	service
This paper	stoch	lost	$L \leqslant R$	P	(R, s, Q)	multi	yes	service

latter policy is its simplicity to understand and implement in practice since order sizes are fixed. However, it does not use the available capacity optimally. The available capacity is utilised more efficiently when the replenishment policy incoorporates an order-up-to level that equals the available capacity. Although such policies require more effort to process the replenishments, we show the benefits in terms of service levels and capacity utilisation compared to fixed order size policies in our case study (Section 5) and our numerical experiments (Section 8).

3. Literature overview

Hardly any literature is available on replenishment policies for inventory systems in a *hospital setting*. The literature overview on operation research in health-care settings by Brandeau *et al* (2004) does not even mention inventory theory. Only a few papers are available about inventory control in a hospital setting. Lapierre and Ruiz (2007) solve a multi-item inventory replenishment problem with storage and manpower capacity restrictions. In their research, demand is assumed to be deterministic and known. They formulate a non-linear mixed-integer problem and solve this with a tabu search metaheuristic. Similar restrictions

are considered by Little and Coughlan (2008). The authors propose an optimisation model based on constraint programming to determine the delivery frequency and order sizes. A single-item inventory problem with capacity restrictions is discussed by Vincent and Ranton (1984) for a hospital environment. They extend the basic EOQ formula with capacity restrictions and focus on a cost objective instead of a service requirement. Order costs are also the main focus in Dellaert and Van de Poel (1996), where joint replenishments at a central storage room in a hospital result in cost savings.

Most of the literature on inventory theory cannot be used in a hospital context, since they are more focused on backorder models contrary to lost-sales models (see, eg, Zipkin, 2000). Janssen *et al* (1998) discuss an (R, s, Q) model with *backorders*. They develop a closed-form expression to calculate the fill rate for given values of R, s and Q. The authors propose a local search procedure to determine the best value of reorder level s subject to a service level constraint. However, we deal with a lost-sales inventory system in this paper. Much less literature is available about such systems, because of the complexity to model the inventory position.

The first study of the (s, Q) policy in a lost-sales context is considered by Hadley and Whitin (1963) for *continuous*

reviews. The majority of the papers that discuss a continuous review lost-sales inventory system extend these preliminary results. A commonly made assumption in literature is to restrict the number of orders outstanding. The restriction $Q > s$ specifies that at most one order may be outstanding at any time (eg, Johansen and Thorstenson, 1996). Hill (1994) analyses a similar inventory model where two orders may be outstanding at any time (ie, $Q \leqslant s < 2Q$). Johansen and Thorstenson (2004) are the only authors who do not consider a restriction on the number of outstanding orders. Almost all papers dealing with (s, Q) policies in a continuous review model with lost sales derive steady-state probabilities to compute the expected total costs and propose a policy-iteration algorithm (PIA) to find near-optimal values of the reorder level s and order quantity Q based on a cost objective. A more detailed overview on lost-sales models with continuous reviews can be found in Bijvank and Vis (2011). Lost-sales models with a fixed order size policy and periodic reviews have received less attention. The (R, s, Q) policy is only considered under the assumption that there is never more than one order outstanding (ie, $Q > s$). Johansen and Hill (2000) propose a PIA to set the values of s and Q for a model with a cost objective.

Order-up-to policies have received more attention for lost-sales inventory systems with periodic reviews. Based on numerical examples it is shown that such policies are not optimal for models with a cost objective in a lost-sales setting (see, eg, Hill, 1999; Hill and Johansen, 2006). Fractional lead times in a lost-sales inventory system with a cost objective are considered by Chiang (2006, 2007). For more information on the cost model, we refer to Hill and Johansen (2006), Bijvank and Johansen (2009), Bijvank and Vis (2011) and the references therein.

When there is a *service level restriction*, Tijms and Groenevelt (1984) propose a procedure for the (R, s, S) policy to determine the value of reorder level s when the order-up-to level S is known. However, the authors briefly mention a lost-sales setting, and they discuss numerical results for only a backorder model. Bijvank and Vis (2010) compare the (R, s, S) policy to an optimal replenishment policy in case of lost sales for a single-item inventory system with no consideration of capacity. For the (R, S) policy, van Donselaar *et al* (1996) determine the order-up-to level dynamically for each review period with a myopic approach based on a service level constraint. Consequently, the order-up-to level varies over time. This is not preferred at hospitals (see also Section 2). Sezen (2006) uses simulation to study the service level performance for an order-up-to policy. The author varies the length of the review interval to consider the impact on the service level. However, no solution method is proposed to find optimal order-up-to levels but only insights are provided. Fractional lead times and a service constraint are studied by Kapalka *et al* (1999) for lost-sales inventory

systems with an (R, s, S) policy. A summary of the aforementioned references is provided in Table 1.

With respect to the existing literature on lost-sales models we add four aspects. First, we consider capacity constraints and develop a simple inventory rule to set the inventory control variables for the capacity model that can be implemented in a spreadsheet-based program like Excel. Second, we embed the capacity model in a multi-item inventory system with lost sales. Most models for multi-item inventory systems consider cost savings when replenishments are coordinated over multiple items instead of capacity constraints (see, for instance, Silver *et al* 1998; Axsäter, 2006). Third, we compare the performance of the (R, s, Q) policy to the (R, s, S) policy in case there is a capacity restriction. Fourth, we derive closed-form expressions to approximate performance measures (eg, fill rate, average costs) for lost-sales inventory systems with either an (R, s, Q) or an (R, s, S) replenishment policy. These expressions can be used in more general settings besides the hospital setting described in Section 2, since there is no assumption on the lead time or on the available capacity.

4. Models

In this section we develop a service model and a capacity model for the inventory control system at a POU location as described in Section 2. Therefore, we decompose the inventory system in single-item models and we embed these models in a multi-item inventory system with a storage capacity constraint. This multi-item model can be used to determine the available capacity for each item in the capacity model.

4.1. Single-item inventory system

We derive a single-item capacity model and a service model for the (R, s, Q) policy with lost sales, fractional lead times and a service level constraint. Kapalka *et al* (1999) propose a Markov decision model to determine the service level based on the steady-state distribution of the on-hand inventory at a review instant for an (R, s, S) replenishment policy. We use a similar approach in this section to develop our model.

