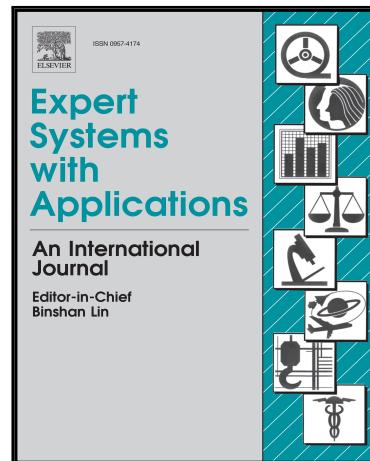


# Accepted Manuscript

A Multi-Objective Production Planning Problem with the Consideration of Time and Cost in Clinical Trials

HUI ZHAO , EDWARD HUANG , RUNLIANG DOU , KAN WU

PII: S0957-4174(19)30038-7  
DOI: <https://doi.org/10.1016/j.eswa.2019.01.038>  
Reference: ESWA 12439



To appear in: *Expert Systems With Applications*

Received date: 14 April 2018  
Revised date: 6 October 2018  
Accepted date: 10 January 2019

Please cite this article as: HUI ZHAO , EDWARD HUANG , RUNLIANG DOU , KAN WU , A Multi-Objective Production Planning Problem with the Consideration of Time and Cost in Clinical Trials, *Expert Systems With Applications* (2019), doi: <https://doi.org/10.1016/j.eswa.2019.01.038>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Highlights

- The proposed algorithm gives better solutions than the direct solution approach.
- The operational cost decreases with the increase of the production quantity.
- We quantify the tradeoff between the duration and the total cost of clinical trials.
- The tradeoff can be used to determine the optimal production quantity.

# A Multi-Objective Production Planning Problem with the Consideration of Time and Cost in Clinical Trials

HUI ZHAO

*School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore,*

[zhaohuabit@gmail.com](mailto:zhaohuabit@gmail.com)

EDWARD HUANG

*Systems Engineering and Operations Research Department, George Mason University, Fairfax, VA 22030,*

*USA, [chuang10@gmu.edu](mailto:chuang10@gmu.edu)*

RUNLIANG DOU

*College of Management and Economics, Tianjin University, Nankai District, Tianjin, China,*

[drl@tju.edu.cn](mailto:drl@tju.edu.cn)

KAN WU\*

*School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore,*

[kan626@gmail.com](mailto:kan626@gmail.com)

\* Corresponding author. Tel: +65-6790-5584; Fax: +65-6792-4062.

## A Multi-Objective Production Planning Problem with the Consideration of Time and Cost in Clinical Trials

Under increasingly challenging circumstances, pharmaceutical companies try to reduce the overproduction of clinical drugs, which is commonly seen in the pharmaceutical industry. When the overproduction is simply reduced without an efficient coordination of the inventories in the supply chain, the stock-outs at clinical sites and clinical trial delay can hardly be avoided. In this study, we propose a multi-objective model to optimize the production quantity, where the clinical trial duration and the total production and operational costs are minimized. The problem is formulated as a multi-stage stochastic programming model to capture the dynamic inventory allocation process in the supply chains. Since this problem's solving time and required memory can increase significantly with the increase of the stage and scenario numbers, the progressive hedging algorithm is applied as the solution approach in this paper. In the numerical experiments, we study this algorithm's performance and compare the solving efficiency with the direct solution approach. In addition, we identify the optimal production quantity of clinical drugs and give a discussion about the tradeoffs between the clinical trial delay and total cost.

**Keywords:** supply chain management; clinical trial; production and distribution planning; multi-stage stochastic programming; progressive hedging algorithm.

### 1. Introduction

The clinical trial is an important step in new drug development due to its high cost, long duration, and low success rate (Laínez, Schaefer, & Reklaitis, 2012; Roberts, Lynch, & Chabner, 2003; Shah, 2004). It tests a new drug on human beings for its safety, efficacy, and dose levels. The drug has the chance to get the approval of Food and Drug Administration (FDA) and enter the drug market of the United States only if it successfully passes all three phases of the clinical trial. The trial will be halted immediately once the drug fails in any phase. Only around five of 5,000 to 10,000 discovered compounds can be selected to take the clinical trials (Profile, 2010). One of the five can finally get into the drug market (DiMasi, Hansen, & Grabowski, 2003). On average, 37% of the total expense (A. Fleischhacker, 2009) and more than 70% of the total time (6 to 8 years (Kaitin & DiMasi, 2011; Shah, 2004) in the 8 to 10 years (Sundaramoorthy, Evans, & Barton, 2012)) in the new drug development are spent on clinical trials.

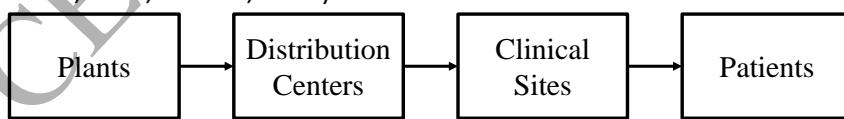
The pharmaceutical companies are facing increasingly challenging circumstances and striving to reduce the cost. The new drug development becomes more and more diverse and complicated (Abdelkafi, Beck, David, Druck, & Horoho, 2009). The investment in research and development (R&D) is growing, but the number of new drugs approved by FDA has declined (Martinez & Goldstein, 2007). A new drug's

effective patent life has also been shortened due to the increase of its development time (Shah 2004). Hence, the profit margin has started to decline and the pharmaceutical companies strive to reduce cost (A. Fleischhacker, 2009; Martinez & Goldstein, 2007).

The reduction of the clinical trial duration has the potential to alleviate the patients' pains earlier or extend more patients' lives, and bring more profit to pharmaceutical companies. According to FDA's regulations, each new drug is granted 20 years of patent protection and the clinical trial duration is included in the patent life (Grabowski & Vernon, 2000). While a drug can make a significant profit in its patent life, the profit will be largely reduced after the patent expiration due to the rise of generic substitutes at much lower prices (Martinez & Goldstein, 2007). During the patent life, the gross margin for a typical drug can be 90% to 95% (Martinez & Goldstein, 2007) and a blockbuster can bring \$1 billion or more in annual sales revenue (Mullin, 2011). The price may drop to a value close to the production cost after the patent expires (Martinez & Goldstein, 2007). On average, the trial delay of a single day can cause a loss of more than \$1 million revenue for a typical drug (Clemento, 1999). Hence, if a clinical trial can be completed earlier, the drug's commercialization time under patent protection may increase and the profit may increase accordingly. As a result, the clinical trial duration should be minimized to gain the most profit.

The clinical trial duration is the time spent on collecting enough valid trial results. Prior to the commencement of each clinical trial, the clinical team determines the patient accrual goal and implementation regulations, such as the dosage levels, treatment interval (the period length between two treatments, e.g. one week), and treatment period (the number of periods required for a patient to participate in the trial, e.g. several weeks). The patient accrual goal is the number of valid trial results required in this clinical trial. If a patient takes the trial according to the implementation regulations, the trial data collected on him/her is deemed valid. Otherwise, the trial data will be invalid if the patient drops out the trial due to the stock-outs at the clinical site, the dissatisfaction of the trial because of no observed improvement, loss of interest or some other reasons.

Clinical trials are conducted at clinical sites. As shown in Figure 1, a clinical trial supply chain consists of plants, distribution centers (DCs), clinical sites, and patients. After the drugs are produced in plants, they are kept at DCs and distributed to clinical sites which are selected from existing hospitals and doctors' offices (A. Fleischhacker, Ninh, & Zhao, 2015). Patients are recruited at clinical sites to take clinical trials.



**Figure 1.** A clinical trial supply chain

The duration of a clinical trial cannot be minimized without an increase of the cost. Due to the patient arrivals at clinical sites which are stochastic, stock-outs may happen if the on-hand inventory is not sufficient. To avoid the clinical trial delay caused by stock-outs, clinical drugs are usually overproduced. However, the pharmaceutical companies are striving to reduce the overproduction in clinical trial supply chains. Rather than the traditional way which is based on chemistry, the current drugs are more developed based on biotechnology. Producing new drugs becomes more expensive. In addition, the leftover drugs cannot be used elsewhere and should be disposed of after an unsuccessful trial. While the

efficient coordination of the inventories in the supply chain can reduce the duration of a clinical trial in case of a smaller production quantity, a greater operational cost may be required in the inventory transshipments among DCs and clinical sites.

The production planning problem in clinical trial supply chains is a multi-objective problem which minimizes the time and cost of clinical trials. The time is the clinical trial duration and the cost consists of the production and inventory allocation costs. Since clinical drugs have no salvage cost and all leftovers should be disposed of after an unsuccessful trial, the production cost should be considered in this problem. The inventory allocation results should also be considered in this problem, since they depend on the production quantity and affect both the duration and operational cost in clinical trials. In addition, since the clinical trial delay can be costly (for a typical drug, more than \$1 million revenue can be lost for a single day's loss in clinical trials (Clemento, 1999)), the duration minimization has a higher priority than the cost minimization in clinical trial supply chains.

In this paper, we consider the production planning problem in clinical trial supply chains. The contributions of this paper are three-fold. First, we propose a multi-objective model to find the optimal production quantity through minimizing the clinical trial duration and cost. Since the duration minimization is usually more important than the cost minimization in clinical trials, this multi-objective model can be solved by the lexicographic method (or the preemptive optimization) (Coello, Lamont, & Van Veldhuizen, 2007; Marler & Arora, 2004). In this method, all objectives are ranked in order of importance and the optimization is performed by considering one objective at a time, based on priorities. Second, due to the patient recruitment uncertainty at clinical sites, this model is formulated as a multi-stage stochastic programming model to capture the dynamic inventory allocation process in the supply chain. Since the computational complexity increases significantly with the increase of stage and scenario numbers in terms of the computational time and required memory, the progressive hedging algorithm is applied as the solution approach. Thirdly, numerical experiments are conducted to study the algorithm's performance and find the optimal production quantity of clinical drugs in a case study. Sometimes the production cost of certain clinical drugs is too expensive that the pharmaceutical companies would rather save the cost with the sacrifice of the minimal clinical trial duration. Hence, the trade-offs between the clinical trial duration and cost are also provided for decision makers in production planning.

The rest of this paper is organized as follows. We make a review of the studies about the production and distribution planning in clinical trial supply chains and other applications in Section 2. The production planning problem is described in detail in Section 3 and formulated as a mathematical model in Section 4. In addition, we discuss the solution approaches for the multi-stage stochastic programming models and describe the progressive hedging algorithm in detail in Section 5. Numerical results are discussed in Section 6. Concluding remarks and future works are given in Section 7.

