Journal of Global Optimization
https://doi.org/10.1007/s10898-021-01058-3

Hospital service levels during drug shortages: Stocking and transshipment policies for pharmaceutical inventory

Cem Deniz Caglar Bozkir[1] **· O. Erhun Kundakcioglu**[1] **· Andrea C. Henry**[2,3]

Received: 17 December 2019 / Accepted: 30 May 2021

Abstract

In this study, we consider a health network that faces uncertain supply disruptions in the form of regional, nationwide, or worldwide drug shortages. Each hospital observes stochastic demand and if the drug is unavailable, patients leave and receive care in another network. As these instances of unavailability diminish the brand value, health networks look for inventory sharing mechanisms among hospitals to mitigate the effect of uncertain supply disruptions. In line with this expectation, we propose a proactive inventory sharing approach for critical drugs to investigate the effect of the inventory-related parameters on service levels.

Keywords Health services optimization · Inventory management · Drug shortages · Service level · Pooling

1 Introduction

Healthcare providers primarily focus on improving society's well-being in contrast to profit-maximizing players of commercial supply chains. Health expenditure constitutes around 10% of countries' GDPs on average, with the healthcare sector's market value estimated at \$8.7 trillion [6]. Hospitals comprise the largest share of healthcare expenditure, ranging from 26 to 53% [19], and spend an estimated 10–18% of their budgets on inventory-related investments [18], which makes inventory one of their most important cost components.

Service level, namely the number of patients served adequately, is an equally important measure for hospitals. A high level of service is crucial in the healthcare system because it represents the proportion of people able to receive timely treatment. Some hospitals share scarce resources (e.g., drugs and vaccines) within a network under a predefined policy that controls the transfer of drugs between facilities. Shortages are thus a major issue for healthcare systems, crippling the services provided to incoming patients. In April 2018, for example,

✉ O. Erhun Kundakcioglu
erhun.kundakcioglu@ozyegin.edu.tr

1 Department of Industrial Engineering, Ozyegin University, Istanbul, Turkey

2 Department of Pharmacy, Pharmacoeconomics and Formulary Management, Harris Health System, Houston, TX, USA

3 RX Management, Magellan Health, Phoenix, AZ, USA

Published online: 30 July 2021

there were about 100 reported drug shortages in the United States [7], in such areas as antibiotic treatments, chemotherapy, pain medication, and drugs used in surgeries. Being unable to provide drugs in a timely manner has a negative impact on a healthcare system (e.g., the hospital chain's brand value). Stock transfers are one way to maintain the required service level while lowering the inventory level; the other is to increase the number of customers served with the same number of items in the inventory of a system [22]. Therefore, hospitals desiring to maintain a certain service level can achieve their targets by allowing transshipments. However, as hospitals manage their inventories using simple procedures based mostly on intuition instead of sophisticated approaches, any improvement in this area will have a direct effect not only on human health through improved services but also on the healthcare delivery system, as any increase in the number of people treated leads to a better image in the public eye. This is our motivation for this study.

There exist several studies in the area of inventory management for systems that consider transshipments. [22] and [28] review these studies in detail. Broadly, the literature can be classified into two categories according to the timing of transshipments. *Proactive* transshipments allow the redistribution of stocks between retailers, whereas *reactive* stock transfers occur only when a stock-out is observed. We categorize studies similar to ours according to these streams of the literature.

Managing transshipments proactively is widely employed for handling stock transfers within a system. Nasr et al. [17] consider a production environment consisting of two locations in which random supply disruptions may be observed. Transshipments are used in a way that when a supply shortage occurs, retailers adjust their safety stock levels instantaneously by sharing inventory to hedge against the risk of a stock-out. van Wijk et al. [27] analyze the proportion of demand satisfied directly from a retailer's own stock, using transshipments, and emergency ordering for a spare parts inventory system adopting continuous review policies and transshipment thresholds, while considering the recovery rate for failed parts. Cheong [5] introduces a newsvendor model for multiple demand locations where proactive transshipments may occur immediately and the perishability of products is included. Feng et al. [8] develop a periodic review base stock policy for a multi-location problem that considers perishability, where proactive transshipments are also permitted, and this is modeled as a Markov decision process. Hochmuth and Köchel [13] consider a similar setting in which retailers adopt periodic review (s, S) policies for ordering and continuous review policies for transshipment decisions. They aim to find the optimal variables for policies using simulations. Zhao et al. [29] propose a queuing theory approach for a system consisting of multiple retailers. Inventory level can increase because of production and decrease because of demand, both with Poisson rates, and retailers can request transshipments even when they have a positive inventory. They aim to find the optimal maximum inventory levels and thresholds for requesting transshipments and accepting transshipment requests. Tagaras and Vlachos [26] evaluate the performance of a system with two locations that replenish their inventories periodically, with respect to the redistribution of stock between the two locations in each replenishment period.

Some studies also consider reactive transshipments. Archibald et al. [2] introduce a periodic review policy using a discrete-time Markov decision process for two retailers and one supplier. They allow transshipments and emergency orders at any time in a period. This study is extended to a multi-location setting by Archibald [1]. Axsäter [4] defines a decision rule for transshipments between retailers that utilize continuous review policies, considering future costs and outstanding orders, with known arrival times. Minner and Silver [16] study a continuous review (R, Q) inventory system with two identical retailers subject to Poisson demand, which can use transshipments in the case of a stock-out until replenishments arrive.

Ramakrishna et al. [24] develop a Markov decision process to decide the periodic review policy parameters and whether to accept a transshipment request according to the time remaining for the replenishment. Herer and Rashit [11] evaluate a two-location single-period newsvendor model with joint and fixed replenishment costs, and transshipments when required. Later, Herer and Tzur [12] offer an extension by modeling this as a network flow problem. Paterson et al. [23] adopt the policy presented by Axsäter [4] to derive a transshipment rule that states that every reactive transshipment serves as an opportunity for the proactive reallocation of the inventory in the system. Glazebrook et al. [9] also consider a hybrid transshipment policy in a periodically replenished system. In Grahovac and Chakravarty [10], a base stock policy with a transshipment threshold is proposed, but transshipments are only allowed when an emergency order cannot be received in time to meet the observed demand. Olsson [20] proposes a (Q, R) policy with reactive transshipments for a two-location system under exponentially distributed lead times.

The closest studies to our approach are those proposing models that consider the service level in a system. Axsäter [3] proposes a continuous review replenishment policy for a network consisting of multiple retailers, where transshipments between retailers are allowed in only one direction. Under normally distributed demand, a non-linear approximation technique based on different fill rates for the sources of demand satisfaction is developed for the problem, and the results of their model are compared with a simulation study. Olsson [21] models the observed demand in Axsäter [3] as a Poisson process and proposes two models based on $(S-1, S)$ and (Q, R) policies. Kukreja and Schmidt [15] investigate a system with multiple retailers at demand locations and a single supplier. Each of the retailers uses continuous review (s, S) policies, and complete inventory pooling between them is allowed. Under the retailers' stochastic demand parameters, they aim to measure the performance of the system using simulations when inventory sharing is allowed, while order quantities are set as EOQ and reorder levels are adjusted iteratively.

The aforementioned review shows that the most dominant metric among transshipment studies is the cost incurred by the considered systems, which consists of different components such as the holding, transshipment, backordering, emergency ordering, and lost sales costs. However, owing to the unique nature of healthcare systems, the achieved service level is at least as much important as the operating cost of the system. This is particularly the case for *critical* drugs that are vital for the patients (e.g., chemotherapeutic drugs). Therefore, in this study, we address this gap in the literature by making meaningful observations about the service level of the system while keeping a realistic amount in stock.