The demand during t time units is modelled as a discrete random variable D_t, which is assumed to be independent for non-overlapping time intervals. The probability distribution function is given by $g_t(d)$. Furthermore,

$$\mathcal{G}_t^0(d) = \sum_{i=0}^{d-1} g_t(i) = P(D_t < d),$$

$$\mathcal{G}_t^1(d) = \sum_{i=1}^{d} \mathcal{G}_t^0(i) = E[(d - D_t)^+],$$

where $(A)^+ = \max\{A, 0\}$. Define X_{n+1} as the on-hand inventory level at the beginning of review $n+1$. The (R, s, Q) policy prescribes to order Q units when the inventory position is at or below reorder level s at a review. The order is delivered after L time units but within the same review period (ie, $L \leqslant R$). Hence,

$$X_{n+1} = \begin{cases} (X_n - D_R)^+, & \text{if } X_n > s, \\ ((X_n - D_L)^+ + Q - D_{R-L})^+, & \text{if } X_n \leqslant s. \end{cases}$$

The random variable X_{n+1} only depends on X_n and the demand during one review period. Thus $X = \{X_n, n \geqslant 0\}$ is a homogeneous, one-dimensional Markov chain with state space $\{0, 1, \ldots, s+Q\}$.

In order to define the transition probabilities $P_{ij} = P(X_{n+1} = j | X_n = i)$, we make a distinction between $X_n \leqslant s$ and $X_n > s$. When $X_n \leqslant s$,

$$P_{ij} = \begin{cases} \sum_{d=0}^{i-1} g_L(d)(1 - \mathcal{G}_{R-L}^0(i - d + Q)) + (1 - \mathcal{G}_L^0(i))(1 - \mathcal{G}_{R-L}^0(Q)), & j = 0, \\ \sum_{d=0}^{i-1} g_L(d) g_{R-L}(i - d + Q - j) + (1 - \mathcal{G}_L^0(i)) g_{R-L}(Q - j), & 0 < j \leqslant Q, \\ g_R(i + Q - j), & j > Q. \end{cases} \tag{1}$$

If $j > Q$ the inventory position can never drop to zero during one review period and therefore P_{ij} is constructed by convolution of D_L and D_{R-L} to D_R. When $X_n > s$,

$$P_{ij} = \begin{cases} 1 - \mathcal{G}_R^0(i), & j = 0, \\ g_R(i - j), & 0 < j \leqslant i, \\ 0, & j > i. \end{cases} \tag{2}$$

This defines the transition matrix P of Markov chain X. The Markov chain X is irreducible and aperiodic since all states communicate. It has a unique stationary distribution $\pi = (\pi(0), \pi(1), \ldots, \pi(s+Q))$, where $\pi(j)$ can be interpreted as the limiting probability that the process is in state j at a review. The stationary probabilities π are given by the solution of

$$\pi(j) = \sum_{i=0}^{s+Q} \pi(i) P_{ij}, \qquad \text{for } 0 \leqslant j \leqslant s + Q$$

$$\sum_{j=0}^{s+Q} \pi(j) = 1.$$

We define $\beta(i)$ as the fraction of demand satisfied in a review period, when i units are on hand at the beginning of the review period. Demand that can be fulfilled is equal to the minimum of the number of units on stock and the actual demand. Hence, when $i \leqslant s$,

$$\beta(i) = 1 - \frac{E[(D_L - i)^+ + (D_{R-L} - ((i - D_L)^+ + Q))^+]}{E[D_R]}$$

$$= \frac{i + Q - \sum_{d=0}^{i-1} g_L(d) \mathcal{G}_{R-L}^1(i - d + Q) - (1 - \mathcal{G}_L^0(i)) \mathcal{G}_{R-L}^1(Q)}{E[D_R]},$$

$$\tag{3}$$

and when $i > s$,

$$\beta(i) = 1 - \frac{E[(D_R - i)^+]}{E[D_R]} = \frac{i - \mathcal{G}_R^1(i)}{E[D_R]}.$$

The average fill rate is denoted by $\beta = \sum_i \pi(i) \beta(i)$. The same analysis can be performed for an (R, s, S) policy. Since the order size in this policy is not fixed, the value of Q has to be replaced by $S - i$ in Equation (1) to Equation (3). Kapalka *et al* (1999) perform a similar analysis for the (R, s, S) policy.

These expressions can be used to formulate the *capacity model* as

$$\text{maximise } \{\beta | s + Q \leqslant C\}, \tag{4}$$

and the *service model* as

$$\text{minimise } \{s + Q | \beta \geqslant \overline{\beta}\}, \tag{5}$$

where $\overline{\beta}$ is a minimum service level. For the (R, s, S) policy, replace $s + Q$ by S. Note that the restriction in the capacity model can be replaced by $s + Q = C$ to make full use of the available capacity. Consequently, the search space of possible solutions for s and Q is a bounded one-dimensional vector (ie, $s \in \{0, 1, \ldots, C\}$). An enumeration procedure over all $C + 1$ values of reorder level s is required to maximise the service level. However, the service model results in an unbounded two-dimensional solution space. How to determine the value of the available capacity C in the capacity model is discussed in the remainder of this section.

4.2. Multi-item inventory system

The inventory control problem in hospitals is more complex than the single-item system considered so far. The capacity limitation for each item is part of a larger inventory system with multiple items, where the limited

capacity is shared by all items stored at a POU location. In this section we use a decomposition method to embed the single-item models in a multi-item inventory system. Let TC denote the total capacity available for all items at a POU location, and $TC(k)$ is the capacity assigned to item k. Hence, $\sum_k TC(k) \leqslant TC$. The capacity is expressed in terms of volume, whereas inventory decisions are expressed in terms of number of units. Items are stored in bins, where each bin for item k uses a fixed storage capacity $BC(k)$ and it can contain at most $C(k)$ units. In the multi-item inventory system, the number of bins assigned to each item has to be determined such that the average service level is maximised within the capacity limitation. When we denote this number by $a(k)$ for item k, then $TC(k) = a(k)BC(k)$ and at most $a(k)C(k)$ units can be stored for item k (this corresponds to the available capacity C in the capacity model of Section 4.1). The average service level of the multi-item inventory system is defined as the demand-weighted average service level,

$$\beta_{total} = \sum_k \frac{E[D^k]}{\sum_l E[D^l]} \beta_k,$$

where $E[D^k]$ is the average demand for item k and β_k is the expected service level for item k. Note that the service level β_k for item k depends on the replenishment policy and the corresponding values of the inventory control variables (see Section 4.1).