## **2. Literature Review**

The production planning problem has been studied jointly with the inventory distribution in clinical trial supply chains. Y. Chen et al. (2012) proposed a decentralized model to optimize the production and distribution phases respectively and combined the solutions in a discrete event simulation to evaluate the supply chain's performance. The model was also extended to a centralized simulation-optimization one, where the  $(Q, r)$  policy and a risk pooling strategy were applied at the clinical sites and distribution centers respectively in the clinical trial supply chain (Ye Chen, Pekny, & Reklaitis, 2012). In addition, A. Fleischhacker et al. (2015) balanced the inventory overage cost against the total shipping cost to find the

optimal production quantity and inventory ordering policies at each location in the supply chain (including the regional distribution centers and clinical sites). Due to the lack of the knowledge of the entire system, decentralized models may not result in optimal solutions. The obtained solutions can also be far from the optimal ones if stationary inventory policies are applied for non-stationary stochastic demands (Tunc, Kilic, Tarim, & Eksioglu, 2011). However, due to the lack of reliable forecasts, the demand at each clinical site is highly uncertain in clinical trial supply chains and usually follows a non-stationary Poisson distribution (Y. Chen et al., 2012). In addition, only the cost but not the clinical trial duration minimization is considered in the production planning of clinical trial supply chains in these researches.

Production planning in clinical trial supply chains has also been studied with the consideration of the drugs' failure rates and demand realization in the literature. The clinical trial may fail in each of its three phases. By considering each drug's failure rate, Gatica, Papageorgiou, and Shah (2003) optimized the production schedule for a certain number of clinical drugs with the finite available manufacturing resources. A. J. Fleischhacker and Zhao (2011) demonstrated that for a clinical drug with deterministic demand, the Wagner-Whitin model could be generalized by incorporating the drug's failure risk and the salvage cost of leftover drugs to find the optimal production decisions. They also considered the demand uncertainty and determined the production time and quantities in a two-period production model based on the first period's demand realization (A. J. Fleischhacker & Zhao, 2013). In the production planning in clinical trial supply chains, none of these researches considers the duration minimization in clinical trials.

The production and distribution planning problem has been widely discussed in other applications. Erengüç, Simpson, and Vakharia (1999) categorized the decisions in such problems by supplier, plant and distribution stages, and reviewed the relevant studies in each stage. Beamon (1998) divided the models of the integrated supply chain design and analysis in four categories: deterministic analytic models, stochastic analytic models, economic models, and simulation models. In addition, Vidal and Goetschalckx (1997) and Mula, Peidro, Díaz-Madroñero, and Vicens (2010) reviewed the studies with strategic and tactical decisions respectively in production and distribution planning problems. Vidal and Goetschalckx (1997) discussed the mixed-integer programming strategic design models and identified the formulations' relevant factors, solution methods' characteristics, and computational times. Mula et al. (2010) selected the mathematical programming models with tactical and/or operational decisions and reviewed them from eight aspects, including the supply chain structure, decision level, modeling approach, purpose, shared information, limitations, novelty, and practical applications.

The tactical production and distribution planning models in the literature have similar (production and distribution) decision variables to our problem. In multi-period multi-retailer supply chains, most studies consider multiple products and minimize the total production, holding and transportation costs subject to certain production capacity and/or resource constraints (e.g. operation time, manpower, and machines). The shortage is disallowed in deterministic models (Boudia, Dauzere-Peres, Prins, & Louly, 2006; Chandra & Fisher, 1994; Jang, Jang, Chang, & Park, 2002) or minimized in stochastic cases (Gupta & Maranas, 2003; Lee & Kim, 2002). In addition, a single product model is considered by Boudia, Louly, and Prins (2008), where the inventory is distributed via a fleet of a limited number of vehicles to completely fulfill the deterministic demand at retailers with the minimal delivery cost. However, since the products in the studies above and clinical drugs are in different phases of new product development, i.e., in the

commercialization and R&D phases respectively, the problem settings in those studies are different from ours.

Since most studies above are about commercial products, they do not consider the overage cost but aim to minimize or eliminate the stock-out cost. For commercial products, the total demand in the entire planning horizon (selling season) is usually greater than the production capacity (Boudia et al., 2006; Boudia et al., 2008; Chandra & Fisher, 1994; Gupta & Maranas, 2003; Jang et al., 2002). In addition, the storage capacity at each retailer may not accommodate all products required in the entire planning horizon (Boudia et al., 2006; Boudia et al., 2008; Lee & Kim, 2002). As a result, such products might be produced very often. The inventory holding and overage costs are reduced or even eliminated consequently (Chandra & Fisher, 1994). Since a salvage cost may be obtained from the leftovers, the total overage cost is negligible for commercial products. On the other hand, failing to fulfill the customer demand may result in loss of immediate revenue or even market share (Gupta & Maranas, 2003; Lee & Kim, 2002). Hence, stock-outs are minimized in these studies.

On the contrary, clinical drugs are still in R&D phases, their overage should be reduced and the stock-out is not considered explicitly in our problem. Due to the significant potential revenue of clinical drugs, the clinical trial duration minimization is a key concern in new drug development. The duration will be minimal if no patient faces stock-outs at clinical sites before the completion of the clinical trial. In other words, stock-outs after the clinical trial's completion will not increase the clinical trial duration. Thus, we should only minimize the stock-outs within the clinical trial duration. Since the clinical trial duration is uncertain due to the patient recruitment uncertainty at clinical sites, the stock-outs to be minimized cannot be expressed explicitly. Rather than the stock-out cost, the overage cost is an important concern in clinical trial supply chains. Compared with the commercial products, the total amount of drugs required in a clinical trial is quite trivial. It can hardly exceed the production or storage capacity. In addition to the notable changeover cost and long supply lead time in drug production, as well as the negligible holding cost due to the small volumes, clinical drugs are usually produced in large batches. Due to the increasing complication in drug production, the production cost of clinical drugs can be considerable, especially for the emerging biotech drugs (Abdelkafi et al., 2009; Martinez & Goldstein, 2007). Furthermore, since the clinical drugs are in the R&D phase, the leftover cannot be used elsewhere and should only be disposed of after an unsuccessful trial. Hence, overproduction should be reduced in clinical trials.

In addition, while inventory policies are widely used in literature, they are analytically intractable in the distribution of clinical drugs due to the specific characteristics of clinical trial supply chains. While stock-out or backorder cost is minimized by conventional inventory policies, the duration minimization is the main concern in clinical trial supply chains. Although no stock-out means the minimal duration, the stock-outs out of the duration will not affect the duration. Even if the objective is to minimize the stock-out or backorder cost, stationary inventory policies may result in a much higher cost in case of non-stationary stochastic demand (Tunc et al., 2011). The time-varying ( $s, S$ ) policy and base-stock policy are proved to be the optimal inventory policies for non-stationary stochastic demand at a single stage or a pure serial system but not in distribution networks (Clark & Scarf, 1960; Scarf, 1959). In addition, optimal inventory policies with transshipments can only be obtained under very restricted conditions, such as single period (Dong & Rudi, 2004; Zhang, 2005), deterministic demand (Herer & Tzur, 2003), few retailers (e.g. two or three) (Herer & Rashit, 1999; Herer & Tzur, 2001; Robinson, 1990; Yang & Qin, 2007), or multiple identical retailers (Robinson, 1990).

### 3. Problem Description

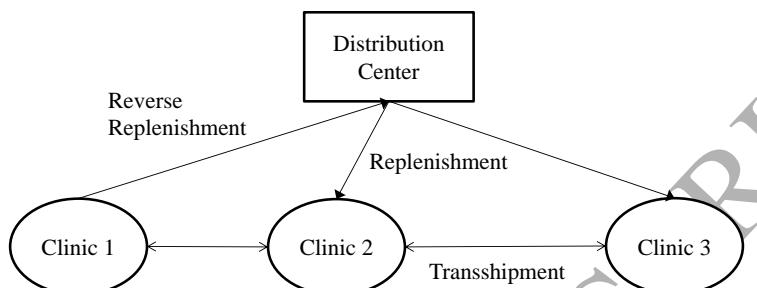
As mentioned in Section 1, clinical drugs are commonly overproduced to hedge the stock-outs at clinical sites. Since clinical trials may fail at each of the three phases and the required drug amount can be considerable in later phases (especially in Phase III), the production of clinical drugs is usually conducted more than once to avoid too much waste. Due to the long supply lead time and notable changeover cost, drugs are commonly produced in large batches and a substantial replenishment cycle exists between the arrivals of two consecutive production batches from a plant to a DC. Hence, the initial inventory at the beginning of a replenishment cycle is the sum of the volume of the production batch and the unused drugs from the previous cycle (if any). The total inventory in the supply chain decreases in a replenishment cycle and will only increase after the next production batch arrives. Due to the uncertainty from the dosage effectiveness and patient recruitment, the demands for clinical drugs at clinical sites are hard to predict (Y. Chen, Mockus, Orcun, & Reklaitis, 2012; Shah, 2004). Stock-outs may occur when the allocated inventory cannot meet the demand at some clinical sites. To reduce the trial delay caused by stock-outs at clinical sites, overproduction is commonly seen in clinical trial supply chains.

The overproduction of clinical drugs is to be reduced in pharmaceutical companies (Abdelkafi et al., 2009). By reducing the overage from an average of over 100% to under 50%, GSK saved \$120 million in 2009 (A. Fleischhacker et al., 2015). Bristol-Myers Squibb also estimated \$40 million could be saved by overage reduction (Powell, 2010). Since the traditional chemistry-based way of developing drugs becomes less efficient (produces fewer and fewer drugs), today's drug development tends to be based on biotechnology (Martinez & Goldstein, 2007). Biotech drugs are usually complicated and expensive to produce (Abdelkafi et al., 2009). Unlike commercial drugs, clinical drugs will be invalid and should be disposed of after an unsuccessful trial. Hence, pharmaceutical companies try to reduce the cost through reducing the overproduction of clinical drugs.

If the production quantity is simply reduced without the transshipments to efficiently coordinate the inventory in a supply chain, the stock-outs at clinical sites and clinical trial delay can hardly be avoided. When the production quantity is closer to the actual demand, stock-outs may occur at clinics even if there is still inventory in the supply chain. This would happen when decision makers want to save the setup (or delivery) cost by shipping large quantities to some clinics, and thus the DC has insufficient inventory to replenish other clinics at the later stage of a replenishment cycle. In this situation, if the inventory at a clinic with excess inventory can be shipped to clinics facing or going to face stock-outs, the clinical trial delay caused by stock-outs may be avoided. According to the discussion by Rudi, Kapur, and Pyke (2001), clinical drugs, which are characterized by long supply lead time, short selling seasons and high demand uncertainty, are eligible for and benefited from transshipments. Hence, transshipments can be applied to coordinate the inventories in clinical trial supply chains and improve the system performance.