The motivation of our study is the need for a sharing infrastructure among hospitals in the same network. Such a network can be a chain of hospitals or a set of regional hospitals managed by the government. The latter practice is common in Europe, especially in non-metropolitan districts. To the best of our knowledge, no study has thus far conducted service level-based analyses of a proactive transshipment policy for continuous review base stock inventory systems. While several studies adopt uncertain lead times, their results cannot be adapted to a system that observes shortages leading back-and-forth with uncertain durations. Our objective in this study is thus to investigate the effect of *inventory and pooling decisions* on the service level of a healthcare system observing shortages. We formulate two nonlinear optimization models that maximize the expected proportion of patient demand satisfied with and without transshipments, which we use to decide on the optimal *transshipment threshold* and *pooled inventory* of each hospital.

The rest of this paper is organized as follows. The considered system is described in more detail in Sect. 2. The proactive transshipment policy is introduced in Sect. 3. A numerical study is presented in Sect. 4 to measure the performance of our proposed methodology and provide managerial insights. We present our concluding remarks in Sect. 5.

2 System properties and key performance indicators

We consider the inventory of critical drugs for a healthcare system with multiple hospitals and a single supplier. Hospitals' inventories are replenished by the supplier according to a continuous review $(S, S-1)$ base stock policy. We assume a zero lead time for replenishments, which is a common assumption because of healthcare systems' frequent replenishment cycles [18,25]. In general, hospitals are visited at least once a day by suppliers for replenishment, if necessary. In other words, when a hospital places an order, the delivery is usually made the same day (or the next day in the worst case) if the supplier has the drugs in stock. However, the supplier is open to disruptions that are stochastic in both duration and frequency. Hospitals need to hedge against the risk of stock-out when a shortage at the supplier exists. Assuming a zero lead time helps us ensure that hospitals' inventory levels are equal to their order-up-to levels at the beginning of a shortage period. Moreover, we assume that transshipments occur instantaneously, which is widely accepted in the literature [14,17].

In practice, the hospitals in a healthcare network do not determine their base stock using complex mathematical inventory policies. Since a higher service level is more desirable than operating with a lower holding cost in healthcare networks, hospitals ignore the holding cost they pay and stock as many drugs as their inventory capacities can handle. Therefore, throughout this study, we do not consider holding costs in our analyses and make no observations or conclusions about base stock policies. This approach allows us to establish decision rules that consider the trade-off between the movement of inventory on hand and hedging against the risk of potential lost sales.

Hospitals are willing to stock as many items as possible to ensure they never run out of the medication needed to treat a patient. Sharing the entire inventory further increases the fill rate of the system. However, capacity restrictions and costs are associated with transshipment. Therefore, hospitals aim to keep the stock transfers and related costs at a certain level. For healthcare systems that face shortages, a target service level is defined at the "strategic" level. Hospitals use inventory and transshipment parameters to ensure that the service level is achieved.

The literature presents different inventory pooling approaches (i.e., retailers' allowed method of making stock transfers), namely *complete* and *partial* pooling policies. As the names imply, these policies define whether retailers share all or some of their stocks, respectively. Complete pooling is more suitable for systems in which the holding and backordering costs are significantly higher than the transshipment costs [22]. This method is similar to the case of the considered system if replenishments can be made regularly. However, when shortages that are unknown in length and frequency are present in the system, there is a trade-off for hospitals between requesting the stock transfer necessary to satisfy demand and retaining their inventory to satisfy their own potential demand. Therefore, we allow partial pooling between hospitals and compare the complete and partial pooling strategies.

2.1 Sharing mechanism

We refer to the levels at which hospitals reject the transshipment requests and keep their remaining inventories to meet their own potential demand as the *transshipment thresholds* or *safety stocks*. In addition, we assume that hospitals' inventory levels are visible to each other; hence, they exactly know which ones can make transshipments when requesting a stock transfer. In line with our zero lead time replenishment, we assume transshipments can also be performed instantaneously, as these times are usually negligible (4–6 hours) in reality.

Within the scope of this study, we investigate a **proactive** approach where hospitals request a stock transfer when they hit their transshipment threshold. The request is granted by another hospital above its transshipment threshold. This means that individual inventories above transshipment thresholds are pooled, acting as a single entity. For prolonged shortages, this ensures that all hospitals are at their thresholds at a point in time, after which they do not share inventory.

Studies considering supply unavailability and demand uncertainty mostly utilize Markov chains because of their interest in the steady-state behavior of systems. However, as we are interested in non-recursive service levels during shortage periods, we use the properties of Poisson processes and exponential distributions to derive these probabilities.

2.2 Service levels for hospitals

In this study, we assume each patient arrival implies demand. No arrival places a batch order as in a commercial supply chain. Therefore, the term *service level* in this study indicates the fill (service) *rate*. We focus on the following two types of service levels hereafter.

2.2.1 Type I service level

For a hospital system, we define the Type I service level as the fill rate of the entire system during shortages. This can be translated into the probability of satisfying patient demand during shortages. Formally, the *Type I service level during shortages* is defined as

$$\alpha_S = P(\text{demand during shortage} \leq \text{inventory on hand at the beginning of shortage period}) \tag{1}$$

2.2.2 Type II service level

For a hospital system, we describe the proportion of demand delivered without delay from stock on hand during shortages. As there is no lead time for transshipments, we formally define the *Type II service level during shortages* as

$$\beta_S = \alpha_S - \frac{E[\text{transshipments during shortage}]}{E[\text{demand during shortage}]} \tag{2}$$

The Type I and II service levels during shortages can be generalized in a straightforward way. The proportion of shortage durations can be computed depending on the shortage and recovery rates. As both service levels are 100% during a non-shortage period, the general Type I (α) and Type II (β) service levels can be computed using conditioning as follows:

$$\alpha = \frac{\alpha_S \times \text{shortage rate} + \text{recovery rate}}{\text{shortage rate} + \text{recovery rate}} \tag{3}$$

$$\beta = \frac{\beta_S \times \text{shortage rate} + \text{recovery rate}}{\text{shortage rate} + \text{recovery rate}} \tag{4}$$

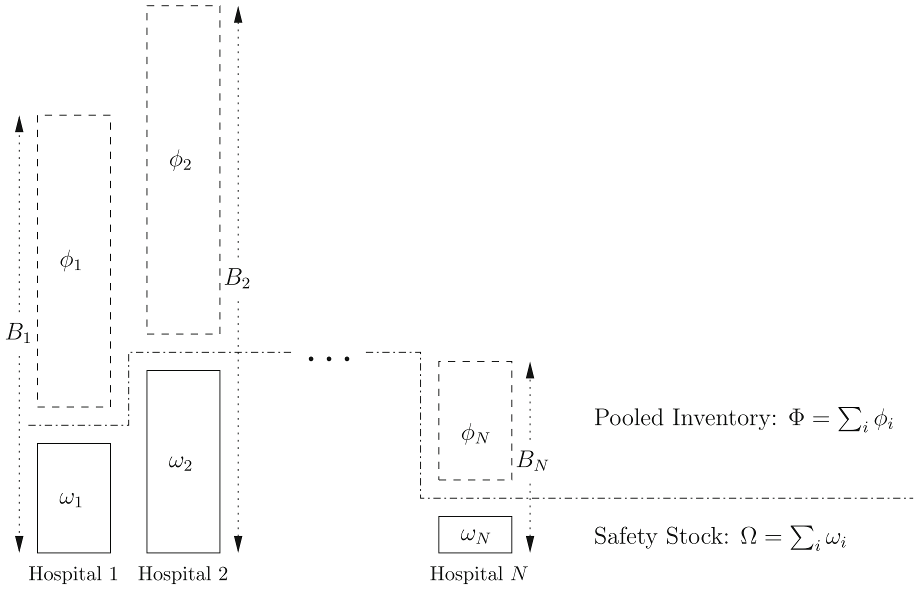


Fig. 1 Allocation of the pooled inventory and safety stock

3 Proactive sharing policy

As described previously, we concentrate on a proactive approach that allows hospitals to make a transfer request upon demand if they already hit their individual transshipment thresholds. The occurrence of a transshipment depends on two factors:

- The hospital requesting a transfer has an inventory level of exactly its transshipment threshold. In other words, a hospital makes a transfer request to ensure a predetermined safety stock on hand.
- A hospital in the system has an inventory level above its threshold to serve the hospital requesting the transfer.