The allocation of the limited storage capacity available and the determination of the values for the inventory control variables can be solved simultaneously. We propose a knapsack kind of approach in which a trade-off has to be made between the increase of the service level for an item and a decrease of the remaining capacity available for the other items. The ratio of this service level increment divided by the extra assigned capacity to this item is computed for each item in every iteration. We assign an extra bin to the item with the highest ratio until all capacity is assigned. To compute the increase in the service level, we determine the optimal control values of the replenishment policy and the corresponding service level as discussed in Section 4.1. The solution procedure for the allocation of limited storage capacity among items in a multi-item inventory system is summarised in Algorithm 1. This procedure can be repeated for each POU location. A similar approach can also be used by a hospital manager to decide how much capacity is required and which items should be stored at each of the POU locations based on a service level constraint similar to our service model. To perform such a procedure, line 2 in Algorithm 1 has to be altered to make sure that a predefined service level is satisfied and the lines 4, 7–9 should be removed since there is no capacity constraint TC.

Algorithm 1: Storage capacity allocation

```
1    set a(k) = 0 for all items k;
2    while ∑_k TC(k) ≤ TC do
3        for each item l do
4            if ∑_k TC(k) + BC(l) ≤ TC then
5                - determine optimal values of the inven-
                   tory control variables for the capacity
                   model of item l where C = (a(l) + 1)C(l)
                   (see Section 4.1);
6                - determine the new corresponding
                   service level β'_l;
7            else
8                β'_l = 0;
9            end
10           Δβ_l = (β'_l − β_l)E[D^l]/∑_k E[D^k];
11           ΔC_l = BC(l);
12           Δ_l = Δβ_l/ΔC_l;
13       end
14       l* = argmax{Δ_l}, a(l*) = a(l*) + 1;
15   end
```

5. Hospital sample data

In this section we illustrate the performance of the (R, s, Q) policy and the (R, s, S) policy based on the models developed in Section 4.1. The main goal of this section is to apply the capacity model and the service model in practice and to demonstrate how the models can be used to increase the service level.

We observed the (R, s, Q) replenishment policy at the VU University Medical Centre (VUmc) in Amsterdam and at Hospital Amstelland in Amstelveen. In particular, we consider many items and report in detail on a representative example about infusion liquids at three POU locations (paediatrics, intensive care and obstetrics).

Our models can be used for any demand distribution. Literature (eg, Duclos, 1993; Epstein and Dexter, 2000) and data from both hospitals (see Figure 2) support the assumption that demand in a hospital environment follows a Poisson distribution with an average of μ_t over t time units. The specific parameter values for infusion liquids at the POU locations in Hospital Amstelland are indicated in Table 2. The final column represents the current fill rate including emergency replenishments. Since this is not preferred in practice, we do not include emergency replenishments in our models.

First, we compare the performance of the (R, s, Q) policy with $Q = C - s$ to the (R, s, S) policy with $S = C$ by considering the fill rate β and the average number of review periods between two consecutive orders $OF = 1/\sum_{i=0}^{s} \pi(i)$ (ie, the inverse of the order frequency). Both performance measures are illustrated in Figure 3 for $s = 0, 1, \ldots, C$. Clearly, the service level for the (R, s, Q)
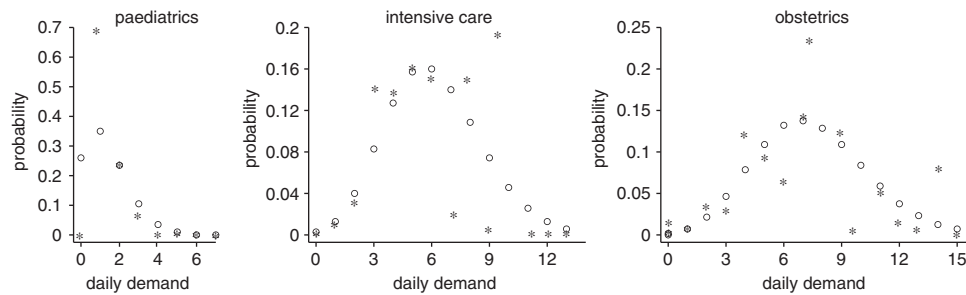
Figure 2 The demand distribution according to the empirical distribution (asterisk) and a Poisson estimation (circle) for three POU locations.

Table 2 The parameter values corresponding to the current situation for infusion liquid at different POU locations in Hospital Amstelland

POU location	L	R	μ_L	μ_R	C	β (%)
Paediatrics	4 h	3 days	0.2	4.1	5	89.9
Intensive care	4 h	3 days	1.0	18.4	40	97.7
Obstetrics	4 h	1 week	1.4	58.9	100	87.8

policy is not a convex function in the reorder level s. Consequently, exhaustive search procedures like enumeration are required to find an optimal value of reorder level s in the capacity model. Let s^* denote this value for the (R, s, Q) policy, whereas $s = C-1$ and $S = C$ are optimal for the (R, s, S) policy (ie, the (R, S) policy with $S = C$). For this latter policy, the order frequency is also the highest (see Figure 3), which results in very frequent deliveries of small orders. This is not practical for hospitals. Therefore, we also derived the value of reorder level s for the (R, s, S) policy with a similar order frequency compared to the best (R, s, Q) policy, denoted by $\bar{s}$. From the results shown in Table 3 we conclude that (R, s, S) policies can improve the service level significantly without an increase of the order frequency.

Based on the results of our experiments we conclude that the (R, s, Q) replenishment policy results in service levels of about 70% up to 98% (without emergency replenishments). The capacity C is insufficient and results in stock outs. There are several solutions to minimise stock-out occurrences:

1. use the alternative (R, s, S) replenishment policy with the same order frequency,
2. shorten the length of the review period,
3. increase the available capacity.

The first solution is illustrated in Table 3. When the length of the review period is shortened by one day, the corresponding service levels are illustrated in Table 4. However, the replenishment process would take more time with such a solution approach. We recommend a

restructuring of the available storage capacity by using the solution procedure for multi-item inventory systems (see Section 4.2). We also solved the service model for the (R, s, Q) policy with a minimal service level of 95 and 98% for infusion liquid at the three POU locations to determine the required capacity. The results are shown in Table 4.

6. Approximation procedure

In this section we propose an approximation procedure in which closed-form expressions are derived for performance measures of interest. Such expressions are much easier to implement in practical applications compared to our Markov models, and large computation times are avoided to set the inventory control variables. Moreover, the expressions can be used for any lost-sales inventory system (even when there is a cost objective). No assumptions are imposed on the lead time or the available capacity. However, we make the assumption that at most one order is outstanding at any time. Such assumptions are common in the literature for lost-sales inventory models (see Section 3). This assumption is always satisfied when the lead time is fractional, otherwise we impose the restriction $Q > s$ or $S - s > s$ on the (R, s, Q) or (R, s, S) policy, respectively. Furthermore, our approximation procedure can be used for different types of replenishment policies.