As shown in Figure 2, transshipments can be longitudinal (cross echelons) or lateral (within the same echelon). With longitudinal transshipments, the inventory at a clinic can be recalled by its upper echelon (the DC) and redistributed to some other clinics. The recalling activity is called *reverse replenishment*. On the other hand, lateral transshipment consists of directly shipping the inventory at a clinic to some other clinics. Due to the double-blind requirement (see Appendix 1) and various regulations in different countries, reverse replenishment is commonly seen but lateral transshipment only exists between a few clinical sites within the same country. Due to the positive lead time of regular deliveries and stochastic

patient arrivals, the clinical trial delay may not be avoided. Hence, in addition to the regular deliveries, the emergency delivery mode is also applicable in each delivery route of the supply chain. Compared with the regular one, the emergency delivery mode has negligible lead time but more expensive cost. The regular and emergency transshipments are equivalent to the proactive and reactive transshipments in the literature (Paterson, Kiesmüller, Teunter, & Glazebrook, 2011). While proactive transshipments are used to arrange the inventory distribution before the stochastic demand is realized, reactive transshipments ship the inventory at a clinic with excess inventory to clinics facing or going to face stock-outs (Paterson et al., 2011).



**Figure 2.** The delivery routes in a clinical trial supply chain

While transshipments and emergency deliveries avoid the clinical trial delay caused by stock-outs, they inevitably increase the delivery cost. However, to avoid the transshipments, the inventory should be held at the DC and the DC has to deliver more frequently with smaller shipping quantities, which will also increase the delivery cost. On the other hand, to avoid the emergency deliveries, enough inventory should be held at clinical sites. Due to the limited inventory in the supply chain and stochastic demand, stock-outs may happen at some clinics but there still exist inventories at other clinics.

As a result, both the production and inventory allocation cost should also be considered jointly with the duration minimization in clinical trial supply chains. Since both the clinical trial duration and operational cost depend on the production quantity, we propose to find the optimal production quantity of clinical drugs in this production planning problem.

#### 4. Model Development

In this section, we present a mathematical formulation for the production planning problem in clinical trial supply chains. With the model assumptions given in Section 4.1, the problem is formulated as a deterministic model first and then extended to a stochastic programming model in Section 4.2.

##### 4.1. Assumptions

Clinical trials are commonly conducted in a multi-echelon supply chain shown in Figure 2. Prior to their commencements, some implementation details should be determined. The model assumptions are summarized as follows.

- (1) The planning horizon of a clinical trial is finite and no longer than the shelf life of clinical drugs.
- (2) Only one production is considered for each clinical trial.
- (3) The patient arrivals at each clinical site are independent and follow a non-stationary Poisson distribution.
- (4) Patients are fulfilled with the same clinical drugs in the whole treatment period.

- (5) The treatment period consists of several continuous periods and the dropout rate is identical in each period.
- (6) The forecast demand is rounded down if it is not an integer.
- (7) The delivery cost is composed of a fixed and a variable element in each route.
- (8) The lateral transshipment has a shorter lead time but a higher cost than the direct replenishment.
- (9) The emergency delivery has negligible lead time but a higher cost than the regular one in each route.
- (10) The holding cost is considered only at the clinical sites.

Since a clinical trial has a predetermined patient accrual goal, its planning horizon is finite. As shown on the website [clinicaltrials.gov](http://clinicaltrials.gov), the planning horizon is determined by a study start date and an estimated study completion date. The shelf life of drugs usually lasts for several years and can cover the entire planning horizon of a clinical trial. In addition to the long production time (about 6 to 9 months) (Laínez et al., 2012), clinical drugs are usually produced in large batches. We only consider one production for each clinical trial.

As discussed in the literature (Y. Chen et al., 2012; A. Fleischhacker et al., 2015), patient arrivals follow an independent non-stationary Poisson process. The forecast demand at each clinical site can be approximated from the record of patients visiting the site and/or the previous data of clinical trials for the same disease (A. Fleischhacker et al., 2015). Patients are required to take the same clinical drugs for the whole treatment period, which usually lasts for three to six weeks. In other words, the data collected on a patient is valid only if the patient has taken the clinical trial once a week and three to six continuous weeks in total. If patients fail to take the trials on time, they should drop out since their data is no longer valid. Similar to the assumption in the literature (Y. Chen et al., 2012), the dropout rate is assumed to follow a uniform distribution and have identical values in each period. In practice, it may decrease in the treatment period since patients who do not drop out early on are more likely to stay later. Due to the dropout rate, the forecast demand of existing patients for follow-up treatments may not be integers. To keep their integral nature, they are rounded down in the mathematical model in the following section.

In this study, two elements (a fixed and variable one) are included in the delivery cost for the sake of extension to general applications. In this problem, our objective is to achieve the minimal clinical trial duration with the minimal total production and inventory distribution costs. Due to the non-negligible lead time of the regular delivery and stochastic patient arrivals, transshipments and emergency deliveries can be required to achieve the minimal clinical trial duration in certain cases. Compared with the direct replenishment and regular deliveries, they have a shorter or negligible lead time but a higher cost. Due to the differences of local regulations, transshipment is only applicable between a few clinical sites. The holding cost is usually greater at the clinical sites than at the DC in terms of the limited storage spaces. Their differences are considered here as the holding costs at the clinical sites for simplicity.

#### **4.2. Production Planning Model**

In this section, the production planning problem in a multi-period multi-retailer clinical trial supply chain is formulated as a mathematical model, which considers the production planning and inventory allocation decisions jointly. The inventory allocation decisions follow a sequential process. With the patient recruitment process revealed gradually over time, they are determined and implemented correspondingly in each period of the planning horizon (consisting of  $T$  periods). This decision process has the following form (Shapiro, Dentcheva, & Ruszczyński, 2009).

decision ( $\theta_1$ ) ~ observation ( $d_1$ ) ~ decision ( $\theta_2$ ) ~ ... ~ decision ( $\theta_T$ ) ~ observation ( $d_T$ ).

If the observations are deterministic (e.g. patients are recruited before the clinical trial commencement and they will arrive by appointment), this decision process can be formulated as a deterministic model. Due to the patient recruitment uncertainty in clinical trials, the production planning problem here is a stochastic one and can be formulated as a multi-stage stochastic programming model. In the following sections, we first present the model under the deterministic setting and extend it to the stochastic setting subsequently.

#### 4.2.1. A Deterministic Model

We consider a production planning problem in a clinical trial supply chain, where the planning horizon consists of  $|T|$  time periods and the patient arrivals in each period are known. The objective is to determine the production quantity of clinical drugs with which the minimal clinical trial duration can be achieved with the minimal total cost. With the notations in Table 1, the model formulation is presented as follows.

**Table 1.** Summary of the notations in the deterministic model

Notation	Description
<b>Sets:</b>	
$I$	The set of clinical sites, $i \in I$
$T$	The set of time periods, $t \in T$ (Please see Appendix 2 for the explanation for the length requirement of the planning horizon.)
<b>Parameters:</b>	
$r$	The unit production cost of clinical drugs
$\alpha f_i$	The fixed delivery cost between the DC and clinical site $i$
$\alpha v_i$	The variable delivery cost between the DC and clinical site $i$
$l_i$	The delivery lead time between the DC and clinical site $i$
$h_i$	The holding cost at clinical site $i$
$d_{t,i}$	The number of forecast new arrival patients at clinical site $i$ in period $t$
$\varphi$	The average dropout rate in each period
$m$	The number of periods in the treatment period
$G$	The patient accrual goal
$M$	A sufficiently large number (E.g. one may define $M = G \sum_{k=1}^m (1 - \varphi)^{k-1} / (1 - \varphi)^{m-1}$ , please see

---

Appendix 3 for explanation.)

**Decision variables:**

$Q$	The production quantity
$f_t$	1 if all clinical sites' total number of patients fulfilled for $m$ periods until period $t$ is less than the patient accrual goal, 0 otherwise
$a_{t,i}$	1 if clinical site $i$ orders from the DC at the beginning of period $t$ , 0 otherwise
$x_{t,i}$	The replenishment quantity of clinical site $i$ at the beginning of period $t$
$ID_t$	The inventory level of the DC in period $t$
$IC_{t,i}$	The inventory level of clinical site $i$ at the beginning of period $t$
$Sd_{t,k,i}$	The number of patients who first arrive at clinical site $i$ in period $t$ and are fulfilled in the clinical trial for at least $k$ periods, $1 \leq k \leq m$ .
$\theta$	All decision variables in the model

---

$$\min Df_1(\theta) = \sum_{t \in T} f_t, \quad (1)$$

$$\min Df_2(\theta) = rQ + \sum_{t \in T, i \in I} (h_i IC_{t,i} + \alpha f_i a_{t,i} + \alpha v_i x_{t,i}), \quad (2)$$

$$s.t. \quad \sum_{\tau=1}^{\max(t-m+1,0)} \sum_i Sd_{\tau,m,i} \geq G(1 - f_t), \quad \forall t \in T \quad (3)$$

$$Sd_{t,1,i} \leq d_{t,i}, \quad \forall i \in I, \forall t \in T \quad (4)$$

$$Sd_{t,k,i} \leq (1 - \varphi) Sd_{t,(k-1),i}, \quad \forall i \in I, \forall k \in \{2, \dots, m\}, \forall t \in T \quad (5)$$

$$\sum_{\tau=\max(t-m+1,1)}^t Sd_{\tau,(t-\tau+1),i} \leq IC_{t,i}, \quad \forall i \in I, \forall t \in T \quad (6)$$

$$ID_0 = Q, \quad (7)$$

$$ID_t = ID_{(t-1)} - \sum_i x_{t,i}, \quad \forall t \in T \quad (8)$$

$$IC_{t,i} = x_{t,i}, \quad \forall i \in I, t = 1 \quad (9)$$

$$IC_{t,i} = IC_{(t-1),i} - \sum_{\tau=\max(t-m,1)}^{t-1} Sd_{\tau,(t-\tau),i} + x_{(t-l_i),i}, \quad \forall i \in I, \forall t \in T \setminus \{1\} \quad (10)$$

$$x_{t,i} \leq Ma_{t,i}, \quad \forall i \in I, \forall t \in T \quad (11)$$

$$x_{t,i}, ID_t, IC_{t,i}, Sd_{t,k,i} \in \mathbb{N}, \quad \forall i \in I, \forall k \in \{1, \dots, m\}, \forall t \in T \quad (12)$$

$$f_t, a_{t,i} \in \{0,1\}. \quad \forall i \in I, \forall t \in T \quad (13)$$

The production planning model has two objectives. The clinical trial duration  $Df_1(\theta)$  and the total production, holding and delivery costs  $Df_2(\theta)$  are minimized in (1) and (2) respectively. Under the deterministic settings, the minimal clinical trial duration can be achieved via the regular deliveries. Constraint (3) checks if the clinical trial's patient accrual goal has been reached within the first  $t$  periods. Constraints (4) to (6) guarantee the fulfilled patients at each clinical site in each period are no greater than the patient arrivals (new arrivals in (4) and existing patients for follow-up treatments in (5)) or the inventory level there. In Constraint (7) the initial inventory at the DC is set to the production quantity  $\mathcal{Q}$ . Constraints (8) to (10) are the inventory balance constraints at the DC and clinical sites. Constraint (11) defines the delivery cost's fixed element at each site. Constraints (12) and (13) enforce the non-negativity (natural numbers) or binary restrictions on the decision variables in the model.

