10^{th} *International Conference on MOdeling, Optimization and SIMlation - MOSIM16 - August 22-24-2016*
Montreal - Canada, "Innovation in Technology for Performant Systems: Challenges and Opportunities"

A multi-period production and distribution optimization model for radiopharmaceuticals

Ioannis AKROTIRIANAKIS, Amit CHAKRABORTY

SIEMENS CORPORATE TECHNOLOGY
755 College Road East
Princeton, NJ 08540, USA
{ioannis.akrotirianakis , amit.chakraborty}@siemens.com

ABSTRACT: *This paper addresses the manufacturing and distribution of short-lived radiopharmaceuticals which are mainly used in diagnostic imaging studies. We develop a mixed integer nonlinear optimization model that is flexible enough to capture the complex underlying physics of the production process of fludeoxyglucose (FDG), which is widely used in oncology and cardiology, as well as the time sensitive constraints of the distribution of the final products to geographically dispersed medical imaging centers. The model synergistically integrates the production and delivery requirements in a multi-period framework. It generates the optimal amount of radioactivity needed to satisfy the demand placed by imaging centers during a full day and provides the minimum cost transportation routes that guarantee the on-time delivery of the doses. We present numerical results that demonstrate the usefulness of the model by substantial cost savings in both the manufacturing and transportation phases.*

KEYWORDS: *Production Planning, Vehicle Routing Problem with Time Windows, Perishable Products, Integer & Nonlinear Optimization.*

1 INTRODUCTION

Supply chain management is a very important aspect of a well-functioning company, involved in the manufacturing and distribution of products. It can be defined as the group of methods used to efficiently connect suppliers, manufacturers, warehouses and customers with the ultimate aim to produce enough quantities, distribute them to customer locations at specified times, and minimize all associated costs while ensuring complex service-oriented constraints are always satisfied (Simchi-Levi et al., 1999). The development of mathematical models and computational techniques for solving complex real-world supply chain problems is currently one of the most active application areas in Operations Research and Management Science (Shen, 2007). However, one of the supply chains that has not been explored yet is the production and distribution of radiopharmaceuticals for Positron Emission Tomography (PET). We believe this is due to the multidisciplinary nature of PET that involves nuclear physics, radiochemistry and advanced instrumentation, resulting in complex processes of manufacturing and distribution. To the best of our knowledge the paper by (Nagurney and Nagurney, 2012) is the only one that studies the supply chain of the radiopharmaceutical industry. The authors have developed a multi-tiered supply chain network design model for a radio-isotope called Technetium. The design of the supply chain is modeled as an optimization problem on a generalized network. The specific losses on an arc of the network are identified using the time decay of the radioisotope. They minimize the cost associated with the operational activities in the network, together with the waste management of radio active products involved. They use a Variational Inequality approach to solve the problem as opposed to an Integer Programming one we use in our paper.

PET is a nuclear medicine procedure that is routinely used by physicians to generate high quality images of functional process in the human body. The main aim of PET is to identify new conditions or to monitor the progress of treatments in various body organs or tissues. PET is mostly used in oncology, cardiology and more recently in neurology. In order to create an image a very small amount of a radioactive tracer is injected into the patient's body. These tracers are made by attaching a radionuclide to a chemical substance that is used during the metabolic process by the organ that needs to be examined. One of the most commonly used radiopharmaceutical in PET is called *Fludeoxyglucose* (FDG). Its sales have been growing since 2010 and are expected to exceed $880 million by 2017, while the market for PET radio-

pharmaceuticals will increase to $3.5 billion by 2017 (Burns, 2010). FDG is used in approximately 95% of all PET studies for patients with cancer.

A dose of FDG is injected rapidly into a saline drip running into a patient's vain. It is made of a radionuclide and a carrier molecule. The role of the former is played by Fluorine-18, which is a radioisotope with extra neutrons in its nucleus. Fluorine-18 is unstable and it will eventually split, releasing positrons in the process[1]. When these positrons come into contact with electrons in neighboring atoms within the patient's body they generate gamma rays which are detected by PET cameras, which are arranged in a ring through which the patient is moved. On the other hand, the role of the carrier molecule is played by glucose,[2] used by the human body to generate energy. Due to their abnormal generation, cancerous cells need more glucose levels than healthy cells and therefore will have a higher concentration of FDG. This will result in brighter areas in a PET image, and the tumors can be identified in early stages.