The approximation procedure is based on the average performance during a replenishment cycle (see Figure 4), whereas our exact models are based on the average performance during a review period. A *replenishment cycle* is the time between two consecutive orders (either order placement or order delivery). The time period from when the inventory position reaches the reorder level to the actual order delivery is called the *risk period*. It consists of the waiting time until the next review plus the lead time. Furthermore, the *undershoot* is defined as the satisfied demand during this waiting time (ie, it represents the difference between the reorder level and the inventory position at the order placement).

The concept of undershoot has been studied by Tijms (1994) and Baganha *et al* (1996) for inventory systems with
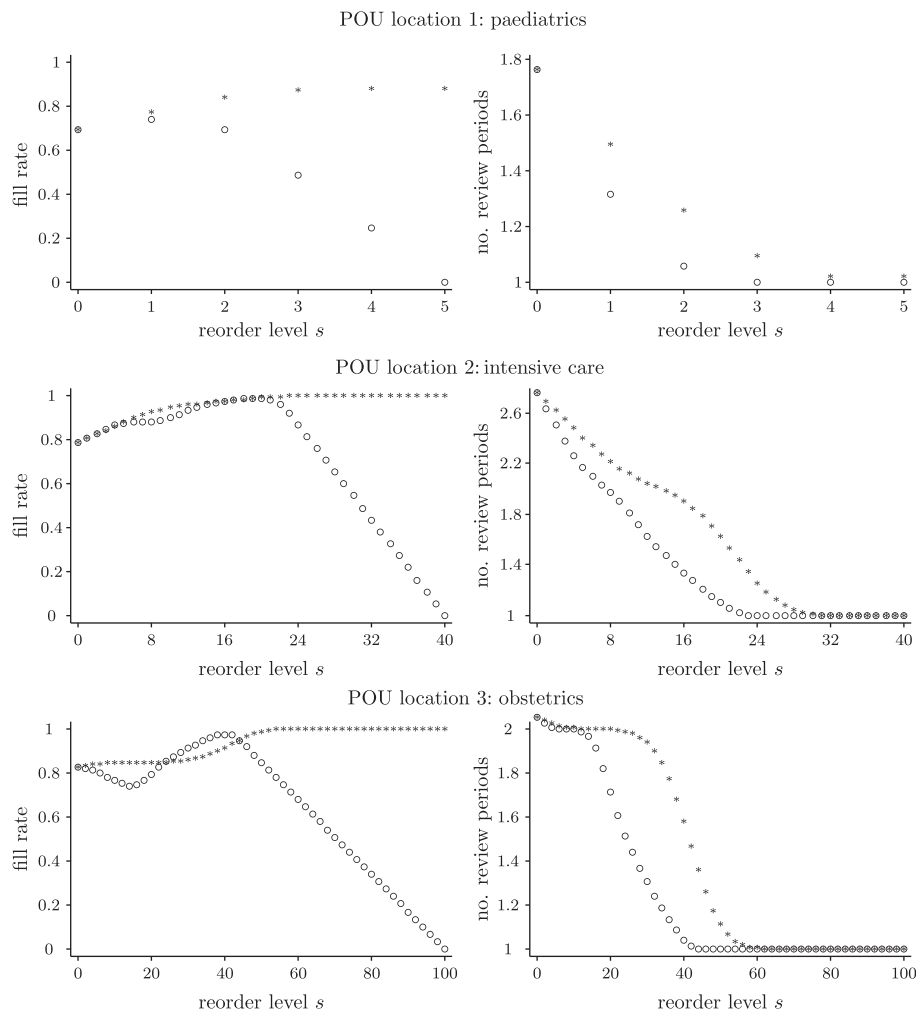
Figure 3 The fill rate and the average number of review periods between two subsequent orders for the (R, s, Q) policy (circle) and the (R, s, S) policy (asterisk) for three different POU locations.

Table 3 The fill rate β and order frequency OF for the optimal (R, s, Q) policy (denoted by s^*) and the (R, s, S) policy with a similar order frequency (denoted by $\bar{s}$)

POU location	s^*	$\beta_{(R,s,Q)}(s^*)$ (%)	$OF_{(R,s,Q)}(s^*)$	$\bar{s}$	$\beta_{(R,s,S)}(\bar{s})$ (%)	$OF_{(R,s,S)}(\bar{s})$
Paediatrics	1	74.2	1.32	2	83.9	1.26
Intensive care	19	98.7	1.16	25	99.9	1.18
Obstetrics	40	97.7	1.04	53	99.6	1.05

Table 4 The results for the capacity model when the review period length is shortened by one day, and the results for the service model

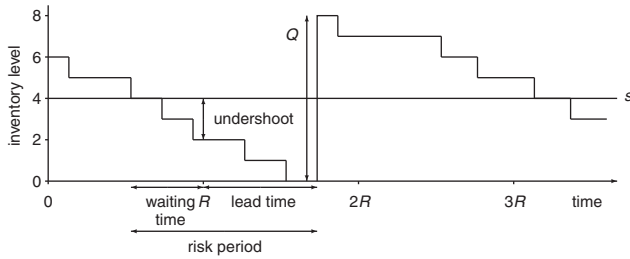| POU location	Shorten R with 1 day			$\bar{\beta} = 95\%$			$\bar{\beta} = 98\%$		
	s^*	$\beta_{(R,s,Q)}(s^*)$ (%)	$OF_{(R,s,Q)}(s^*)$	s^*	Q^*	C^*	s^*	Q^*	C^*
Paediatrics	1	85.8	1.28	5	5	10	6	6	12
Intensive care	19	99.9	1.39	14	19	33	18	20	38
Obstetrics	40	99.1	1.06	26	58	84	43	60	103

Figure 4 Illustration of the definitions during a risk period.

backorders. The authors derived approximations for the average undershoot,

$$\mu_u = \frac{E[D_R^2]}{2E[D_R]} - \frac{1}{2}, \tag{6}$$

and the variance of the undershoot,

$$\sigma_u^2 = \frac{E[D_R^3]}{3E[D_R]} - \left(\frac{E[D_R^2]}{2E[D_R]}\right)^2 - \frac{1}{12}. \tag{7}$$