Since minimizing the duration  $Df_1(\theta)$  is more important than minimizing the total cost  $Df_2(\theta)$  in clinical trials, this model can be solved by the lexicographic method. At first, we obtain the minimal clinical trial duration  $Df_1^*$  in the following program (14). By setting the clinical trial duration  $Df_1(\theta)$  to its minimal value  $Df_1^*$ , we find the optimal production quantity through minimizing the total production and operational costs  $Df_2(\theta)$  in (15).

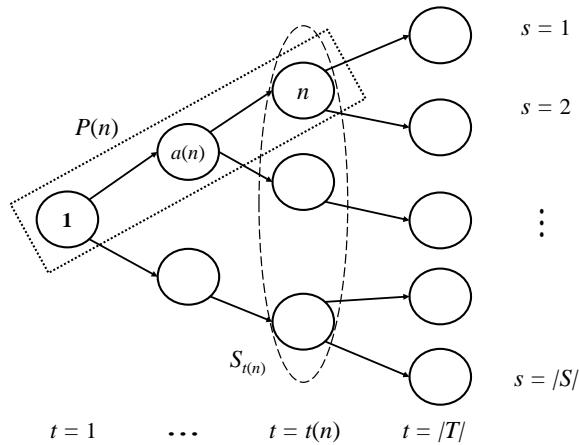
$$Df_1^* = \min\{Df_1(\theta), s.t. (3) \sim (13)\}. \quad (14)$$

$$Q^* = \arg \min_{\mathcal{Q}} \{Df_2(\theta), s.t. (3) \sim (13), Df_1(\theta) = Df_1^*\}. \quad (15)$$

The production planning problem is NP-hard. If we only consider single period, then  $f_1^*$  is equal to 0 to achieve the minimal duration in the first objective. Constraints (3) to (6) are equivalent to  $\sum_i Sd_i \geq G$  and  $Sd_i = \min(d_i, x_i)$ . Due to Constraint (11), we have  $Sd_i = \min(d_i, x_i) = \min(d_i a_i, x_i)$ . Since  $\sum_i Sd_i = \sum_i \min(d_i a_i, x_i) \leq \sum_i d_i a_i$ ,  $\sum_i Sd_i \geq G$  can be represented as  $\sum_i d_i a_i \geq \sum_i Sd_i \geq G$ . If we further set all of the production, holding, and variable delivery costs to zero, then the model of the second objective can be reduced to the following min-knapsack problem,  $\min\{\sum_i \alpha f_i a_i : s.t. \sum_i d_i a_i \geq G, a_i \in \{0,1\} \forall i \in I\}$ . The min-knapsack problem is NP-hard (Csirik, Frenk, Labb  , & Zhang, 1990). Unless P = NP, it is not possible to develop an algorithm which can solve this multi-period problem in polynomial time.

#### 4.2.2. Stochastic Programming Extensions

The patient recruitment in clinical trials is a stochastic process over the planning horizon. As discussed in Section 4.1, the patient arrivals follow an independent Poisson process. They have a finite number of possible realizations with specific probabilities in each period and can be represented as a scenario tree shown in Figure 3. This scenario tree consists of  $|N|$  nodes,  $|T|$  periods, and  $|S|$  scenarios. Each scenario represents a full evolvement of the stochastic process over the entire planning horizon, i.e., a scenario consists of the path from the root node to a leaf node (Huang, 2005). If the patient arrivals are independent in each period, then we have  $|S| = \prod_{t=1}^{|T|} |S(t)|$ , where  $|S(t)|$  represents the number of patient arrivals' possible values in period  $t$ .



**Figure 3.** A scenario tree

Each node  $n$  represents a possible state of the stochastic patient arrivals with associated probability  $p_n$  in a specific period. The root node represents the initial state of the stochastic process. Node  $a(n)$  is the direct ancestor of node  $n$ . A node's direct descendants are called its children. The set of all nodes along the path from the root node to the node  $n$  is denoted by  $P(n)$ . We say that  $n \in s$  if node  $n$  belongs to scenario  $s$ . The notations  $t(n)$  and  $S_t$  represent the time period of node  $n$  and the set of all nodes in period  $t$  respectively. The nodes in the tree are usually indexed in the increasing order of the time period (Huang, 2005), i.e.,  $t(1) \leq t(2) \leq \dots \leq t(|N|)$ . The root node has an index of 1. The indices of the nodes in the same period are not specified. In addition, the sum of the probabilities of the nodes in each period is equal to 1, i.e.,  $\sum_{n \in S_t} p_n = 1, \forall t$ . A node's probability is equal to the total probability of all its child nodes, i.e.,  $p_n = \sum_{m \in S_{t(n)+1}, a(m)=n} p_m, \forall n$ .

**Table 2.** Summary of the notations in the stochastic model

Notation	Description
<b>Sets:</b>	
$N$	The set of nodes in the scenario tree, $n \in N$
$S$	The set of forecast patient arrival scenarios, $s \in S$
<b>Parameters:</b>	
$\beta f_{i,j}$	The fixed regular transshipment cost between clinical site $i$ and $j$ , $i \neq j$
$\beta v_{i,j}$	The variable regular transshipment cost between clinical site $i$ and $j$ , $i \neq j$
$eaf_i$	The fixed emergency delivery cost between the DC and clinical site $i$
$eav_i$	The variable emergency delivery cost between the DC and clinical site $i$

---

$e\beta f_{i,j}$	The fixed emergency transshipment cost between clinical site $i$ and $j$ , $i \neq j$
$e\beta v_{i,j}$	The variable emergency transshipment cost between clinical site $i$ and $j$ , $i \neq j$
$lb_{i,j}$	The regular delivery lead time between clinical site $i$ and $j$ , $i \neq j$
$d_{n,i}$	The number of forecast new arrival patients at clinical site $i$ at node $n$
$p_n$	The probability of node $n$ with $\sum_{n \in S_t} p_n = 1$ for $\forall t \in T$
$d_{n,i}$	The number of forecast new arrival patients at clinical site $i$ at node $n$
<b>Decision variables:</b>	
$f_n$	1 if all clinical sites' total number of patients fulfilled for $m$ periods until node $n$ is less than the patient accrual goal, 0 otherwise
$a_{n,i}$	1 if clinical site $i$ orders from the DC at node $n$ using the regular delivery, 0 otherwise
$b_{n,i}$	1 if the DC is replenished from clinical site $i$ at node $n$ using the regular delivery, 0 otherwise
$c_{n,i,j}$	1 if there exists a regular transshipment from clinical site $i$ to clinical site $j$ at node $n$ , 0 otherwise
$ea_{n,i}$	1 if clinical site $i$ orders from the DC at node $n$ using the emergency delivery, 0 otherwise
$eb_{n,i}$	1 if the DC is replenished from clinical site $i$ at node $n$ using the emergency delivery, 0 otherwise
$ec_{n,i,j}$	1 if there exists an emergency transshipment from clinical site $i$ to $j$ at node $n$ , 0 otherwise
$x_{n,i}$	The regular replenishment quantity of clinical site $i$ at node $n$
$y_{n,i}$	The regular reverse replenishment quantity of clinical site $i$ at node $n$
$z_{n,i,j}$	The regular transshipment quantity from clinical site $i$ to clinical site $j$ at node $n$
$ex_{n,i}$	The emergency replenishment quantity of clinical site $i$ at node $n$
$ey_{n,i}$	The emergency reverse replenishment quantity of clinical site $i$ at node $n$
$ez_{n,i,j}$	The emergency transshipment quantity from clinical site $i$ to clinical site $j$ at node $n$
$ID_n$	The inventory level of the DC at node $n$
$IC_{n,i}$	The inventory level of clinical site $i$ at node $n$
$Sd_{n,k,i}$	The number of patients who first arrive at clinical site $i$ at node $n$ and are fulfilled in the clinical trial for at least $k$ periods, $1 \leq k \leq m$ .
$\theta$	All decision variables in the model

---

When the patient recruitment process is specified by a scenario tree and the inventory allocation decisions are determined correspondingly in each period, the production planning problem in clinical trial supply chains can be formulated as a multi-stage stochastic programming model. The production quantity is determined at the first stage, and the inventory allocation decisions can be determined in each of the multiple stages. In this model, we let  $a^k(n)$  represent the  $k$ -th ancestor of node  $n$ .  $a^0(n)$  is the node  $n$  itself.  $a(1)$  does not exist. A patient who has finished the whole treatment period (consisting of  $m$  periods) at node  $n$  participated the clinical trial at the node  $a^{(m-1)}(n)$  for the first time. With the notations in Table 2, the multi-stage stochastic model is presented as follows.

$$\min Sf_1(\theta) = \sum_{n \in N} p_n f_n, \quad (16)$$

$$\begin{aligned} \min Sf_2(\theta) = & rQ + \sum_{n \in N, i \in I} p_n (h_i IC_{n,i} + \alpha f_i(a_{n,i} + b_{n,i}) + \alpha v_i(x_{n,i} + y_{n,i}) + e\alpha f_i(ea_{n,i} + eb_{n,i}) + \\ & e\alpha v_i(ex_{n,i} + ey_{n,i}) + \sum_{j \in I, j \neq i} (\beta f_{i,j} c_{n,i,j} + e\beta f_{i,j} ec_{n,i,j} + \beta v_{i,j} z_{n,i,j} + e\beta v_{i,j} ez_{n,i,j})), \end{aligned} \quad (17)$$

$$s.t. \quad \sum_{\tau \in P(n), \tau \leq a^{(m-1)}(n)} \sum_i Sd_{\tau,m,i} \geq G(1 - f_n), \quad \forall n \in N \quad (18)$$

$$Sd_{n,1,i} \leq d_{n,i}, \quad \forall i \in I, \forall n \in N \quad (19)$$

$$Sd_{n,k,i} \leq (1 - \varphi) Sd_{n,(k-1),i}, \quad \forall i \in I, \forall k \in \{2, \dots, m\}, \forall n \in N \quad (20)$$

$$\sum_{k=1}^m Sd_{a^{(k-1)}(n),k,i} \leq IC_{n,i}, \quad \forall i \in I, \forall n \in N \quad (21)$$

$$ID_0 = Q, \quad (22)$$

$$ID_n = ID_{(n-1)} - \sum_i (x_{n,i} + ex_{n,i}) + \sum_i (y_{a^l(n),i} + ey_{n,i}), \quad \forall n \in N \quad (23)$$

$$IC_{n,i} = x_{n,i}, \quad \forall i \in I, n = 1 \quad (24)$$

$$IC_{n,i} = IC_{a(n),i} - \sum_{k=1}^m Sd_{a^k(n),k,i} + (x_{a^l(n),i} + ex_{n,i}) - (y_{n,i} + ey_{n,i}) + \sum_{j \in I, j \neq i} (z_{a^l b_{j,i}(n),j,i} - z_{n,i,j} + ez_{n,i,j} - ez_{n,i,j}), \quad \forall i \in I, \forall n \in N \setminus \{1\} \quad (25)$$

$$x_{n,i} \leq Ma_{n,i}, \quad \forall i \in I, \forall n \in N \quad (26)$$

$$ex_{n,i} \leq Mea_{n,i}, \quad \forall i \in I, \forall n \in N \quad (27)$$

$$y_{n,i} \leq Mb_{n,i}, \quad \forall i \in I, \forall n \in N \quad (28)$$

$$ey_{n,i} \leq Meb_{n,i}, \quad \forall i \in I, \forall n \in N \quad (29)$$

$$z_{n,i,j} \leq Mc_{n,i,j}, \quad \forall i, j \in I, i \neq j, \forall n \in N \quad (30)$$

$$ez_{n,i,j} \leq Mec_{n,i,j}, \quad \forall i, j \in I, i \neq j, \forall n \in N \quad (31)$$

$$ex_{n,i} = 0, y_{n,i} = 0, ey_{n,i} = 0, z_{n,i,j} = 0, ez_{n,i,j} = 0, \quad \forall i, j \in I, i \neq j, n = 1 \quad (32)$$

$$x_{n,i}, ex_{n,i}, y_{n,i}, ey_{n,i}, z_{n,i,j}, ez_{n,i,j}, ID_n, IC_{n,i}, Sd_{n,k,i} \in \mathbb{N}, \quad \forall i, j \in I, i \neq j, \forall k \in \{1, \dots, m\}, \forall n \in N \quad (33)$$

$$f_n, a_{n,i}, ea_{n,i}, b_{n,i}, eb_{n,i}, c_{n,i,j}, ec_{n,i,j} \in \{0,1\}. \quad \forall i, j \in I, i \neq j, \forall n \in N \quad (34)$$