In other words, hospitals in the network make transshipments to consume the partially pooled inventory together at the beginning. When all their inventory levels become equal to their transshipment thresholds, they begin to satisfy the observed demand from their own inventories, if possible. Any unsatisfied demand becomes lost (i.e., backordering is prohibited).

Let us introduce the following notation:

$\mathcal{B}$: total base stock,

B_i: base stock at hospital i,

N: number of hospitals in the system,

Ω: total safety stock,

γ_i: proportion of the total safety stock for hospital i, $\sum \gamma_i = 1$,

ω_i: transshipment threshold for hospital i, $\omega_i = \gamma_i \Omega$,

Φ: pooled inventory,

ϕ_i: inventory reserved for pooling at hospital i,

λ_i: demand rate for hospital i,

μ: recovery rate for shortage occurrences.

3.1 Inventory parameters that maximize the Type I service level

We consider a system of N hospitals, each observing Poisson-distributed demand at the rate λ_i ([21,29]). Further, the shortage durations are exponentially distributed at the rate μ [25]. Hospitals regularly stock $\mathcal{B}$ items, and their inventories are continuously replenished according to a base stock policy. When a shortage occurs, they reserve Ω items as safety, and $0 \leq \gamma_i \leq 1$ is the proportion of the safety stock held by hospital i, where $\sum_i \gamma_i = 1$. During shortages, hospitals consume the system's inventory together by making transshipments when needed, until *all* their inventory levels become their transshipment threshold (individual safety stock), namely $\gamma_i \Omega$, $\forall i$. When *all* hospitals' inventories reach their respective thresholds, they begin to satisfy any observed demand from their own inventories. Any unsatisfied patient demand is assumed to be lost. This happens when the safety stock is depleted, as no more transshipments are allowed. Figure 1 summarizes the system and related variables.

First, we formally define the equation for the Type I service level as a function of the inventory variables.

Theorem 1 *For a system of hospitals, given the total base stock* ($\mathcal{B}$)*, total safety stock value* (Ω)*, and proportion of each hospital in the total threshold* (γ_i)*, the Type I service level is*

$$\alpha_S = 1 - \left(\frac{\sum_j \lambda_j}{\mu + \sum_j \lambda_j}\right)^{\mathcal{B}-\Omega} \sum_i \left(\frac{\lambda_i}{\lambda_i + \mu}\right)^{\gamma_i \Omega} \frac{\lambda_i}{\sum_j \lambda_j}.$$

An increase in $\mathcal{B}$ increases the service level achieved. However, since we aim to keep transshipments between hospitals at a certain level, we require Ω to be a positive value. The allocation of the total threshold between hospitals can be determined by finding the value for each γ_i that maximizes the service level equation.

Corollary 1 *As* Ω *decreases, for fixed* $\mathcal{B}$*, the Type I service level increases.*

The proof of Corollary 1 is straightforward and can be easily observed from Theorem 1. Therefore, it can be stated that the optimal Type I service level is achieved when Ω is zero. However, from Theorem 1, no conclusion on the Type II service level can be made.

Corollary 2 *For a total base stock of* $\mathcal{B}$*, the best possible Type I service level is attained under complete pooling, where*

$$\alpha_S^{\max} = 1 - \left(\frac{\sum_j \lambda_j}{\mu + \sum_j \lambda_j}\right)^{\mathcal{B}}.$$

Moreover, for a total base stock of $\mathcal{B}$, the service level when no transshipments are allowed is

$$\alpha_S^{\text{No Transshipment}} = 1 - \sum_i \left(\frac{\lambda_i}{\lambda_i + \mu}\right)^{\gamma_i \mathcal{B}} \frac{\lambda_i}{\sum_j \lambda_j}.$$

Next, we present the optimal threshold levels for each hospital.

Theorem 2 *For a system of hospitals, given the total safety stock* (Ω)*, to maximize the Type I service level, the proportions of the transshipment threshold for any two hospitals k and m must hold the following equality:*

$$\gamma_m \ln\left(\frac{\lambda_m}{\mu + \lambda_m}\right) - \gamma_k \ln\left(\frac{\lambda_k}{\mu + \lambda_k}\right) = \frac{1}{\Omega} \ln\left(\frac{\lambda_k \ln\left(\frac{\lambda_k}{\mu+\lambda_k}\right)}{\lambda_m \ln\left(\frac{\lambda_m}{\mu+\lambda_m}\right)}\right).$$

From Theorem 2, the proportions of the threshold for hospital i, γ_i, can be found by solving a set of linear equations including $\sum_i \gamma_i = 1$.

3.2 Amount to pool at each hospital to maximize the Type II service level

After defining the service level for the given base stock and total threshold, the question becomes how to determine the inventory to be pooled. Knowing the amounts to stock at each location when pooling the inventory, the total threshold among hospitals can be allocated using Theorem 2.

In practice, hospitals proactively share their inventories for a predefined period when a shortage occurs. They begin to satisfy patient demand from their own safety stock if the shortage continues when the sharing period is over. Working this backward, we can determine the number of items required for a given sharing duration from the properties of the Poisson process.

Next, the allocation of the pooled inventory among locations must be decided in a way that minimizes the expected number of transshipments in the system to maximize efficiency. We first determine the expected demand satisfied in any system given an inventory level.

Lemma 1 *For a system with the demand rate λ, recovery rate μ, and current inventory level of I,*

$$E[\textit{demand satisfied by the system}] = \frac{\lambda}{\mu}\left(1 - \left(\frac{\lambda}{\mu+\lambda}\right)^{I}\right).$$

Using Lemma 1, given the inventory reserved for sharing, we can find the total demand satisfied by the system and by each hospital during the sharing period. Assuming a hospital that transfers an item cannot request a transshipment, we can calculate the expected total demand satisfied by the system using transshipments.

Theorem 3 *For a system of i hospitals, the expected demand satisfied by transshipment can be calculated as follows:*

$$E[\textit{satisfied with transshipment}] = \sum_i \frac{\lambda_i}{\mu}\left(\left(\frac{\lambda_i}{\mu+\lambda_i}\right)^{\phi_i} - \left(\frac{\sum_j \lambda_j}{\mu+\sum_j \lambda_j}\right)^{\Phi}\right).$$

Theorem 4 *The expected demand satisfied from hospitals' own inventories, namely the Type II service level, is calculated as follows:*

$$\beta_S = 1 - \left(\frac{\sum_j \lambda_j}{\mu+\sum_j \lambda_j}\right)^{\mathcal{B}-\Omega} \sum_i \left(\frac{\lambda_i}{\lambda_i+\mu}\right)^{\gamma_i\Omega} \frac{\lambda_i}{\sum_j \lambda_j}$$
$$- \sum_i \frac{\lambda_i}{\sum_j \lambda_j}\left(\left(\frac{\lambda_i}{\mu+\lambda_i}\right)^{\phi_i} - \left(\frac{\sum_j \lambda_j}{\mu+\sum_j \lambda_j}\right)^{\Phi}\right).$$

The inventory levels in each location that minimize the expected number of transshipments can be found using Theorem 3.