We approximate the demand during the risk period with a normal distribution with mean $\mu = \mu_u + \mu_L$ and variance $\sigma^2 = \sigma_u^2 + \sigma_L^2$. Consequently, we formulate the approximate expected lost sales as

$$
\begin{aligned}
ELS &= \int_s^\infty (x - s)f(x)\mathrm{d}x \\
&= \sigma\left[\frac{\mu - s}{\sigma}\left[1 - \Phi\left(\frac{s - \mu}{\sigma}\right)\right] + \phi\left(\frac{s - \mu}{\sigma}\right)\right], \quad (8)
\end{aligned}
$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ represent the probability density function and cumulative distribution function of the standard normal distribution, respectively. Furthermore, the average inventory level just before order delivery can be approximated as

$$IL^- = \int_{-\infty}^s (s - x)f(x)\mathrm{d}x = s - \mu + ELS.$$

Note that these expressions are independent of the replenishment policy. The approximate average inventory level just after order delivery is denoted by IL^+ and does depend on the replenishment policy. For the (R, s, Q) policy $IL^+ = IL^- + Q$, whereas for an (R, s, S) policy $IL^+ = IL^- + S - [s - \mu_u + ELS_u]$, where ELS_u is the expected lost sales during the waiting time for a new order placement after the inventory position has reached reorder level s similar to Equation (8). Furthermore, the average inventory level can be approximated by $IL = (IL^+ + IL^-)/2$.

Note that Equation (8) assumes that the inventory position after ordering exceeds reorder level s. This is satisfied in case $Q > s$ or $S - s > s$. However, when $L \leqslant R$, we do not impose restrictions on these inventory control variables, and we have to correct Equation (8) for the case that $\mu_R > Q$ and $Q \leqslant s$. In this special case the on-hand

inventory level before ordering is likely to be zero, and the length of a replenishment cycle is one review period. Therefore, Equation (8) becomes $ELS = \mu_R - Q$. No correction is required for the (R, s, S) policy, since the inventory position after ordering equals $S > s$.

The fill rate β is expressed as the satisfied demand divided by the total demand in a replenishment cycle. For the (R, s, Q) policy, this can be approximated by

$$\hat{\beta} = \frac{Q}{Q + ELS}. \tag{9}$$

Similarly, for the (R, s, S) policy,

$$\hat{\beta} = \frac{S - s + \mu_u - ELS_u}{S - s + \mu_u - ELS_u + ELS}. \tag{10}$$

In case there is a cost objective, the approximation of the expected total costs for the (R, s, Q) policy equals

$$\hat{C} = \frac{K + pELS}{(Q + ELS)/\mu_R} + hIL, \tag{11}$$

and for the (R, s, S) policy

$$\hat{C} = \frac{K + pELS}{(S - s + \mu_u - ELS_u + ELS)/\mu_R} + hIL. \tag{12}$$

The approximations of Equations (9) and (10) can be used to formulate the capacity model and service model as non-linear integer models to find optimal values of s and Q, which can be solved with standard software packages or spreadsheet-based programs like Excel. Therefore, we show in Section 8 that the approximation models derived in this section are appealing to be used in real-world applications and that they can be extended easily for models with a cost objective in other settings, such as a retail environment.

7. Inventory rule

The goal of this section is to develop a heuristic inventory rule for the capacity model with an (R, s, Q) replenishment policy that can easily be understood by hospital staff to decide upon the reorder level and the order size. This inventory rule can also be used in the multi-item model of Section 4.2.

The inventory rule consists of several tests. First, we check if the capacity C is sufficient to satisfy the demand. If the capacity is restrictive, we need to check whether it is likely that this restriction results in out-of-stock occurrences. Therefore, we examine if the reorder level s could be sufficient to be used as safety stock in order to fulfil demand until the next delivery. If this seems to be sufficient, we can determine the value of s such that stock outs are minimised. Otherwise, we need to find a balance between the reorder level and the order quantity.

The capacity is not restrictive when the order quantity is at least the average amount that is asked for during a

review period, that is $Q \geqslant \mu_R$. Another characteristic for this situation is that when no order is placed (ie, inventory level larger than s) the remaining inventory is sufficient to fulfil the demand until the next possible order delivery (ie, the demand until the next review and order delivery), or $s + 1 \geqslant \mu_R + \mu_L$. Since $s + Q = C$, the capacity is not restrictive if $C + 1 \geqslant 2\mu_R + \mu_L$. Therefore, we can set $s \in [\mu_R + \mu_L - 1; C - \mu_R]$ to obtain high service levels. We have chosen to set the value of s equal to the middle of this interval.

When there is a shortage of capacity, we want to order at least the average number of units that are asked for during a review period, that is $Q = \mu_R$. This order quantity is on average sufficient to satisfy demand between two order deliveries when orders are placed every review period. Owing to the stochastic nature of the demand, we cannot guarantee that an order is placed at each review. Therefore, we introduce an approximation for the probability that orders are placed every two succeeding reviews. This is only likely when $Q \leqslant \mu_R$ (ie, we assume the inventory level to be zero when an order arrives). A new order is placed when the delivered quantity minus the demand between order delivery and the next review is equal to or less than s. This is expressed by

$$P(Q - D_{R-L} \leqslant s) \geqslant \alpha \Leftrightarrow P(D_{R-L} \geqslant Q - s) \geqslant \alpha. \quad (13)$$

The tail probability of a Poisson distribution can be approximated by a normal distribution. Therefore Equation (13) is approximated by

$$1 - \Phi\left(\frac{Q - s - \mu_{R-L}}{\sqrt{\mu_{R-L}}}\right) \geqslant \alpha.$$

Now, we set α sufficiently large such that orders are placed every review period with a high probability ($\alpha = 0.98, \Phi^{-1}(1 - \alpha) \approx -2$). When the following inequality is satisfied there is a high probability that an order is placed each review period,

$$\frac{Q - s - \mu_{R-L}}{\sqrt{\mu_{R-L}}} \leqslant -2. \quad (14)$$

We can substitute $Q = \mu_R$ and $s = C - Q$ and check whether Equation (14) is satisfied. If it is, these parameter values are most likely to result in a high service level. Otherwise, we have to increase the reorder level s (and decrease order quantity Q) until Equation (14) is satisfied, that is,

$$\frac{C - 2s - \mu_{R-L}}{\sqrt{\mu_{R-L}}} \leqslant -2 \Leftrightarrow s \geqslant \frac{1}{2}\{C - \mu_{R-L} + 2\sqrt{\mu_{R-L}}\}.$$

The inventory rule can be summarised as follows:

1. If $C + 1 \geqslant 2\mu_R + \mu_L$, we set s equal to $(C + \mu_L)/2$ rounded to the nearest integer.
   Otherwise go to step 2.