Due to the patient recruitment uncertainty, this stochastic production planning model is constructed based on the scenario tree shown in Figure 3. All decision variables and constraints are indexed and presented for each node of the scenario tree. The two objectives (16) and (17) represent the minimization of the expected clinical trial duration and the total production and expected operational costs at all nodes of the scenario tree. Since the minimal clinical trial duration or cost may not be achieved with only the regular delivery under the stochastic setting, transshipments and emergency deliveries are also applied in the model.

## 5. Solution Approach

Since the computational complexity increases with the number of scenarios in stochastic programming models, decomposition and heuristic algorithms are widely studied as the solution approaches (Birge & Louveaux, 2011). By dividing a stochastic model into a number of smaller and more manageable sub-problems, decomposition algorithms can provide practical approaches for models with a large number of scenarios (Sen, 2005). The Benders decomposition (Birge, 1985) and progressive hedging algorithm (PHA) (Rockafellar & Wets, 1991) are two commonly used solution approaches in stochastic programming models. They decompose the entire stochastic model by stage and scenario respectively. Since the dual information of the programs in stages later than the first is applied in Benders decomposition, the convexity property is required in those stages (Carøe & Tind, 1998). In addition, although Benders decomposition can be applied in multi-stage stochastic models via a nested approach, the implementation is quite cumbersome in practice (Sherali & Fraticelli, 2002). Due to the binary restrictions on the decision variables in the model in Section 4, Benders decomposition is not applicable in this problem. On the other hand, the PHA is proved to converge theoretically in convex models and can be used as an effective heuristic in nonconvex ones (Fan & Liu, 2010; Watson & Woodruff, 2011). Parallel computing can also be applied in this algorithm to reduce the computational time significantly (Fan & Liu, 2010; Gade et al., 2016; Ryan, Wets, Woodruff, Silva-Monroy, & Watson, 2013). As shown in the numerical results, the computational time and required memory increase dramatically with the increase of the stage and scenario numbers in this problem. Hence, the PHA is applied as the solution approach in this paper.

The PHA is an effective method for solving multi-stage stochastic models, especially for those with integer decision variables in each stage (Gade et al., 2016). Based on the scenario tree, it decomposes a stochastic model into a number of independent deterministic models (scenario sub-problems) and solves these scenario sub-problems iteratively. Scenario sub-problems are obtained by fixing the random variables to the corresponding realizations in each scenario. All decision variables and constraints in the scenario sub-problems remain the same as those in the multi-stage stochastic models except the non-anticipativity constraints. With the non-anticipativity constraints, the decisions at a certain period should be identical among a set of scenarios if the random variables in those scenarios are indistinguishable up to the specific period (Rockafellar & Wets, 1991). Instead of these constraints, two penalty terms (a linear and a quadratic one) are added to the objective functions of the scenario sub-problems to gradually enforce the convergence of the decision variables' solutions restricted by the non-anticipativity constraints. Meanwhile, according to the solutions obtained in each scenario sub-problem, the coefficients of the linear term are updated in each iteration. This iterative solving process will not stop until the termination criteria are met.

The basic procedures of the PHA can be stated as follows. At the initial step, the scenario sub-problems are solved independently, where the vectors  $\chi_s$  and  $\chi_s^\nu$  represent all decision variables and the corresponding solutions in scenario  $s$  of the multi-stage stochastic model.  $c_s$  is a vector representing the corresponding parameters in the objective function. With the average value  $\bar{\chi}_n^\nu$  at node  $n$  in the  $\nu$ -th iteration, we calculate the errors (e.g. Euclidean distance) of the solutions in all scenarios and check if the convergence is achieved, where  $p_s$  is the probability of scenario  $s$ . If the error is no greater than the predetermined tolerance threshold  $\epsilon$ , it is regarded that the convergence is achieved, i.e., the non-anticipativity constraints are fulfilled. Otherwise, two penalty terms (a linear and a quadratic one) are added in the objective function, where  $w_s^\nu$  and  $\rho/2$  are the corresponding coefficients. In practice, the termination criteria can also depend on the total elapsed computational time due to the possible non-convergence (Løkketangen & Woodruff, 1996). This procedure is repeated for more iterations and  $w_s^\nu$  is updated in each iteration. Since both penalty terms depend on the parameter  $\rho$ , the algorithm's performance can be sensitive to  $\rho$ 's value (Watson & Woodruff, 2011). We will discuss  $\rho$ 's effects on the algorithm's convergence and this problem's solution quality in the numerical experiments.

---

#### Progressive Hedging Algorithm

1.  $\nu = 0, w_s^\nu = 0;$
  2.  $\chi_s^\nu = \arg \min_{\chi_s} (c_s \chi_s), \text{ for } \forall s \in S;$
  3.  $\bar{\chi}_n^\nu = \sum_{s \in S, n \in s} p_s \chi_s^\nu / p_n, \text{ for } \forall n \in N;$
  4. based on the scenario tree structure, the node-based value  $\bar{\chi}_n^\nu$  is transformed into the scenario-based value  $\bar{\chi}_s^\nu$  for  $\forall s \in S$ ;
  5.  $d^\nu = \sum_s p_s \|\chi_s^\nu - \bar{\chi}_s^\nu\|;$
  6. if  $d^\nu \leq \epsilon$ , then terminate; otherwise, go on;
  7.  $\nu = \nu + 1;$
  8.  $w_s^\nu = w_s^{(\nu-1)} + \rho (\chi_s^{(\nu-1)} - \bar{\chi}_s^{(\nu-1)}), \text{ for } \forall s \in S;$
  9.  $\chi_s^\nu = \arg \min_{\chi_s} (c_s \chi_s + w_s^\nu \chi_s + \rho/2 \|\chi_s - \bar{\chi}_s^{(\nu-1)}\|^2), \text{ for } \forall s \in S;$
  10. go to step 3.
- 

Cycling behavior may be detected in the algorithm for stochastic mixed-integer programs (Watson & Woodruff, 2011). Watson and Woodruff (2011) proposed a mechanism to detect and avoid cycles based on  $w_s^\nu$ .  $w_s^\nu$  is the coefficient of the linear term added in the objective function and updated in each iteration until convergence according to the solutions obtained in previous iterations. Due to the integrality restrictions, the decision variables' solutions and average values may stagnate for many iterations. As a result, the variation of  $w_s^\nu$  is much easier to capture than that of its corresponding decision variable's solution  $\chi_s^\nu$ . Hence,  $w_s^\nu$  is used in cycle detection. Assume  $w_{s,i}$  and  $\chi_{s,i}$  are the  $i$ -th element of the vectors  $w_s$  and  $\chi_s$  respectively. For computational efficiency, the values  $w_{s,i}^\nu$  of all

scenarios are jointly considered to generate a hash value for the variable  $\chi_{s,i}$  by applying a simple hashing scheme proposed in the literature (Woodruff & Zemel, 1993). If a variable's hash value in an iteration is identical to that obtained in a previous iteration, a cycle is detected. The solution of  $\chi_{s,i}$  is then fixed to  $\max_{s \in S} \chi_{s,i}^v$  to break the cycle. Since cycling behaviors are detected in the numerical experiments of Section 6, the cycle detection and avoidance mechanism is applied in this study. In addition, the technique of variable fixing (fixing the values of variables those have converged for a number of iterations) proposed in the literature (Watson & Woodruff, 2011) is also applied in the numerical experiment to accelerate the convergence.

## 6. Numerical Results

In this section, we present a case study of the production planning problem in clinical trial supply chains. Since the minimal clinical trial duration in each scenario sub-problem can be achieved as long as each patient arrival is fulfilled, it only depends on the patient arrivals and can be simply obtained in (14). As a result, only the cost minimization needs to be achieved by the PHA. Since the performance of the PHA is affected by the parameter  $\rho$ , with the parameters and implementation details in Section 6.1, the model is solved with different  $\rho$  values in Section 6.2 to check  $\rho$ 's effects on the algorithm's convergence and the solution quality. In addition, we solve the problem as a whole by CPLEX and using the PHA respectively for different stage and scenario numbers. The solving efficiency is compared in Section 6.3 in terms of the computational time, memory, and the optimality gap. Furthermore, by setting the production quantity to a number of possible values, we evaluate the expected total inventory allocation cost and identify the production quantity value with the minimal total cost, where the clinical trial duration is minimized as well. Since the inventory distribution may become less critical and the cost may be reduced when the clinical trial delay is allowed, we also provide the trade-offs between the clinical trial duration and cost for decision makers in production planning in Section 6.4.

### 6.1. Case study

In this study, we consider the production and distribution planning problem in a multi-echelon clinical trial supply chain consisting of a DC and 20 clinical sites. The patient accrual goal is set as 200, i.e., 200 pieces of valid trial data are required to complete the clinical trial. To get a valid trial data, a qualified patient should take the trial for three continuous periods (weeks). However, 20% of the patients may drop out in each period due to various reasons. The new patient arrivals at clinical sites are assumed to follow Poisson distributions (Y. Chen et al., 2012) and the arrival rates range from two to five in each period. The planning horizon consists of ten periods (the length requirement in Appendix 2 is satisfied). In addition, the lead times of regular deliveries from the DC to each clinical sites and between eligible clinical sites are set as one period respectively. Emergency deliveries are assumed to have negligible lead time. The fixed and variable regular delivery costs are randomly generated from 20 to 60 and 2 to 6 respectively. The emergency delivery is assumed to have a double cost of the regular one in each route.