Radionuclides are produced in special equipments, called cyclotrons, from bombardment of a target material with charged particles (Jacobson, 2012). A cyclotron is a particle accelerator. Negatively charged particles are placed at its center and a strong magnetic field increases theier kinetic energy. The particles are then collide with a target material yielding positron emitting radionuclides. For the production of Fluorine-18 the accelerated particles are usually isotopes of Hydrogen and the target material is Oxygen-18 enriched water. The longer the bombardment process lasts the more Fluoride-18 isotopes are produced which in turn will be used to synthesize larger quantities of FDG. Due to their unstable nature, Fluorine-18 radionuclides will undergo radioactive decay soon after their production. During the decay process their radioactivity levels decrease. The length of time for the radioactivity to reach half of its initial amount is called *half life* and for Fluorine-18 it is 109.771 minutes. This makes Fluorine-18 very useful tracer, since after almost 10 hours it will be completely out of the patients body. However, the decay process poses a difficult and complex *logistical* problem when it needs to be distributed to imaging centers or hospitals that are not close to the production facility (radiopharmacy). The production of Fluorine-18 as well as the distribution of the FDG need to be *coordinated* in such a way that the radioactivity at the end of the bombardment is high enough allowing it to decay to the levels that are acceptable for injection before the patient enters the PET scanner.

In this paper we focus on a single radiopharmacy, which serves a number of imaging centers. Each radiopharmacy is assumed to have up to two cyclotrons and as such its capacity is limited. The planning horizon is one day where production may take place during several times periods depending on the demand. Typically, production takes place two to four times a day. The imaging centers place their orders the previous day and production starts soon after midnight of the next day. Due to the radioactive decay of Fluorine-18 and its short half-life, it is not possible to store it and use it in several future periods. Each dose ordered by an imaging center has a specific radioactivity level and injection time. The doses are assigned to a time period based on their injection times. This defines the demand for a specific time period. FDG is transported as a single patient dose in syringes or in a multi-dose vial that can serve many patients. The syringes or the vials are placed in heavy metallic containers that prevent radiation from spreading to personnel that handles their deliveries. A dose must arrive to its destination a certain amount of time before its injection time, otherwise the radiopharmacy has to replace it (free of charge) during a later time or day. These requirements make both the production and distribution processes complex and *time-sensitive*. The optimization model developed in this paper belongs to the category of Mixed Integer Nonlinear Programming problems (Floudas, 1995). The optimal solution *integrates* four major decisions: (a) how long the bombardment time should be in order to generate enough radioactivity which will satisfy the total demand in each time period, (b) how many cyclotrons and targets should be used to generate that radioactivity, (c) which routes the delivery vehicles should follow in order to minimize total transportation cost, and (d) the arrival time at every imaging center, which must not be later than the injection time of all the doses ordered by it. The set of periods is denoted by $\mathcal{T}$.

2 RADIOACTIVE DECAY

We briefly present the underlying physics of radioactive decay, because it plays a fundamental role in the construction of the optimization model. Radioactive decay is a stochastic process and as such we can only determine the probability when a radioisotope decays (Sprawls, 1995). However, for a group of radioisotopes we can determine its decay rate as a whole. The *half-life* of a group of radioisotopes, denoted by $t_{1/2}$, is the time after which, on average, half of them will have decayed and the activity of the group will be half of its initial value. The *decay constant*, denoted by λ, expresses the probability that a nucleus will undergo a transition in a specific time period. The relation between half-life and decay constant is given by $t_{1/2} = \frac{ln(2)}{\lambda}$. The decay constant of Fluorine-18 is $\lambda_F = 0.006311$. For a given sample that contains Q radioisotopes, the number of decay events, $-dQ$,

[1] a positron is a positively charged electron
[2] glucose is a type of sugar

which is expected to happen during a short interval $[t, t+dt]$ is proportional to Q and dt, that is,

$$-dQ \propto Q\, dt \tag{1}$$

where the minus sign indicates that the number of radioactive atoms decreases. We can turn the proportionality in (1) into an equality by using the decay constant λ as follows

$$-dQ = \lambda\, Q\, dt \;\Rightarrow\; \frac{dQ}{Q} = \lambda\, dt \tag{2}$$

By integrating (2) in a time interval $[T_0, T_1]$ we obtain

$$\int_{Q^{T_0}}^{Q^{T_1}} \frac{dQ}{Q} = \lambda \int_{T_0}^{T_1} dt \Rightarrow Q^{T_1} = Q^{T_0} e^{-\lambda(T_1 - T_0)} \tag{3}$$

where Q^{T_0} and Q^{T_1} denote the number of radioactive nuclei at time points T_0 and T_1, respectively. Equation (3) is known as the *radioactive decay law*.

3 PRODUCTION PROCESS

The production of FDG requires the use of complex equipment the most important of which are: (i) one or more cyclotrons for producing Fluorine-18, (ii) one or more FDG synthesizers, (iii) an analytical laboratory where all the quality assurance tests are performed (e.g., testing batches for sterility, chemical purity, etc), and (iv) a dispenser which is responsible for separating the total quantity of FDG into the unit doses which in turn will be delivered to imaging centers. An overview of a typical work flow for the production of FDG can be seen in Figure 1. The process consists of many stages. The completion times of each stage (except the bombardment, denoted by the dashed boxes) are known and their length is shown in minutes. The optimal bombardment time is determined by our model. In the remaining of this paper we will denote by $\mathcal{C} = \{1, \ldots, N^C\}$ the set of available cyclotrons and by $\mathcal{R} = \{1, \ldots, N^R\}$ the set of available targets.