Table 1 Annual demand rates for hospitals in the system

Drug name	λ_1	λ_2	λ_3
Bleomycin	530	210	94
Doxorubicin 50 mg	940	380	9
Etoposide 50 mg	5980	440	14
Mitomycin 5 mg	760	430	–

Theorem 5 *For a system of hospitals, to maximize the Type II service level, the allocation of the inventory pool at locations k and m must satisfy the equality:*

$$\phi_m \ln\left(\frac{\lambda_m}{\mu+\lambda_m}\right) - \phi_k \ln\left(\frac{\lambda_k}{\mu+\lambda_k}\right) = \ln\left(\frac{\lambda_k \ln\left(\frac{\lambda_k}{\mu+\lambda_k}\right)}{\lambda_m \ln\left(\frac{\lambda_m}{\mu+\lambda_m}\right)}\right).$$

Similar to the allocation of the safety stock among hospitals, the inventory pool allocation can be determined by solving the set of linear equations derived from Theorem 5.

Our analyses and optimality conditions for the nonlinear optimization problems presented in Theorems 2 and 5 imply the following important analytical results:

- Among the hospitals of a health network, the optimal allocation mechanisms for safety stock (maximizing the Type I service level) and pooled inventory (maximizing the Type II service level) are the same.
- The optimal allocation of the pooled inventory and safety stock are not linearly proportional to the demand rates of hospitals.
- The sensitivity of the pooled inventory and safety stock allocation also differ with respect to shortage recovery rate.

4 Case study for Harris Health System

To measure the effects of the system parameters on the key performance indicators, namely the Type I and Type II service levels, we obtained real-life data about four chemotherapy drugs (Bleomycin, Doxorubicin 50 mg, Etoposide 50 mg, and Mitomycin 5 mg) from Harris Health System in Houston, TX. The aim of this numerical study is to observe the amounts to be allocated as pooled inventories and transshipment thresholds under different scenarios for demand, shortage duration, and total inventory.

According to the available data, the hospital chain consists of two hospitals and a small clinic, namely Lyndon B. Johnson Hospital, Ben Taub Hospital, and Smith Clinic. Table 1 shows the demand rates observed in these locations for each of the drugs. We evaluate the considered system under different values for the parameters related to shortage duration and total inventory. In terms of shortages, we analyze the system under exponentially distributed durations with varying expectations between two to twelve months (i.e., recovery rate of 1 to 6 per year). Moreover, to investigate the effect of total inventory (base stock level) on service levels, we assume three, six, nine, and 12 months of inventory within the system to be allocated. We measure the Type I and Type II service levels for the different proportions of the pooled inventory within the total inventory held, from nothing to the entire inventory in increments of 25% of total amount stocked. This study can be seen as a guideline for determining the allocation of total inventory, pooled amounts, and thresholds among hospitals when national drug shortages are taken into account.

We present our results for the different drugs by observing the various demand rates, recovery rates, months of inventory kept, and changing ratio of pooled inventory. We examine the case of Bleomycin for verification and illustrate how various parameters affect expected service levels. The pattern for the expected behaviors are very similar for the other drugs, thus are not presented. The results show that the Type I service level increases with the pooled inventory, whereas the Type II service level follows a convex curve as the allocation policy changes from no pooling to complete pooling and is slightly better in all cases when complete pooling policy is applied. As an example, we provide Fig. 2a for Bleomycin, considering an expected shortage duration of two months and an initial inventory of three months of demand. The plot reflects the changes in both service levels with respect to the proportion of the pooled inventory among the total inventory in the system. In fact, in all these cases, the service levels behave similarly when the ratio of pooled inventory increases, suggesting that for healthcare systems observing shortages, the best proactively pooling solution can be achieved by allowing hospitals to share their inventories to the fullest in terms of both the Type I and the Type II service levels. The increase in the expected service level can be subtle for some parameters as shown in Fig. 2b, but the system is always better off with complete pooling. The improvement in especially Type II service levels can be more dramatic for some drugs, as shown in the simulation results next.

The increased recovery rates and base stock levels unsurprisingly increase the service levels, converging to 100% on the abundance of inventory relative to the shortage duration. We illustrate the effect of recovery rate in Fig. 2c with a base stock of one years' demand, and the impact of base stock level in Fig. 2d, where the expected shortage duration is one year (recovery rate of 1 per year). In Fig. 2c and d, there is complete pooling, and Type I and II service levels are close to each other and overlap on the plot.

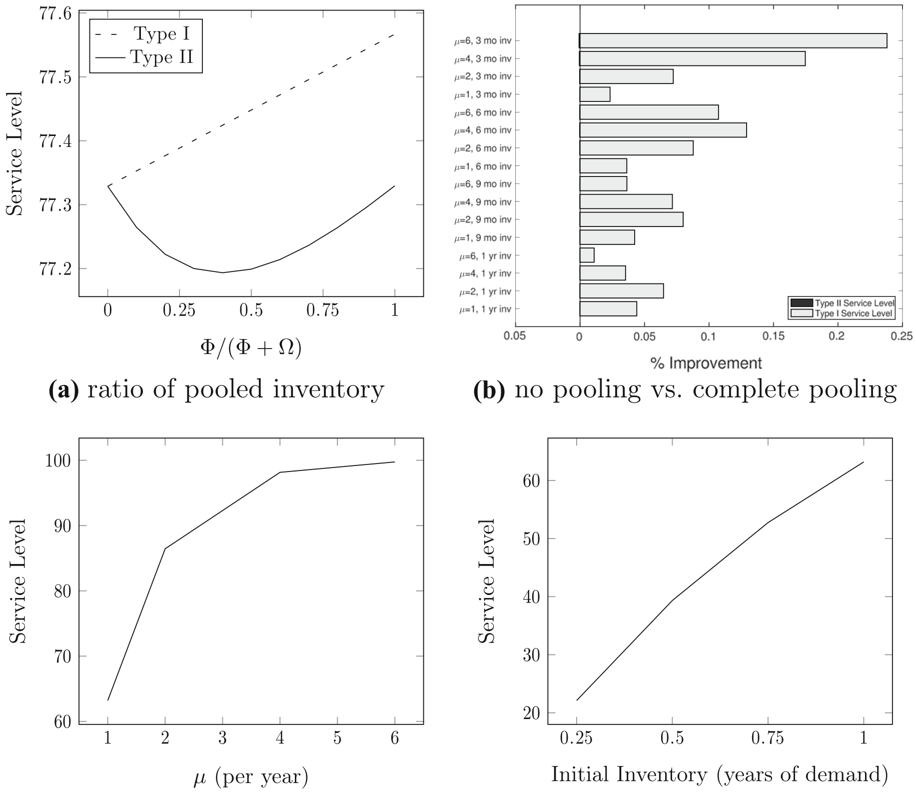
We have the following important observations as a result of our numerical analysis on all drugs using various parameters:

- Increased expected shortage duration results in a higher deviation from the allocation that is linearly proportional to the demand rates for both the pooled inventory and the total safety stock.
- Higher variance in hospitals' demand rates results in a higher deviation from the linearly proportional allocation for both the pooled inventory and the total safety stock.

The difference between proportional allocation and optimal allocation is not significant in quantity, especially considering the effect of rounding to the nearest integer. However, our experiments show the optimal allocation improves both service levels by as much as 1.6% compared to the proportional distribution.

Next, we focus on a simulation of these results for the 4 drugs presented, where we use varying levels of recovery rates from 1 to 6 per year, total stock as a fraction of annual demand in increments of 25%, and fraction of total stock pooled in increments of 25%. We randomly generate 5000 replications for each presented scenario[1]. We present simulation results using box plots, with the inclusion of a diamond marker to denote the average values and solid (dotted) horizontal lines to indicate the mean (median) of the leftmost box plot for reference.

[1] Our code is available at https://github.com/OEKundakcioglu/HospitalInventorySimulation.