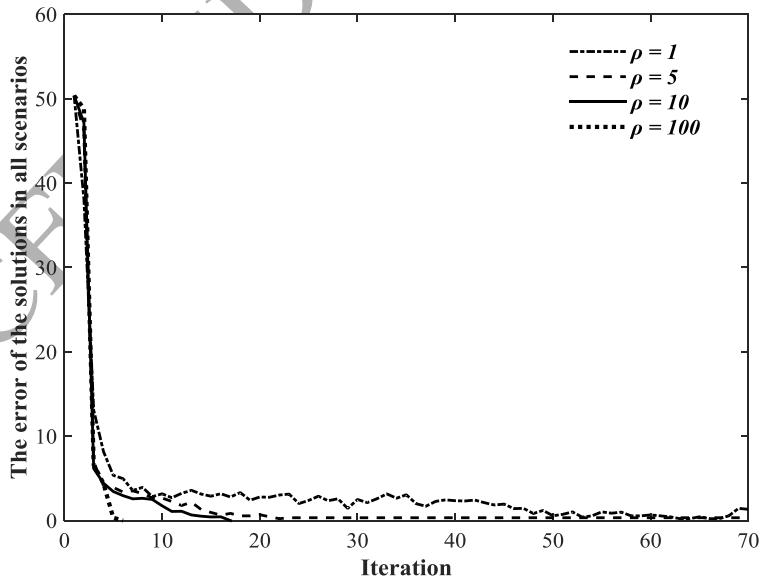
The program is implemented in C# and CPLEX 12.6 is used as a callable library to solve the mixed-integer problem. This algorithm is executed on a windows server with two Intel® Xeon® E5-2620 v2 2.1GHz processors and 120GB memory.

## 6.2. Effects of $\rho$ on Convergence and Solution Quality

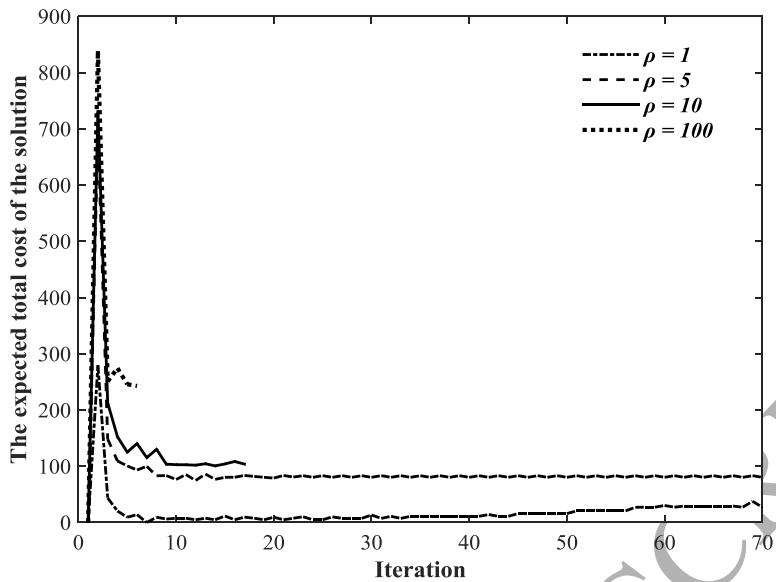
The progressive hedging algorithm's convergence and solution quality are sensitive to the choice of  $\rho$  (Fan & Liu, 2010; Ryan et al., 2013; Watson & Woodruff, 2011). Poor choices can lead to non-convergence or many iterations' to convergence (Ryan et al., 2013). Watson and Woodruff (2011) proposed an effective formula to determine the appropriate value of  $\rho$ . The formula  $\rho(i) = \frac{c(i)}{\max((\sum_{s \in S} p_s |\chi_{s,i}^0 - \bar{\chi}_{s,i}^0|), 1)}$  sets  $\rho$  as a value which is proportional to the variable cost of the associated decision variable  $\chi_{s,i}$ . To check  $\rho$ 's effects on the algorithm's convergence and solution quality, a small size case is studied in this section, where  $\rho$  is set as different values around the value defined by the formula. In this case study, the stochastic model consists of three stages and nine demand scenarios. Each of the second and third stages has three scenarios and the stochastic demands are independent between stages.

The algorithm's convergence and solution quality should be quantified to check  $\rho$ 's effects. The convergence can be described by the total iteration number and the errors of the solutions in all scenarios (Fan & Liu, 2010). The solution quality can be represented by the associated expected total cost of all scenarios in this problem. In each iteration of the algorithm, the average solution of all scenarios is a feasible solution of the original multi-stage stochastic problem and its associated total cost works as an upper bound of the problem.

Figure 4 and Figure 5 illustrate the effects of  $\rho$  on the algorithm's convergence and solution quality respectively. When  $\rho$  is equal to 10 or 100, the error of the solutions in all scenarios decreases to zero in 17 and 6 iterations respectively. The convergence is achieved in the two cases. On the other hand, when  $\rho$  is reduced to 5 or 1, although the error decreases to a small number in only a few iterations, it does not decrease to zero even after quite a number of iterations.



**Figure 4.** The errors of the solutions in all scenarios obtained in PHA



**Figure 5.** The expected total cost of the solution obtained in PHA

As shown in the figures, while fewer iterations are required with the increase of  $\rho$ 's value, the solution quality is getting worse (i.e., a greater upper bound is achieved in case of a greater  $\rho$ 's value). As a result, with the consideration of the trade-off between the convergence and solution quality, an intermediate value of  $\rho$  should be preferred in the algorithm. This conclusion is consistent to that obtained in the literature (Fan & Liu, 2010; Mulvey & Vladimirou, 1991).  $\rho$  is set as 10 in the PHA in the following sections.

### 6.3. The Solving Efficiency Comparison

Although it is claimed that the computational complexity increases significantly with the increase of the stage and scenario numbers in general stochastic programming models, it is not clear if it is true for our problem. In this section, we investigate it by solving the problem as a whole by CPLEX for different stage and scenario numbers. This solution process is referred to as the “direct solution approach” in this section. The problem is also solved by the PHA for comparison. The initial inventory is set as 850 in these cases.

Table 3 compares the solving efficiency between the direct solution approach and the PHA. In the direct solution approach, the solving time and memory required increase dramatically with the increase of the stage and scenario numbers. The optimality gap increases as well. While the problem with only 2 stages and 2 scenarios is easily solved in 50 seconds with negligible memory, the problem with more stages and/or scenarios can only be solved using the direct solution approach with a considerable optimality gap in ten hours. The direct solution approach also require more than 107GB memory in this case. In other words, the direct solution approach is difficult to solve a problem with a large number of stages and scenarios within a reasonable time and the available memory. On the other hand, the PHA gives a better solution in a less solving time within the limited available memory, especially in large cases. Since the solution obtained in PHA is a feasible solution, the corresponding objective value is an upper bound of the problem. As shown in the table, although the PHA gives a larger upper bound in small cases, the objective value obtained in the PHA is much better than it in the direct solution approach in large

cases. As a result, in large cases, the feasible solution obtained in the PHA has a better optimality gap than that obtained in the direct solution approach. In addition, the PHA is not constrained by the available computational resources. The PHA decomposes the whole problem into multiple scenario problems. Each scenario problem can be solved independently, the memory required in the PHA is largely reduced. By applying the parallel computing in the algorithm, the computational time in larger cases can be reduced significantly. Even though a number of iterations and then a longer solving time may be required for larger problems to converge, the PHA can work well with the available server (with the limited memory) to provide a better solution than the direct solution approach. Hence, the progressive hedging algorithm is applied as the solution approach in this paper.

**Table 3.** The solving efficiency comparison

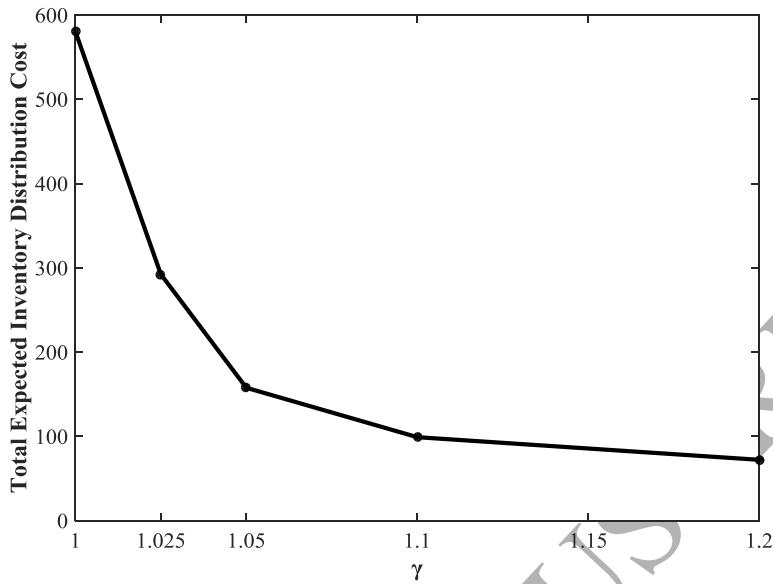
Stage number	Scenario number	Direct solution approach				Memory	PHA	
		Lower bound	Upper bound	Optimality gap	Solving time		Objective value	Solving time
2	2	198.99	211.00	5.69%	50 s	Negligible	504	1h47m
2	3	167.84	184.33	8.95%	~ 47 h	> 107 G	328.67	6h12m
3	4	201.34	277.25	27.38%	~ 17 h	> 90 G	352.75	7h25m
3	9	196.87	383.67	48.69%	~ 9 h	> 64 G	313.22	~ 9 h
4	8	124.3	302	58.84%	~ 21 h	> 107 G	296.38	~ 21 h
4	27	285.08	854.74	66.65%	~ 27 h	> 107 G	379.67	~ 27 h

#### 6.4. Production and Distribution Cost

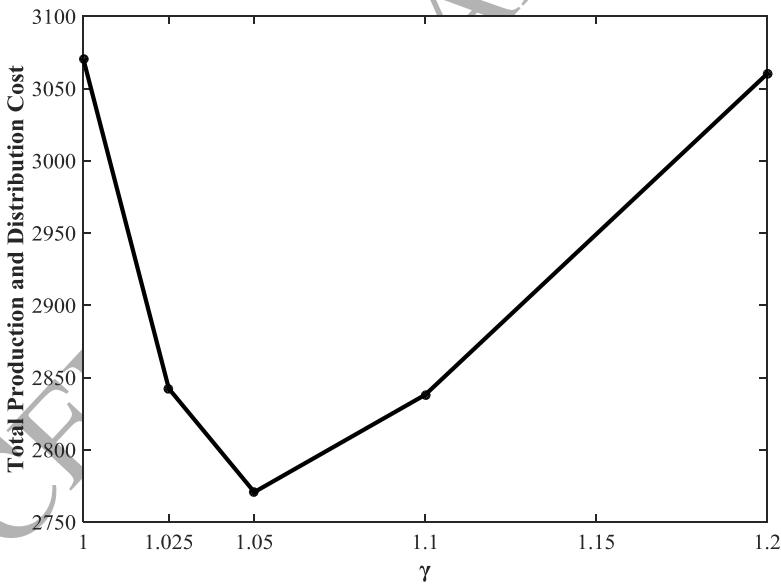
The optimal production quantity can be obtained by balancing the production cost against the expected total inventory allocation cost in clinical trial supply chains. The inventory allocation decisions and associated cost may vary greatly when the production quantity has different values. In this experiment, the production quantity is set to a number of values no less than  $Q_m$ , where  $Q_m$  is the minimal production quantity value required to complete the clinical trial in all scenarios.  $\gamma$  is the ratio of the production quantity value to  $Q_m$ . The unit production cost is set to 3 in this case study. Since the inventory allocation decisions also depend on the clinical trial duration, the experiment is conducted in two cases where the delay is allowed or not in clinical trials. The stochastic model consists of four stages and 27 demand scenarios, where each of the stages later than the first has three scenarios and the stochastic demands are independent between stages.