2. If $(2\mu_R - \mu_{R-L} - C)/(\sqrt{\mu_{R-L}}) \leqslant -2$, we set s equal to $C - \mu_R$ rounded to the nearest integer.
   Otherwise go to step 3.

3. We set s equal to $1/2\{C - \mu_{R-L} + 2\sqrt{\mu_{R-L}}\}$ rounded to the nearest integer.

In all situations $Q = C - s$. Note, when $\mu_{R-L} = 0$ the second test should be $2\mu_R \leqslant C$ since $D_{R-L} = 0$ and Equation (13) specifies $Q \leqslant s$.

The inventory rule can be implemented with the use of a simple spreadsheet program and is, therefore, very appealing to be applied in many hospitals. When we apply this inventory rule to the data of Section 5 on infusion liquid, we obtain the following results for reorder level s: for paediatrics 3, for intensive care 20 and for obstetrics 41. When we compare these results to the optimal values in Table 3, we conclude that the results are very close.

8. Numerical results

The goal of this section is to illustrate the performance of the approximation procedure (Section 6) and the inventory rule (Section 7). Furthermore, we include a comparison of the (R, s, Q) policy and the (R, s, S) policy similar to the case study (Section 5). In order to test the performance for a wide range of settings that correspond to the characteristics of Section 2, we specify test instances in which the average demand in a review period equals 5 to 30 units (with steps of 5 units), while the lead time varies between 0.125 and one times the review period length. In order to test all three situations of the inventory rule, the capacity ranges from μ_R to $3\mu_R$ with steps of $0.5\mu_R$. This results in 240 test instances for the capacity model. In the service model, there is no capacity limitation but a minimum service level $\bar{\beta}$ is specified (see Section 4), where $\bar{\beta}$ varies from 90 to 98%. In the final setting we demonstrate the performance of the approximation procedure in more general settings than Section 2. Therefore, we consider a cost objective and a lead time equal to two review periods.

8.1. Capacity model

The reorder level that maximises the fill rate in the capacity model for an (R, s, Q) policy is denoted by s^*, and $\bar{s}$ represents the value of the reorder level for the (R, s, S) policy where the order frequency is similar as for the (R, s, Q) policy with reorder level s^* (similar to Section 5). The values of the reorder level based on the approximation procedure and heuristic rule are denoted by $\hat{s}_1$ and $\hat{s}_2$, respectively. Table 5 shows the aggregated results over the eight different values of the lead time because they represent a similar capacity limitation. The average fill rate for the (R, s, Q) policy where $s = s^*$ and $Q = C - s^*$ is

Table 5 The average fill rate for using the (R, s, Q) policy with $s = s^*$ and the increase and decrease of the service level for using the (R, s, S) policy with $s = \bar{s}$ and the (R, s, Q) policy with $s = \hat{s}$ respectively

μ_R	C	$\beta_{(R, s, Q)}(s^*)$ (%)	$\Delta\beta(\bar{s}, s^*)$ (%)	$\Delta\beta(s^*, \hat{s}_1)$ (%)	$\Delta\beta(s^*, \hat{s}_2)$ (%)
5	5	52.26	8.07	0.36	10.02
5	8	74.35	9.24	0.78	1.64
5	10	83.65	8.29	1.65	0.29
5	13	92.98	4.67	1.87	0.29
5	15	96.54	2.49	3.09	0.21
10	10	56.90	9.43	0.00	3.67
10	15	75.27	9.58	0.27	1.05
10	20	87.68	7.09	0.70	1.05
10	25	94.97	3.84	2.00	0.39
10	30	98.45	1.35	2.29	0.22
15	15	57.90	2.71	1.05	2.51
15	23	78.86	8.67	0.27	0.27
15	30	89.67	6.40	0.90	1.52
15	38	96.55	2.85	1.59	0.22
15	45	99.07	0.84	1.82	0.19
20	20	59.88	1.23	0.00	1.21
20	30	79.48	8.28	0.60	0.04
20	40	90.96	6.18	0.99	1.85
20	50	97.00	2.50	1.43	0.29
20	60	99.36	0.59	1.52	0.15
25	25	60.37	8.30	0.82	1.39
25	38	81.39	7.24	0.27	0.07
25	50	91.93	5.48	1.38	2.13
25	63	97.60	2.08	1.24	0.17
25	75	99.52	0.45	1.31	0.15
30	30	61.21	0.73	0.00	0.62
30	45	81.65	7.19	0.23	0.18
30	60	92.60	5.02	1.46	2.28
30	75	97.80	1.89	1.02	0.24
30	90	99.62	0.35	1.15	0.13

The results are aggregated over $L/R = 1/8, \ldots, 1$.

denoted by $\beta_{(R, s, Q)}$, whereas $\Delta\beta(\bar{s}, s^*) = \beta_{(R, s, S)}(\bar{s}) - \beta_{(R, s, Q)}(s^*)$ and $\Delta\beta(s^*, \hat{s}) = \beta_{(R, s, Q)}(s^*) - \beta_{(R, s, Q)}(\hat{s})$.

Based on the results of Table 5 we conclude that the fill rate can increase significantly (about 5%) when the (R, s, S) policy is applied compared to the (R, s, Q) policy. However, when the service level is high, the difference between the policies is less. The approximation procedure and the inventory rule are very effective to find good values for the reorder level in all situations. Only when the capacity is equal to the average demand in a review period for a slow moving item the heuristic rule performs badly. However, such situations should be avoided by hospitals at all times, since the average fill rate is below 60% in those situations. In general we conclude that the approximation procedure and the inventory rule perform on average within 1–2% from the optimal service level.

8.2. Service model

The same notation and test bed are used for the service model as for the capacity model. Recall from Section 7 that the inventory rule is designed for the capacity model and is therefore not included in these results. The results are aggregated over the lead time (similar to Table 5) and they are shown in Table 6, where $C^* = s^* + Q^*$ and $\hat{C} = \hat{s}_1 + \hat{Q}_1$. In most of the numerical results we observe that $\hat{C} \geqslant C^*$, that is the approximation model underestimates the service level. Based on the results in Table 6 we conclude that the approximation procedure performs well in case the service level is rather high (at least 95%). Furthermore, we note that the order frequency for the service model is close to one review period since the service model minimises inventory costs. As mentioned by Tijms and Groenevelt (1984), Equations (6) and (7) perform better if the average order size is larger than 1.5 times the average demand.