(a) ratio of pooled inventory

(b) no pooling vs. complete pooling

(c) recovery rate

(d) base stock level

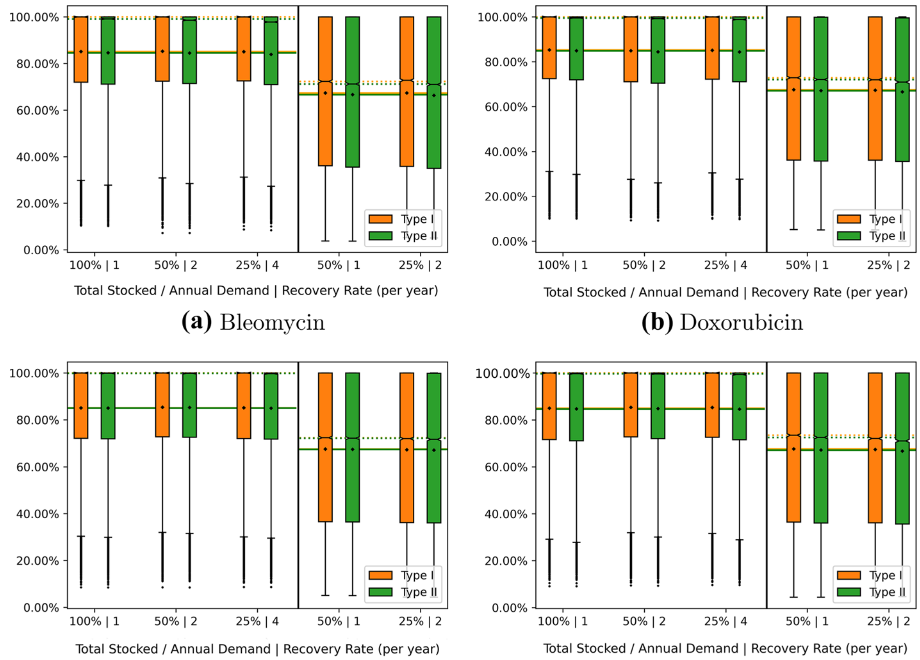
Fig. 2 Changes in the expected Type I and Type II service levels with respect to pooled inventory and problem parameters for Bleomycin

The service levels rise with increased inventory and decreased shortage durations, both in terms of expectation and experimentally. First, we use our simulation results to answer the following question: *under a constant total quantity stocked to shortage duration ratio, how do the service levels change?*

Figure 3 shows that for the same ratio, the Type I and Type II service levels behave similarly. On the left side of each plot, we respectively present one year, six months, and three months of inventory and expected shortage durations. On the right side, as another example, we show half the stocked inventory of what is demanded during the expected shortage duration, hence lower service levels. There is a modest decrease in the median and mean service levels towards smaller inventory and shorter duration in all cases due to increased risk of stock-outs.

Next, we explore the pooling decisions to see if complete or partial pooling is advantageous in simulated instances.

Figure 4 shows the service levels under various percentages of inventory reserved for pooling. Here, the expected shortage duration is one year, and three months of demand is available to stock. There is no particular trend in the results due to similar service level expectations for any fraction pooled (e.g., expected improvements of 0.1–0.2%). Still, pooling (even partially) improves the service levels in some cases, particularly for Doxorubicin, Mitomycin,



(a) Bleomycin **(b)** Doxorubicin

(c) Etoposide **(d)** Mitomycin

Fig. 3 Change in the service levels under two total quantity stocked to expected shortage duration ratios

and Etoposide with a small margin. What these drugs have in common is high variability among demand in different hospitals. Nevertheless, we do not derive a firm conclusion on pooling because all fractions presented can yield satisfactory service levels, as long as we allocate inventory optimally.

We present the service levels obtained with optimally allocated inventory (safety stock or pooled) in the analyses above. Next, we deviate from optimality using a parameter, which identifies the inventory to be reallocated. We sort hospitals using their demand rates, and starting with largest to smallest, we iteratively reallocate a fraction of inventory from next largest to next smallest. The larger the demand is, the larger the inventory reallocated. A larger parameter converges to a solution where all hospitals stock similar amounts, regardless of their needs. We report results on this re-adjustment separately for pooled inventory, as well as safety stock.

Figure 5 shows how the service levels change with a reallocation of 10, 20, and 30% of pooled inventory (P) and safety stock (S). Here expected shortage duration is one year, and the total amount stocked is six months of demand, half of which is pooled. For verification, the Type I service level should be unaffected by any reallocation in pooled inventory, which we can observe for all drugs. The presented deviation's effect is quite dramatic for Etoposide due to the considerable variation among hospital demands. One of the most surprising observations is that *the optimal allocation of pooled inventory is more important for Type II service level than safety stock*. Although we find optimal safety stock and pooling quantity that respectively maximizes Type I and Type II service levels, considering the nature of the consumption, it is exciting to observe how small an effect safety stock has on Type II service level.