Figure 6 and Figure 7 demonstrate the inventory distribution cost and the total production and distribution cost obtained in supply chains with different production quantity values, where the minimal clinical trial duration is achieved in each case. The minimal production quantity  $Q_m$  is 830 in this case study. As shown in Figure 6, the expected total inventory distribution cost decreases monotonically with the increase of the production quantity. By adding it with the linear production cost, we can get the total production and inventory distribution cost. Figure 7 shows the total minimal cost under different production quantity values and the production quantity with the minimal total cost can be identified in

this figure. The optimal  $\gamma$  in this figure is equal to 1.05, and the optimal production quantity is 871 in this case study.



**Figure 6.** The expected total inventory distribution cost



**Figure 7.** The production and inventory distribution cost

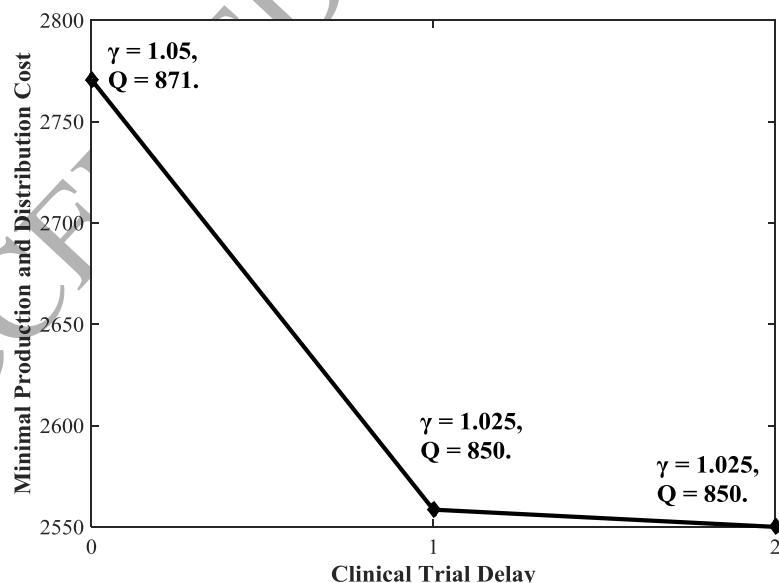
When the clinical trial duration is not too crucial, prolonging it may save the production and distribution costs. The inventory distribution cost is evaluated under different values of the allowed clinical trial delay and production quantity. As shown in Table 4, the total inventory distribution cost decreases with the increase of the production quantity for each clinical trial delay value. It also decreases with the increase of the allowed clinical trial delay for each production quantity. As a result, due to the decreasing inventory distribution cost and increasing production cost, the total cost is convex to the

production quantity for each delay value. It also decreases monotonously with the increase of the clinical trial delay for each production quantity value.

**Table 4.** The production and inventory distribution cost

$\gamma$	$Q$	Production cost	Inventory distribution cost			Production and distribution cost		
			Delay = 0	Delay = 1	Delay = 2	Delay = 0	Delay = 1	Delay = 2
1	830	2490	580.5	180.5	65.75	3070.5	2670.5	2555.75
1.025	850	2550	292.375	8.5	0	2842.375	2558.5	2550
1.05	871	2613	157.75	6.5	0	2770.75	2619.5	2613
1.1	913	2739	99.125	0	0	2838.125	2739	2739
1.2	996	2988	72.25	0	0	3060.25	2988	2988

Figure 8 shows the trade-off between the minimal total production and distribution costs and the allowed delay in clinical trials. As shown in the figure, the total cost decreases with the increase of the allowed clinical trial delay. When the allowed clinical trial delay is increased from 0 to 1, the total cost is reduced by about 7.67% and the production quantity with the minimal total cost also drops from 871 to 850. On the contrary, when the delay is increased from 1 to 2, the total cost is reduced by only 0.33% and the optimal production quantity keeps unchanged. This time-cost trade-off in this figure can provide some guidance for decision makers to determine the production quantity. For example, in cases where the clinical trial duration is less critical than the cost (e.g. the budget is very tight in a pharmaceutical company), the production quantity and the total cost can be reduced with the sacrifice of a short delay in clinical trials.



**Figure 8.** The trade-off between the total cost and clinical trial delay

## 7. Conclusion

This paper considers a production planning problem in clinical trial supply chains. New drug development becomes increasingly complicated that the production cost of drugs can be significant nowadays. Since leftover clinical drugs are useless and should be disposed of after an unsuccessful clinical trial, overproduction should be minimized in clinical trials. When the production quantity is reduced, it becomes more challenging to distribute the inventory among clinical sites. Due to the patient recruitment uncertainty, inappropriate distribution may increase the operational cost or even the clinical trial duration. Since the clinical trial delay can be costly, the production quantity should be optimized so that the production and inventory allocation costs are minimized without increasing the clinical trial duration.

The production planning problem in clinical trial supply chains has been studied in the literature. Although some studies consider the balance between the production cost and the inventory allocation cost, they commonly assume a stationary inventory policy at each location in the supply chain. Since the patient arrivals in clinical trial supply chains follow non-stationary distributions (Y. Chen et al., 2012), the solution can be far from optimal ones if stationary inventory policies are applied in the supply chains. Some researchers also studied the production planning problem considering the failure rates of clinical trials. However, the clinical trial duration is considered in none of these researches. In addition, the widely discussed production and distribution problem in other applications have different problem settings from ours. While those studies neglect overage and mainly consider stock-out minimization, our problem minimizes the overproduction and only considers the stock-out minimization in terms of the clinical trial duration minimization.

In this paper, we find the optimal production quantity in clinical trial supply chains through a multi-objective multi-stage stochastic programming model, where the clinical trial duration and total production and distribution costs are minimized sequentially. Due to the significantly increasing computational complexity in multi-stage stochastic models, the progressive hedging algorithm is applied as the solution approach. As discussed in the numerical experiments, the algorithm's convergence and solution quality are significantly affected by the parameter  $\rho$ . In addition, the inventory distribution cost decreases monotonously with the increase of the production quantity. By balancing it with the increasing production cost, we can find the production quantity with the minimal total cost in clinical trial supply chains. The trade-offs between the total cost and clinical trial delay are also provided as a guidance for decision makers to determine the production quantity, especially when the clinical trial cost is too high to afford for pharmaceutical companies. In this situation, the cost can be reduced significantly with the sacrifice of only a few periods' clinical trial delay.

Since the number of clinical sites in the supply chain affects the production quantity due to the square root law of inventory (Croxton & Zinn, 2005), a direction for future research is to extend the production planning problem by jointly considering the network design in clinical trial supply chains. In addition, this problem's solving efficiency should be further improved. Although the solving efficiency is largely improved by applying the progressive hedging algorithm, solving mixed-integer quadratic scenario sub-problems may still require considerable computational resources (e.g. CPU and memory), especially for those multi-stage stochastic models with a large number of scenarios. More efficient acceleration techniques should be developed to further improve the solving efficiency.

The author contributions (based on CRediT) are summarized as follows:

1. Hui Zhao
  - Data curation
  - Formal analysis
  - Writing - original draft
  - Writing - review & editing
2. Edward Huang
  - Supervision
  - Methodology
  - Validation
  - Writing - review & editing
3. Runliang Dou
  - Methodology
  - Resources
4. Kan Wu:
  - Conceptualization
  - Investigation
  - Supervision
  - Project administration
  - Funding acquisition

## Acknowledgements

Declarations of interest: none. This research is supported in part by the GSK-Singapore Partnership for Green and Sustainable Manufacturing under Grant M406884.

## Appendix

1. The explanation for “double-blind requirement”.

To test a new drug's safety and efficacy, besides the new drug, a placebo (the product without the active pharmaceutical ingredients (API)) and a comparator (a commercial drug with the same targeting disease) are also produced and used in the clinical trial. When a treatment is given, neither the patient nor the doctor does not know what exactly it is. This kind of double-blind study is designed to guarantee the impartiality and integrity of the test. In order to assist the double-blind study and avoid psychological biases, the placebo and the comparator have the same form, packaging and labeling as the new drug.

2. The explanation for the length requirement of the planning horizon.

In the production and distribution problem, the planning horizon should be long enough to guarantee the patient accrual goal can be achieved in the planning horizon with the probability 100% under the settings of the stochastic patient recruitment process. We have the equation  $P(X \geq G) = 100\%$ , where  $G$

represents the patient accrual goal,  $X$  is a random number representing the number of patients who have finished the entire treatment until the end of the planning horizon. Due to the dropout rate  $\varphi$  in each period during the treatment period (consisting of  $m$  periods),  $X$  follows a Poisson distribution and has a mean  $\sum_{t=1}^{l-m+1} \sum_i \lambda_{t,i} (1 - \varphi)^{m-1}$ , where  $l$  is the number of periods in the planning horizon and  $\lambda_{t,i}$  represents the patient arrival rate at clinical site  $i$  in period  $t$ .

3. The explanation for defining the sufficiently large number  $M = G \sum_{k=1}^m (1 - \varphi)^{k-1} / (1 - \varphi)^{m-1}$ . The defined value for the sufficiently large number  $M$  is actually the minimal production quantity of clinical drugs. For each clinical trial, the production quantity should be large enough to guarantee the clinical trial's patient accrual goal is reached. Due to the dropout rate  $\varphi$  in each period during the treatment period (consisting of  $m$  periods), only  $(1 - \varphi)^{m-1}$  of the recruited patients would finish the entire treatment and deliver valid trial data. Hence, the number of recruited patients should be no less than  $G / (1 - \varphi)^{m-1}$ . In addition,  $(1 - \varphi)^{k-1}$  of them would attend the clinical trial for at least  $k$  periods,  $k \in \{1, \dots, m\}$ . As a result, to reach the patient accrual goal  $G$ , the production quantity should be at least  $\sum_{k=1}^m (1 - \varphi)^{k-1}$  times of the number of patients required, i.e.,  $Q \geq G \sum_{k=1}^m (1 - \varphi)^{k-1} / (1 - \varphi)^{m-1}$ .