8.3. General setting

We demonstrate the performance of the approximation procedure of Section 6 for more general inventory systems than described in Section 2. In particular we consider systems with a lead time equal to two review periods and a cost objective with holding cost $h = 1$ per unit time and fixed order cost $K = 25, 50, 100$. Furthermore, demand

Table 6 The minimal required capacity for the (R, s, Q) policy to satisfy the service level constraint $\overline{\beta}$ for the exact service model (C^*), and the capacity increase when the approximate service model is used

μ_R	$\overline{\beta} = 90\%$			$\overline{\beta} = 95\%$			$\overline{\beta} = 98\%$		
	C^*	$\lvert\hat{C} - C^*\rvert$	$\lvert\hat{C} - C^*\rvert / C^*$ (%)	C^*	$\lvert\hat{C} - C^*\rvert$	$\lvert\hat{C} - C^*\rvert / C^*$ (%)	C^*	$\lvert\hat{C} - C^*\rvert$	$\lvert\hat{C} - C^*\rvert / C^*$ (%)
5	12.4	0.4	3.35	14.3	1.0	7.12	16.5	2.0	12.38
10	21.4	1.1	5.20	24.9	0.5	2.02	28.6	2.1	7.67
15	30.4	1.6	5.26	35.1	1.0	2.89	40.0	1.4	3.61
20	38.5	4.1	10.65	45.5	1.4	2.91	51.8	1.1	2.29
25	46.5	5.8	12.45	54.8	2.6	4.78	63.0	0.6	1.04
30	54.5	8.1	15.05	64.1	3.9	6.00	74.1	0.6	0.82

The results are aggregated over $L/R = \{1/8, \ldots, 1\}$.

Table 7 The performance of the approximation procedure for the (R, s, Q) policy and the (R, s, S) policy in more general inventory settings

$\overline{\beta}$ (%)	K	(R, s, Q) policy						(R, s, S) policy					
		s^*, Q^*	$C(s^*, Q^*)$	$\beta(s^*, Q^*)$ (%)	$\hat{s}_1, \hat{Q}_1$	$C(\hat{s}_1, \hat{Q}_1)$	$\beta(\hat{s}_1, \hat{Q}_1)$ (%)	$\bar{s}, \overline{S}$	$C(\bar{s}, \overline{S})$	$\beta(\bar{s}, \overline{S})$ (%)	$\hat{s}_2, \hat{S}_2$	$C(\hat{s}_2, \hat{S}_2)$	$\beta(\hat{s}_2, \hat{S}_2)$ (%)
75	25	7,17	12.37	75.28	8,15	12.58	76.26	7,22	12.56	75.76	8,20	12.63	75.77
75	50	5,23	17.25	75.38	6,20	17.36	75.38	5,26	17.36	75.63	6,24	17.48	75.75
75	100	2,32	24.19	75.29	3,29	24.29	75.32	1,35	24.26	75.16	4,28	24.71	75.45
80	25	9,16	13.42	80.74	9,16	13.42	80.74	9,22	13.46	80.39	9,23	13.61	81.29
80	50	7,23	18.47	80.48	8,19	18.73	80.30	7,27	18.44	80.10	8,25	18.82	80.68
80	100	5,31	25.87	80.49	6,27	26.21	80.52	5,33	25.81	80.14	6,31	26.28	80.75
85	25	10,18	14.51	85.50	10,18	14.51	85.50	10,25	14.50	85.27	11,23	14.75	85.65
85	50	9,22	19.75	85.24	9,22	19.75	85.24	9,29	19.88	85.52	9,29	19.88	85.52
85	100	7,32	27.42	85.15	7,32	27.42	85.15	7,37	27.54	85.33	7,37	27.54	85.33
90	25	12,16	15.91	90.37	12,17	16.04	90.84	12,26	15.92	90.34	12,26	15.92	90.34
90	50	11,22	21.43	90.36	11,22	21.43	90.36	11,30	21.43	90.14	11,31	21.51	90.55
90	100	9,35	29.40	90.18	10,29	29.67	90.53	9,41	29.36	90.05	10,36	29.71	90.38
95	25	14,18	18.01	95.32	14,17	17.79	94.93	14,30	18.02	95.32	14,29	17.82	94.98
95	50	13,25	23.58	95.10	13,25	23.58	95.10	13,36	23.65	95.19	13,35	23.51	95.00
95	100	12,34	31.82	95.06	12,34	31.82	95.06	12,44	31.89	95.13	12,44	31.89	95.13
98	25	16,20	20.28	98.01	16,19	20.05	97.88	16,34	20.30	98.02	16,33	20.08	97.90
98	50	16,22	26.29	98.15	15,29	26.18	97.89	16,35	26.24	98.11	16,35	26.24	98.11
98	100	15,31	34.56	98.02	15,31	34.56	98.02	15,44	34.54	98.05	15,44	34.54	98.05

follows a Poisson process with mean $\mu_R = 5$ and a service level restriction $\overline{\beta} = 75\%, 80\%, 85\%, 90\%, 95\%, 98\%$.

The values of the inventory control variables that minimise the average holding and order costs subject to the service level constraint are denoted by s^* and Q^* for the (R, s, Q) policy, and $\bar{s}$ and $\overline{S}$ for the (R, s, S) policy. Furthermore, we determine the inventory control variables based on the approximation procedure of Section 6, denoted by $\hat{s}_1$ and $\hat{Q}_1$ for the (R, s, Q) policy, and $\hat{s}_2$ and $\hat{S}_2$ for the (R, s, S) policy. The results are shown in Table 7. In 15 of the 18 test instances the approximation procedure results in inventory control variables that satisfy the service level constraint, whereas the costs are on average within 1% from the optimal costs. Based on these results we conclude that our approximation procedure performs very well, and can be used in many inventory control settings with lost sales such as retail environments.

9. Conclusion

The inventory replenishment system at POU locations in hospitals can be classified as a lost-sales inventory system where the lead time is shorter than the length of a review period and the focus is on service levels. Another characteristic of hospital inventory management is the lack of available storage capacity. We developed capacity and service models for such inventory systems and compared their performances for (R, s, Q) policies and (R, s, S) policies. Both types of replenishment policies are common in hospitals. The fixed order size policy results in a more insightful replenishment process for hospitals with the use of bar codes. However, the (R, s, S) policy uses the capacity more efficiently. If inventory levels are monitored automatically (for instance, when RFID-chips are used) such policies are recommended. However, both replenish-

ment policies perform equally well when the service level is rather high.

We also derived closed-form expressions that approximate performance measures (like the service level) in order to set the inventory control variables. We demonstrated that this approximation procedure can also be used in more general settings other than a hospital inventory system, including a cost objective. The procedure performs very good when the average order size is larger than 1.5 times the average demand in a review period. Furthermore, we developed a simple inventory rule that finds near-optimal values for the reorder levels and order quantities for the capacity model. This inventory rule can easily be embedded in multi-item algorithms that assign items to the available capacity at different POU locations. It can also be used to determine the required capacity.