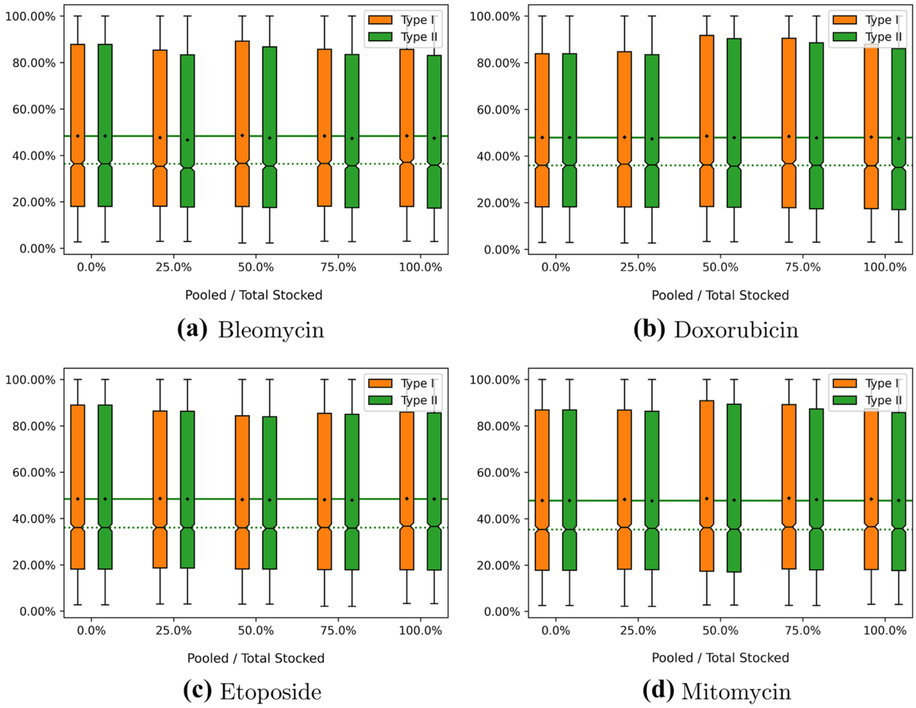


Fig. 4 Effect of pooling part of available inventory on the service levels

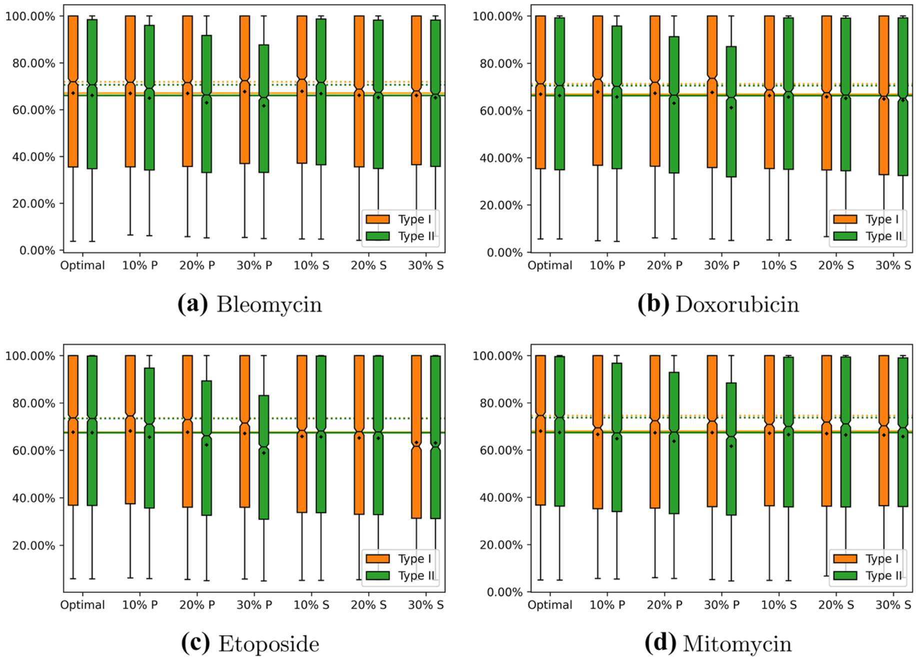
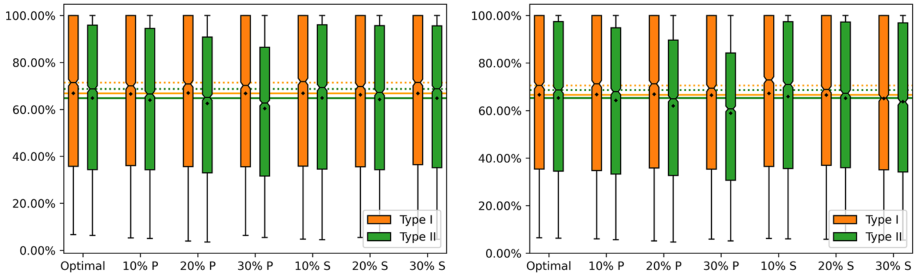


Fig. 5 Effect of deviation from the optimal allocation on the service levels

Table 2 Two synthetically created instances with six hospitals

	Demand Rates (per year)					
Low variability	50	60	70	150	160	170
High variability	50	60	70	350	360	370



(a) Low variability in demand

(b) High variability in demand

Fig. 6 Results on networks of six hospitals

To further support the above observation, we synthetically create two more datasets for six hospitals. All other parameters being the same, demands have the rates in Table 2.

Figure 6 supports the above result, showing the importance of optimal allocation of pooled inventory on Type II service level. Optimality of safety stock has minimal impact on the service levels, especially in the low variability scenario.

5 Conclusion

In this study, we analyze the inventory allocation for a healthcare system during a shortage that lasts for an uncertain length. We investigate two quantities of interest for critical drugs during such a shortage: (i) the expected rate of demand satisfied and (ii) the expected rate of demand satisfied without transshipment and the associated cost/wait/clinical drawbacks. Our analyses show that each hospital's inventory level is not directly proportional to its demand rate in contrast to common wisdom. Further, the theory shows that inventories should be completely pooled to maximize the two service levels. We provide a numerical study with real-life data obtained from a hospital network, showing that optimal allocation always provides acceptable service levels, regardless of how many items are pooled. These numerical experiments allow us to provide managerial insights into the sensitivity of inventory allocations and shortage rates. To summarize:

- Among the hospitals of a health network, the optimal allocation mechanisms for safety stock (maximizing the Type I service level) and pooled inventory (maximizing the Type II service level) are the same.
- Optimal allocation of pooled inventory is far more critical than the distribution of safety stock.
- Complete pooling, as expected, maximizes the service levels for a health network, yet the margins are small.
- Networks should carefully assess service levels under a set of initial inventory scenarios.

- The inventory within the network can be distributed proportionally to hospitals' demand rates as a good approximation if total demand and the variance among hospitals are low.
- The service levels are slightly more challenging for shorter periods of shortage with proportionally smaller inventory.

We expect our stylized model and analytical results to provide input for decision support systems that consider multiple healthcare objectives. Aligned with this goal, there exist potential research directions that require further analyses to understand how such healthcare systems behave, when additional realistic assumptions are introduced. The effects of randomness, capacity considerations, and expiration dates might be worth investigating, especially for drugs with prolonged shortage periods. Within the scope of this study, we assume that transshipments occur one-by-one; however, in reality, systems mostly desire transshipments in bulk quantities. Therefore, decisions on transshipment quantities in systems suffering shortages are open to investigation. Another issue worth a deeper analysis is the effect of the transshipment/order lead time on performance metrics.

Acknowledgements This research was supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK) Grant 115M564.

Appendix

Proofs of the Theorems and Lemmas

A Proof of Theorem 1

N hospitals consume $\mathcal{B}-\Omega$ items together first. They satisfy the observed demand from their own inventory. Any demand when a hospital reaches an inventory level of zero becomes lost.

The random variable that denotes the time until the next patient arrival at hospital i is denoted by T_i, where $T_i \sim \text{Exponential}(\lambda_i)$. The random variable T_d denotes the time until the next demand occurrence in the system and $T_d \sim \text{Exponential}(\sum_j \lambda_j)$. W denotes the time until the supplier becomes available and $W \sim \text{Exponential}(\mu)$.

We define Ω as the total transshipment thresholds, and the threshold for hospital i is defined as $\gamma_i \Omega$, where $\sum_i \gamma_i = 1$.

The total expected lost sales during a shortage can be written using conditioning as follows:

$$\begin{aligned} E[L] &= P(T_d = \min(T_d, W))^{\mathcal{B}-\Omega} \sum_i P(T_i = \min(T_i, W))^{\gamma_i \Omega} E \\ &\quad [\text{arrivals with rate } \lambda_i \text{ before shortage ends}] \\ &= \left(\frac{\sum_j \lambda_j}{\mu + \sum_j \lambda_j} \right)^{\mathcal{B}-\Omega} \sum_i \left(\frac{\lambda_i}{\lambda_i + \mu} \right)^{\gamma_i \Omega} \frac{\lambda_i}{\mu}. \end{aligned} \tag{5}$$

Service level, denoted by α_S, can be calculated as follows:

$$\alpha_S = 1 - \frac{E[L]}{E[total\ demand\ during\ shortage]} = 1 - \frac{\left(\frac{\sum_j \lambda_j}{\mu + \sum_j \lambda_j} \right)^{\mathcal{B}-\Omega} \sum_i \left(\frac{\lambda_i}{\lambda_i + \mu} \right)^{\gamma_i \Omega} \frac{\lambda_i}{\mu}}{\frac{\sum_j \lambda_j}{\mu}} \tag{6}$$

$$= 1 - \left(\frac{\sum_j \lambda_j}{\mu + \sum_j \lambda_j} \right)^{\mathcal{B}-\Omega} \sum_i \left(\frac{\lambda_i}{\lambda_i + \mu} \right)^{\gamma_i \Omega} \frac{\lambda_i}{\sum_j \lambda_j}. \tag{7}$$

B Proof of Theorem 2

We can remove the $\sum_i \gamma_i = 1$ constraint by substituting the N-th hospital's threshold with $\gamma_N = 1 - \sum_{j=1}^{N-1} \gamma_i$. Keeping in mind $0 \leq \gamma_i \leq 1$ for all i, the Type I service level can be written as

$$\alpha_S = 1 - \left(\frac{\sum_j \lambda_j}{\mu + \sum_j \lambda_j} \right)^{\mathcal{B}-\Omega} \left(\sum_{i=0}^{N-1} \left(\frac{\lambda_i}{\lambda_i + \mu} \right)^{\gamma_i \Omega} \frac{\lambda_i}{\sum_j \lambda_j} \right.$$
$$\left. + \left(\frac{\lambda_N}{\lambda_N + \mu} \right)^{(1-\sum_{j=1}^{N-1} \gamma_i)\Omega} \frac{\lambda_N}{\sum_j \lambda_j}. \right. \tag{8}$$