## References

- Abdelkafi, C., Beck, B. H., David, B., Druck, C., & Horoho, M. (2009). Balancing risk and costs to optimize the clinical supply chain—a step beyond simulation. *Journal of Pharmaceutical Innovation*, 4(3), 96-106.
- Beamon, B. M. (1998). Supply chain design and analysis:: Models and methods. *International journal of production economics*, 55(3), 281-294.
- Birge, J. R. (1985). Decomposition and partitioning methods for multistage stochastic linear programs. *Operations Research*, 33(5), 989-1007.
- Birge, J. R., & Louveaux, F. (2011). *Introduction to stochastic programming*: Springer Science & Business Media.
- Boudia, M., Dauzere-Peres, S., Prins, C., & Louly, M. A. O. (2006). *Integrated optimization of production and distribution for several products*. Paper presented at the 2006 International Conference on Service Systems and Service Management.
- Boudia, M., Louly, M. A. O., & Prins, C. (2008). Fast heuristics for a combined production planning and vehicle routing problem. *Production Planning and Control*, 19(2), 85-96.
- Carøe, C. C., & Tind, J. (1998). L-shaped decomposition of two-stage stochastic programs with integer recourse. *Mathematical Programming*, 83(1-3), 451-464.
- Chandra, P., & Fisher, M. L. (1994). Coordination of production and distribution planning. *European Journal of Operational Research*, 72(3), 503-517.

- Chen, Y., Mockus, L., Orcun, S., & Reklaitis, G. V. (2012). Simulation-optimization approach to clinical trial supply chain management with demand scenario forecast. *Computers & Chemical Engineering*, 40, 82-96. doi:DOI 10.1016/j.compchemeng.2012.01.007

Chen, Y., Pekny, J. F., & Reklaitis, G. V. (2012). Integrated planning and optimization of clinical trial supply chain system with risk pooling. *Industrial & engineering chemistry research*, 52(1), 152-165.

Clark, A. J., & Scarf, H. (1960). Optimal policies for a multi-echelon inventory problem. *Management Science*, 6(4), 475-490.

Clemento, A. (1999). New and integrated approaches to successful accelerated drug development. *Drug information journal*, 33(3), 699-710.

Coello, C. A. C., Lamont, G. B., & Van Veldhuizen, D. A. (2007). *Evolutionary algorithms for solving multi-objective problems* (Vol. 5): Springer.

Croxton, K. L., & Zinn, W. (2005). Inventory considerations in network design. *Journal of Business Logistics*, 26(1), 149-168.

Csirik, J., Frenk, J. B. G., Labb  , M., & Zhang, S. (1990). *Heuristics for the 0-1 min-knapsack problem*: European Institute for Advanced Studies in Management Erasmus Universiteit Rotterdam.

DiMasi, J. A., Hansen, R. W., & Grabowski, H. G. (2003). The price of innovation: new estimates of drug development costs. *Journal of health economics*, 22(2), 151-185.

Dong, L., & Rudi, N. (2004). Who benefits from transshipment? Exogenous vs. endogenous wholesale prices. *Management Science*, 50(5), 645-657.

Ereng  c,   . S., Simpson, N. C., & Vakharia, A. J. (1999). Integrated production/distribution planning in supply chains: An invited review. *European Journal of Operational Research*, 115(2), 219-236.

Fan, Y., & Liu, C. (2010). Solving stochastic transportation network protection problems using the progressive hedging-based method. *Networks and Spatial Economics*, 10(2), 193-208.

Fleischhacker, A. (2009). *An investigation of clinical trial supply chains*. Rutgers University-Graduate School-Newark.

Fleischhacker, A., Ninh, A., & Zhao, Y. (2015). Positioning Inventory in Clinical Trial Supply Chains. *Production and Operations Management*, 24(6), 991-1011.

Fleischhacker, A. J., & Zhao, Y. (2011). Planning for demand failure: A dynamic lot size model for clinical trial supply chains. *European Journal of Operational Research*, 211(3), 496-506.

Fleischhacker, A. J., & Zhao, Y. (2013). Balancing learning and economies of scale for adaptive clinical trials. *Operations Research for Health Care*, 2(3), 42-51.

Gade, D., Hackebeil, G., Ryan, S. M., Watson, J.-P., Wets, R. J.-B., & Woodruff, D. L. (2016). Obtaining lower bounds from the progressive hedging algorithm for stochastic mixed-integer programs. *Mathematical Programming*, 157(1), 47-67.

Gatica, G., Papageorgiou, L., & Shah, N. (2003). Capacity planning under uncertainty for the pharmaceutical industry. *Chemical Engineering Research and Design*, 81(6), 665-678.

Grabowski, H. G., & Vernon, J. M. (2000). Effective patent life in pharmaceuticals. *International Journal of Technology Management*, 19(1-2), 98-120.

Gupta, A., & Maranas, C. D. (2003). Managing demand uncertainty in supply chain planning. *Computers & Chemical Engineering*, 27(8), 1219-1227.

Herer, Y. T., & Rashit, A. (1999). Lateral stock transshipments in a two - location inventory system with fixed and joint replenishment costs. *Naval Research Logistics (NRL)*, 46(5), 525-547.

Herer, Y. T., & Tzur, M. (2001). The dynamic transshipment problem. *Naval Research Logistics (NRL)*, 48(5), 386-408.

Herer, Y. T., & Tzur, M. (2003). Optimal and heuristic algorithms for the multi-location dynamic transshipment problem with fixed transshipment costs. *IIE Transactions*, 35(5), 419-432.

Huang, K. (2005). *Multi-stage stochastic programming models in production planning*. Georgia Institute of Technology.

Jang, Y.-J., Jang, S.-Y., Chang, B.-M., & Park, J. (2002). A combined model of network design and production/distribution planning for a supply network. *Computers & Industrial Engineering*, 43(1), 263-281.

- Kaitin, K. I., & DiMasi, J. A. (2011). Pharmaceutical innovation in the 21st century: new drug approvals in the first decade, 2000–2009. *Clinical pharmacology & therapeutics*, 89(2), 183-188.
- Laínez, J. M., Schaefer, E., & Reklaitis, G. V. (2012). Challenges and opportunities in enterprise-wide optimization in the pharmaceutical industry. *Computers & Chemical Engineering*, 47, 19-28.
- Lee, Y. H., & Kim, S. H. (2002). Production–distribution planning in supply chain considering capacity constraints. *Computers & Industrial Engineering*, 43(1), 169-190.
- Løkketangen, A., & Woodruff, D. L. (1996). Progressive hedging and tabu search applied to mixed integer (0, 1) multistage stochastic programming. *Journal of Heuristics*, 2(2), 111-128.
- Marler, R. T., & Arora, J. S. (2004). Survey of multi-objective optimization methods for engineering. *Structural and multidisciplinary optimization*, 26(6), 369-395.
- Martinez, B., & Goldstein, J. (2007). Big pharma faces grim prognosis. *Wall Street Journal A*, 1, 12.
- Mula, J., Peidro, D., Díaz-Madroñero, M., & Vicens, E. (2010). Mathematical programming models for supply chain production and transport planning. *European Journal of Operational Research*, 204(3), 377-390.
- Mullin, R. (2011). Before the Storm: AMER CHEMICAL SOC 1155 16TH ST, NW, WASHINGTON, DC 20036 USA.
- Mulvey, J. M., & Vladimirou, H. (1991). Applying the progressive hedging algorithm to stochastic generalized networks. *Annals of Operations Research*, 31(1), 399-424.
- Paterson, C., Kiesmüller, G., Teunter, R., & Glazebrook, K. (2011). Inventory models with lateral transshipments: A review. *European Journal of Operational Research*, 210(2), 125-136.
- Powell, M. (2010). Presentation by Senior Vice President, *Pharmaceutical Development*, Bristol-Myers Squibb. Argyle Executive Forum's 2010 Leadership in pharmaceuticals & biotechnology conference. June 22, 2010. New York, NY. Paper presented at the Argyle Executive Forum.
- Profile, P. I. (2010). Pharmaceutical Research and Manufacturers of America. *Washington DC, US, PhRMA March*.
- Roberts, T. G., Lynch, T. J., & Chabner, B. A. (2003). The phase III trial in the era of targeted therapy: unraveling the “go or no go” decision. *Journal of Clinical Oncology*, 21(19), 3683-3695.

- Robinson, L. W. (1990). Optimal and approximate policies in multiperiod, multilocation inventory models with transshipments. *Operations Research*, 38(2), 278-295.
- Rockafellar, R. T., & Wets, R. J.-B. (1991). Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of operations research*, 16(1), 119-147.
- Rudi, N., Kapur, S., & Pyke, D. F. (2001). A two-location inventory model with transshipment and local decision making. *Management Science*, 47(12), 1668-1680.
- Ryan, S. M., Wets, R. J.-B., Woodruff, D. L., Silva-Monroy, C., & Watson, J.-P. (2013). *Toward scalable, parallel progressive hedging for stochastic unit commitment*. Paper presented at the Power and Energy Society General Meeting (PES), 2013 IEEE.
- Scarf, H. (1959). The optimality of (5, 5) policies in the dynamic inventory problem.
- Sen, S. (2005). Algorithms for stochastic mixed-integer programming models. *Handbooks in operations research and management science*, 12, 515-558.
- Shah, N. (2004). Pharmaceutical supply chains: key issues and strategies for optimisation. *Computers & Chemical Engineering*, 28(6), 929-941.
- Shapiro, A., Dentcheva, D., & Ruszczyński, A. (2009). *Lectures on stochastic programming: modeling and theory*: SIAM.
- Sherali, H. D., & Fraticelli, B. M. (2002). A modification of Benders' decomposition algorithm for discrete subproblems: An approach for stochastic programs with integer recourse. *Journal of Global Optimization*, 22(1-4), 319-342.
- Sundaramoorthy, A., Evans, J. M. B., & Barton, P. I. (2012). Capacity Planning under Clinical Trials Uncertainty in Continuous Pharmaceutical Manufacturing, 1: Mathematical Framework. *Industrial & Engineering Chemistry Research*, 51(42), 13692-13702. doi:Doi 10.1021/ie300324h
- Tunc, H., Kilic, O. A., Tarim, S. A., & Eksioglu, B. (2011). The cost of using stationary inventory policies when demand is non-stationary. *Omega*, 39(4), 410-415.
- Vidal, C. J., & Goetschalckx, M. (1997). Strategic production-distribution models: A critical review with emphasis on global supply chain models. *European Journal of Operational Research*, 98(1), 1-18.

- Watson, J.-P., & Woodruff, D. L. (2011). Progressive hedging innovations for a class of stochastic mixed-integer resource allocation problems. *Computational Management Science*, 8(4), 355-370.
- Woodruff, D. L., & Zemel, E. (1993). Hashing vectors for tabu search. *Annals of Operations Research*, 41(2), 123-137.
- Yang, J., & Qin, Z. (2007). Capacitated production control with virtual lateral transshipments. *Operations Research*, 55(6), 1104-1119.
- Zhang, J. (2005). Transshipment and its impact on supply chain members' performance. *Management Science*, 51(10), 1534-1539.