One possible aspect for future research is the influence that substitution products have on the service level in case of a stock out. Another interesting aspect to investigate would be the interaction of the inventory control between the POU locations and the higher-level stock points like the central storage room. Especially how to set the lead time and the review period length, since they influence the performance of the POU location but they are typically determined by the supplier (ie, the higher-level stock point).

References

Axsäter S (2006). *Inventory Control*. Chapter 7. Springer: New York.

Baganha MP, Pyke DF and Ferrer G (1996). The undershoot of the reorder point: Tests of an approximation. *Int J Prod Econ* **45**: 311–320.

Bijvank M and Johansen SG (2009). *Periodic review lost-sales inventory models with non-fractional lead times*. Working paper.

Bijvank M and Vis IFA (2010). *Lost-sales inventory systems with a service level criterion*. Working paper.

Bijvank M and Vis IFA (2011). Lost-sales inventory theory: A review. *Eur J Opl Res*. doi:10.1016/j.ejor.2011.02.004.

Brandeau ML, Sainfort F and Pierskalla WP (2004). *Operations Research and Health Care—A Handbook of Methods and Applications*. Kluwer Academic Publishers: Dordrecht, MA.

Chiang C (2006). Optimal ordering policies for periodic-review systems with replenishment cycles. *Eur J Opl Res* **170**: 44–56.

Chiang C (2007). Optimal ordering policies for periodic-review systems with a refined intra-cycle time scale. *Eur J Opl Res* **177**: 872–881.

Danas K, Ketkidis P and Roudsari A (2002). A virtual hospital pharmacy inventory: An approach to support unexpected demand. *Int J Med Market* **2**: 125–129.

Dellaert N and Van de Poel E (1996). Global inventory control in an academic hospital. *Int J Prod Econ* **46–47**: 277–284.

Downs B, Metters R and Semple J (2001). Managing inventory with multiple products, lags in delivery, resource constraints, and lost sales: A mathematical programming approach. *Mngt Sci* **47**: 464–479.

Duclos LK (1993). Hospital inventory management for emergency demand. *Int J Purch Mater Mngt* **29**: 30–37.

Epstein RH and Dexter F (2000). Economic analysis of linking operating room scheduling and hospital material management information systems for just-in-time inventory control. *International Anesthesia Research Society* **91**: 337–343.

Hadley G and Whitin TM (1963). *Analysis of Inventory Systems*. Prentice-Hall: Englewood Cliffs, New York.

Hill RM (1994). Continuous review lost sales inventory models where two orders may be outstanding. *Int J Prod Econ* **35**: 313–319.

Hill RM (1999). On the suboptimality of $(S-1, S)$ lost sales inventory policies. *Int J Prod Econ* **59**: 387–393.

Hill RM and Johansen SG (2006). Optimal and near-optimal policies for lost sales inventory models with at most one replenishment order outstanding. *Eur J Opl Res* **169**: 111–132.

Janssen F, Heuts R and De Kok T (1998). On the (R, s, Q) inventory model when demand is modelled as a compound Bernoulli process. *Eur J Opl Res* **104**: 423–436.

Johansen SG and Hill RM (2000). The (r, Q) control of a periodic-review inventory system with continuous demand and lost sales. *Int J Prod Econ* **68**: 279–286.

Johansen SG and Thorstenson A (1996). Optimal (r, Q) inventory policies with poisson demands and lost sales: Discounted and undiscounted cases. *Int J Prod Econ* **46–47**: 359–371.

Johansen SG and Thorstenson A (2004). *The (r, q) policy for the lost-sales inventory system when more than one order may be outstanding*. Working paper L-2004-03, Aarhus School of Business.

Johnston DH (1992). Primary supplier relationships for all inventory items can reduce a hospital's operating costs. *Hosp Mater Mngt* **17**: 17–18.

Kapalka BA, Katircioglu K and Puterman ML (1999). Retail inventory control with lost sales, service constraints, and fractional lead times. *Prod Opns Mngt* **8**: 393–408.

Katsaliaki K and Brailsford SC (2007). Using simulation to improve the blood supply chain. *J Opl Res Soc* **58**: 219–227.

Lapierre SD and Ruiz AB (2007). Scheduling logistic activities to improve hospital supply systems. *Comp Opns Res* **34**: 624–641.

Little J and Coughlan B (2008). Optimal inventory policy within hospital space constraints. *Health Care Mngt Sci* **11**: 177–183.

Moon S (2004). Taking cost off supply shelf. *Modern Healthcare* **34**: 26–28.

Nathan J and Trinkaus J 1996. Improving health care means spending more time with patients and less time with inventory. *Hosp Mater Mngt Q* **18**: 66–68.

Nicholson L, Vakharia AJ and Erenguc SS (2004). Outsourcing inventory management decisions in healthcare: Models and application. *Eur J Opl Res* **154**: 271–290.

Prastacos GP (1984). Blood inventory management: An overview of theory and practice. *Mngt Sci* **30**: 777–798.

Rivard-Royer H, Landry S and Beaulieu M (2002). Hybrid stockless: A case study; lessons for health-care supply chain integration. *Int J Opns Prod Mngt* **22**: 412–424.

Sezen B (2006). Changes in performance under various lengths of review periods in a periodic review inventory control system with lost sales—A simulation study. *Int J Phys Distrib Logist Mngt* **36**: 360–373.

Silver EA, Pyke DF and Peterson R (1998). *Inventory Management and Production Planning and Scheduling*. Chapter 11. John Wiley & Sons: New York.

Tijms HC (1994). *Stochastic Models*. Wiley: Chichester.

Tijms HC and Groenevelt H (1984). Simple approximations for the reorder point in periodic and continuous review (s, S) inventory systems with service level constraints. *Eur J Opl Res* **17**: 175–190.

van Donselaar K, de Kok T and Rutten W (1996). Two replenishment strategies for the lost sales inventory model: A comparison. *Int J Prod Econ* **46–47**: 285–295.

van Houtum GJ and Zijm WHM (2000). On the relationship between cost and service models for general inventory systems. *Statistica Neerlandica* **54**: 127–147.

Vincent V and Ranton M (1984). Hospital pharmacy inventory management: Economic order quantity model with space limitation. *Hosp Mater Mngt Q* **5**: 82–96.

Zipkin PH (2000). *Foundations of Inventory Management*. McGraw-Hill: New York.