To find the globally optimal γ_i values, we find the Jacobian, where the i-th entry is

$$\frac{\partial \alpha_S}{\partial \gamma_i} = - \left(\frac{\sum_j \lambda_j}{\mu + \sum_j \lambda_j} \right)^{\mathcal{B}-\Omega} \left[\Omega \ln \left(\frac{\lambda_i}{\lambda_i + \mu} \right) \frac{\lambda_i}{\sum_j \lambda_j} \left(\frac{\lambda_i}{\lambda_i + \mu} \right)^{\gamma_i \Omega} \right.$$
$$\left. - \Omega \ln \left(\frac{\lambda_N}{\lambda_N + \mu} \right) \frac{\lambda_N}{\sum_j \lambda_j} \left(\frac{\lambda_N}{\lambda_N + \mu} \right)^{(1-\sum_{j=1}^{N-1} \gamma_j)\Omega} \right]. \tag{9}$$

Equating the Jacobian to zero yields the following condition for critical points:

$$\lambda_i \ln \left(\frac{\lambda_i}{\lambda_i + \mu} \right) \left(\frac{\lambda_i}{\lambda_i + \mu} \right)^{\gamma_i \Omega} = \lambda_N \ln \left(\frac{\lambda_N}{\lambda_N + \mu} \right) \left(\frac{\lambda_N}{\lambda_N + \mu} \right)^{(1-\sum_{j=1}^{N-1} \gamma_j)\Omega} \quad \forall i. \tag{10}$$

Therefore, for hospitals k and m, at the critical point, we have

$$\lambda_k \ln \left(\frac{\lambda_k}{\lambda_k + \mu} \right) \left(\frac{\lambda_k}{\lambda_k + \mu} \right)^{\gamma_k \Omega} = \lambda_m \ln \left(\frac{\lambda_m}{\lambda_m + \mu} \right) \left(\frac{\lambda_m}{\lambda_m + \mu} \right)^{\gamma_m \Omega}, \tag{11}$$

which leads to

$$\gamma_m \ln \left(\frac{\lambda_m}{\mu + \lambda_m} \right) - \gamma_k \ln \left(\frac{\lambda_k}{\mu + \lambda_k} \right) = \frac{1}{\Omega} \ln \left(\frac{\lambda_k \ln \left(\frac{\lambda_k}{\mu + \lambda_k} \right)}{\lambda_m \ln \left(\frac{\lambda_m}{\mu + \lambda_m} \right)} \right). \tag{12}$$

Next, we use the Hessian to prove the concavity of the Type I service level function, where

$$\frac{\partial^2 \alpha_S}{\partial \gamma_i^2} = - \left(\frac{\sum_j \lambda_j}{\mu + \sum_j \lambda_j} \right)^{\mathcal{B}-\Omega}$$
$$\frac{\Omega^2 \left(\lambda_i \ln^2 \left(\frac{\lambda_i}{\lambda_i + \mu} \right) \left(\frac{\lambda_i}{\lambda_i + \mu} \right)^{\gamma_i \Omega} + \lambda_N \ln^2 \left(\frac{\lambda_N}{\lambda_N + \mu} \right) \left(\frac{\lambda_N}{\lambda_N + \mu} \right)^{(1-\sum_{j=1}^{N-1} \gamma_j)\Omega} \right)}{\sum_j \lambda_j} \tag{13}$$

and

$$\frac{\partial^2 \alpha_S}{\partial \gamma_i \partial \gamma_l} = -\left(\frac{\sum_j \lambda_j}{\mu + \sum_j \lambda_j}\right)^{\mathcal{B}-\Omega} \frac{\Omega^2 \left(\lambda_N \ln^2\left(\frac{\lambda_N}{\lambda_N+\mu}\right)\left(\frac{\lambda_N}{\lambda_N+\mu}\right)^{(1-\sum_{j=1}^{N-1}\gamma_j)\Omega}\right)}{\sum_j \lambda_j}. \tag{14}$$

At the critical point, all off-diagonal entries of the Hessian are negative constants, whereas each diagonal entry depends on the individual demand rates; however, they are definitely less than the off-diagonal constants. It can easily be shown that all the eigenvalues of this matrix are negative, guaranteeing that the Hessian is negative definite and α_S is concave. Thus, the critical point is the global maximizer for α_S.

C Proof of Lemma 1

For a system with a demand rate of λ, a recovery rate of μ, and a total inventory of ϕ, the expected demand satisfied can be written as

$$\begin{aligned} E[\text{demand satisfied}] &= \sum_{j=0}^{\phi-1} j \times P(\text{Demand} = j) + \phi \times P(\text{Demand} \geq \phi) \\ &= \sum_{j=0}^{\phi-1} j \left(\frac{\lambda}{\mu+\lambda}\right)^j \frac{\mu}{\mu+\lambda} + \phi\left(1 - \sum_{j=0}^{\phi-1}\left(\frac{\lambda}{\mu+\lambda}\right)^j \frac{\mu}{\mu+\lambda}\right) \\ &= \frac{\mu\lambda}{(\mu+\lambda)^2} \sum_{j=0}^{\phi-1} j \left(\frac{\lambda}{\mu+\lambda}\right)^{j-1} + \phi - \phi\left(1 - \left(\frac{\lambda}{\mu+\lambda}\right)^{\phi}\right). \end{aligned} \tag{15}$$

Knowing that

$$\sum_{j=0}^{\phi-1} j p^{j-1} = \sum_{j=0}^{\phi-1} \frac{\partial p^j}{\partial p} = \frac{\partial}{\partial p} \sum_{j=0}^{\phi-1} p^j = \frac{\partial}{\partial p} \frac{1-p^{\phi}}{1-p} = \frac{-\phi p^{\phi-1}(1-p) + 1 - p^{\phi}}{(1-p)^2}, \tag{16}$$

the expected demand satisfied can be written as follows:

$$\begin{aligned} E[\text{demand satisfied}] &= \frac{\mu\lambda}{(\mu+\lambda)^2} \frac{-\phi\left(\frac{\lambda}{\mu+\lambda}\right)^{\phi-1}\left(\frac{\mu}{\mu+\lambda}\right) + 1 - \left(\frac{\lambda}{\mu+\lambda}\right)^{\phi}}{\left(\frac{\mu}{\mu+\lambda}\right)^2} \\ &\quad + \phi - \phi\left(1 - \left(\frac{\lambda}{\mu+\lambda}\right)^{\phi}\right) \\ &= \frac{\lambda}{\mu}\left(1 - \left(\frac{\lambda}{\mu+\lambda}\right)^{\phi}\right). \end{aligned} \tag{17}$$

D Proof of Theorem 3

Let us introduce the following notation for this part. First, the *random variables*:

D: demand during shortages,

H: demand during sharing,

P: demand satisfied from hospitals' pooled inventories, ϕ_i, directly,
T: demand satisfied by transshipment,
S: demand satisfied from each hospital's safety stock, $\gamma_i \Omega$,
L: unsatisfied demand.
For a system of hospitals, demand during shortages can be calculated as

$$D = P + T + S + L$$
$$E[D] = E[P] + E[T] + E[S] + E[L].$$

Alternatively, we can use demand during sharing

$$E[H] = E[P] + E[T].$$

From Lemma 1, since we know that

$$E[H] = \frac{\sum_i \lambda_i}{\mu} \left(1 - \left(\frac{\sum_i \lambda_i}{\mu + \sum_i \lambda_i}\right)^{\sum_i \phi_i}\right)$$

and

$$E[P] = \sum_i \frac{\lambda_i}{\mu} \left(1 - \left(\frac{\lambda_i}{\mu + \lambda_i}\right)^{\phi_i}\right),$$

the expected demand satisfied by transshipment can be calculated as follows:

$$E[T] = \sum_i \frac{\lambda_i}{\mu} \left(\left(\frac{\lambda_i}{\mu + \lambda_i}\right)^{\phi_i} - \left(\frac{\sum_j \lambda_j}{\mu + \sum_j \lambda_j}\right)^{\sum_j \phi_j}\right). \tag{18}$$

E Proof of Theorem 5

The Type II service level can be maximized when the expected number of transshipments ($E[T]$) is minimized because demand during shortages is not affected by the allocation of the pooled inventory. Thus, we can find the optimal allocation of the inventory pool to minimize the number of transshipments within the system by taking the derivative of Equation (18) with respect to ϕ_i:

$$\frac{\partial E[T]}{\partial \phi_i} = \frac{\lambda_i}{\mu} \ln\left(\frac{\lambda_i}{\mu + \lambda_i}\right)\left(\frac{\lambda_i}{\mu + \lambda_i}\right)^{\phi_i} - \frac{\sum_j \lambda_j}{\mu} \ln\left(\frac{\sum_j \lambda_j}{\mu + \sum_j \lambda_j}\right)\left(\frac{\sum_j \lambda_j}{\mu + \sum_j \lambda_j}\right)^{\sum_j \phi_j}. \tag{19}$$

Using (19), the obtained Jacobian can be set to zero, which yields

$$\lambda_i \ln\left(\frac{\lambda_i}{\mu + \lambda_i}\right)\left(\frac{\lambda_i}{\mu + \lambda_i}\right)^{\phi_i} = \sum_j \lambda_j \ln\left(\frac{\sum_j \lambda_j}{\mu + \sum_j \lambda_j}\right)\left(\frac{\sum_j \lambda_j}{\mu + \sum_j \lambda_j}\right)^{\sum_j \phi_j}, \tag{20}$$

where $\sum_j \phi_j = \Phi$. Therefore, for hospitals k and m, the critical points for the allocated inventory pool to minimize the expected number of transshipments must hold the equality below:

$$\lambda_k \ln\left(\frac{\lambda_k}{\mu+\lambda_k}\right)\left(\frac{\lambda_k}{\mu+\lambda_k}\right)^{\phi_k} = \lambda_m \ln\left(\frac{\lambda_m}{\mu+\lambda_m}\right)\left(\frac{\lambda_m}{\mu+\lambda_m}\right)^{\phi_m}$$

$$\phi_m \ln\left(\frac{\lambda_m}{\mu+\lambda_m}\right) - \phi_k \ln\left(\frac{\lambda_k}{\mu+\lambda_k}\right) = \ln\left(\frac{\lambda_k \ln\left(\frac{\lambda_k}{\mu+\lambda_k}\right)}{\lambda_m \ln\left(\frac{\lambda_m}{\mu+\lambda_m}\right)}\right). \tag{21}$$

References

1. Archibald, T.W.: Modelling replenishment and transshipment decisions in periodic review multilocation inventory systems. J. Op. Res. Soc. **58**(7), 948–956 (2007)
2. Archibald, T.W., Sassen, S.A.E., Thomas, L.C.: An optimal policy for a two depot inventory problem with stock transfer. Manag. Sci. **43**(2), 173–183 (1997)
3. Axsäter, S.: Evaluation of unidirectional lateral transshipments and substitutions in inventory systems. Eur. J. Op. Res. **149**(2), 438–447 (2003a)
4. Axsäter, S.: A new decision rule for lateral transshipments in inventory systems. Manag. Sci. **49**(9), 1168–1179 (2003b)
5. Cheong, T.: Joint inventory and transshipment control for perishable products of a two-period lifetime. Int. J. Adv. Manuf. Technol. **66**(9–12), 1327–1341 (2013)
6. Deloitte. Global Health Care Sector Outlook. 2017. URL https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Life-Sciences-Health-Care/gx-lshc-2017-health-care-outlook-infographic.pdf
7. FDA. Drug Shortages, 2018. URL https://www.accessdata.fda.gov/scripts/drugshortages/default.cfm#B. Retrieved: 2021/07/12 17:45:11
8. Feng, P., Fung, R.Y.K., Wu, F.: Preventive transshipment decisions in a multi-location inventory system with dynamic approach. Comput. Ind. Eng. **104**, 1–8 (2017)
9. Glazebrook, K., Paterson, C., Rauscher, S., Archibald, T.: Benefits of hybrid lateral transshipments in multi-item inventory systems under periodic replenishment. Prod. Op. Manag. **24**(2), 311–324 (2015)
10. Grahovac, J., Chakravarty, A.: Sharing and lateral transshipment of inventory in a supply chain with expensive low-demand items. Manag. Sci. **47**(4), 579–594 (2001)
11. Herer, Y.T., Rashit, A.: Lateral stock transshipments in a two-location inventory system with fixed and joint replenishment costs. Naval Res. Logist. **46**(5), 525–547 (1999)
12. Herer, Y.T., Tzur, M.: The dynamic transshipment problem. Naval Res. Logist. **48**(5), 386–408 (2001)
13. Hochmuth, C.A., Köchel, P.: How to order and transship in multi-location inventory systems: The simulation optimization approach. Int. J. Product. Econ. **140**(2), 646–654 (2012)
14. Hu, J., Watson, E., Schneider, H.: Approximate solutions for multi-location inventory systems with transshipments. Int. J. Prod. Econ. **97**(1), 31–43 (2005)
15. Kukreja, A., Schmidt, C.P.: A model for lumpy demand parts in a multi-location inventory system with transshipments. Comput. Op. Res. **32**(8), 2059–2075 (2005)
16. Minner, S., Silver, E.A.: Evaluation of two simple extreme transshipment strategies. Int. J. Prod. Econ. **93**, 1–11 (2005)
17. Nasr, W.W., Salameh, M.K., Moussawi-Haidar, L.: Transshipment and safety stock under stochastic supply interruption in a production system. Comput. Ind. Eng. **63**(1), 274–284 (2012)
18. Nicholson, L., Vakharia, A.J., Erenguc, S.S.: Outsourcing inventory management decisions in healthcare: Models and application. Eur. J. Op. Res. **154**(1), 271–290 (2004)
19. OECD. Health at a Glance 2017: OECD Indicators. 2017. URL http://www.oecd.org/health/health-systems/health-at-a-glance-19991312.htm
20. Olsson, F.: Optimal policies for inventory systems with lateral transshipments. Int. J. Prod. Econ. **118**(1), 175–184 (2009)
21. Olsson, F.: An inventory model with unidirectional lateral transshipments. Eur. J. Op. Res. **200**(3), 725–732 (2010)
22. Paterson, C., Kiesmüller, G., Teunter, R., Glazebrook, K.: Inventory models with lateral transshipments: A review. Eur. J. Op. Res. **210**(2), 125–136 (2011)
23. Paterson, C., Teunter, R., Glazebrook, K.: Enhanced lateral transshipments in a multi-location inventory system. Eur. J. Op. Res. **221**(2), 317–327 (2012)
24. Ramakrishna, K.S., Sharafali, M., Lim, Y.F.: A two-item two-warehouse periodic review inventory model with transshipment. Annals Op. Res. **233**(1), 365–381 (2015)

25. Saedi, S., Kundakcioglu, O.E., Henry, A.C.: Mitigating the impact of drug shortages for a healthcare facility: An inventory management approach. Eur. J. Op. Res. **251**(1), 107–123 (2016)
26. Tagaras, G., Vlachos, D.: Effectiveness of stock transshipment under various demand distributions and nonnegligible transshipment times. Prod. Op. Manag. **11**(2), 183–198 (2002)
27. van Wijk, A.C.C., Adan, I.J.B.F., van Houtum, G.J.: Approximate evaluation of multi-location inventory models with lateral transshipments and hold back levels. Eur. J. Op. Res. **218**(3), 624–635 (2012)
28. Yao,M., Minner,S.: Review of multi-supplier inventory models in supply chain management: An update. Technical report, Technische Universität München, (2017)
29. Zhao, H., Deshpande, V., Ryan, J.K.: Emergency transshipment in decentralized dealer networks: When to send and accept transshipment requests. Naval Res. Logist. **53**(6), 547–567 (2006)