3.1 Cyclotrons and targetry

The production rate of Fluorine-18 depends on the beam current, the enrichment of the target material and the saturation yield. The beam current is the flux of the bombarding particles. The end of bombardment (EOB) occurs when the beam current is set to zero. In order to produce larger quantities of Fluorine-18 it is important to use water enriched in Oxygen-18 as target material. The rate of production of Fluorine-18 is also affected by the fact that the resulting nuclide is also radioactive and its decay is expected to start any time after it is produced. The rates of production and decay will eventually reach an equilibrium after sufficiently long bombardment times. The point where equilibrium is achieved is called saturation and it means that there is no additional Fluorine-18 produced and as a result the bombardment will have to stop. The following equation defines the quantity of Fluorine-18 produced at the end of the bombardment (EOB) of a single target

$$Q^{EOB} = b(SY)(EN)(1 - e^{-\lambda(T^{EOB} - T^{BOB})}) \tag{4}$$

where b represents the beam current, SY denotes the saturation yield for the target, EN represents the enrichment of the Qxygen-18 water and T^{BOB} and T^{EOB} represent the beginning and end of bombardment of a target, respectively. The bombardment time may range from 30 minutes to 4 hours.

Each cyclotron can accelerate particles in two different lines (some times called *beam-lines*) and may bombard one or two targets. Thus, it is possible to use up to four targets when two cyclotrons are available. The target loading process must be completed before applying beam to the target. When first applying the beam to a target, a tuning process is required before the beam can be operated at full power. The beginning of bombardment (BOB) occurs after the beam is at full power on the target. The end of bombardment (EOB) occurs when the beam current is set to zero. If two targets are used in the bombardment process, the EOB will be the same for both targets. In order to generate enough quantity of Fluorine-18 ion and be able to satisfy all the orders placed by imaging centers we have to determine the optimal number of targets, the beam current, and the bombardment time for each time period.

3.2 Target unload process

The total amount of Fluoride-18 produced by each target is consolidated into a single product during the unload process. To accomplish this, each target in a cyclotron is sequentially unloaded to a collection vial. Targets may not be unloaded simultaneously. The total amount of Fluorine-18 at the end of unload (EOU) of both targets at the first cyclotron is

$$Q_1^{EOU} = Q_1^{EOU_2} + Q_1^{EOU_1} e^{-\lambda(T^{UNLD})} \tag{5}$$

where $Q_1^{EOU_1} + Q_1^{EOU_2}$ is the radioactivity produced by targets 1 and 2, respectively, T^{UNLD} is the time needed to unload a target. A similar equation can be derived for the second cyclotron. After completing the unload process, a target may be immediately reloaded in preparation for the bombardment for the next time period. The time required for the reload process is the same as the time for the target load process during the first time period.

3.3 Fluorine-18 ion transfer process

After the target unload process is complete, the vial containing the produced Fluorine-18 ion is transferred

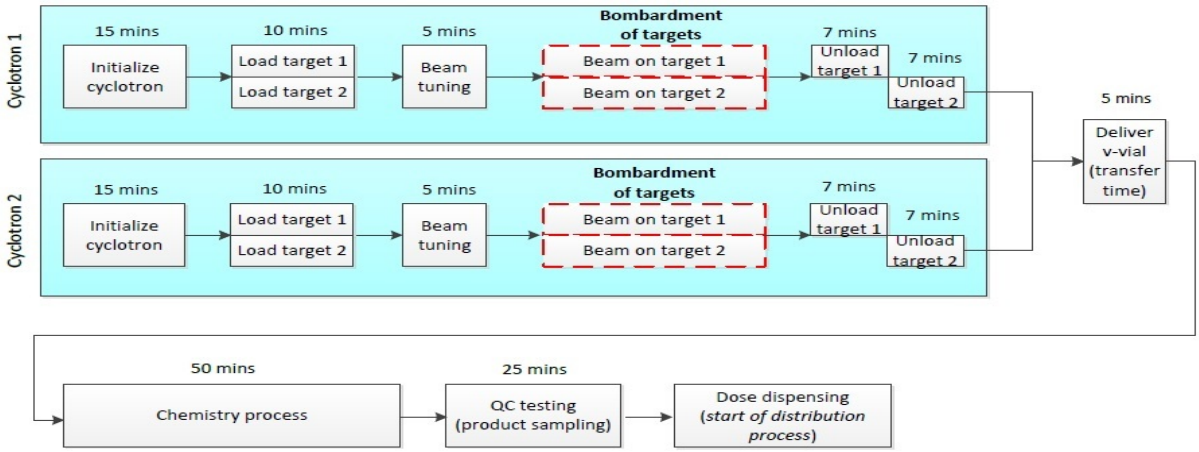

Figure 1: Production workflow with two cyclotrons each employing two targets

to the chemistry module. A small percentage of the total Fluorine-18 ion is lost during the transfer process. The time when the transfer process is complete is defines the beginning of synthesis (BOS) and the quantity of Fluoride-18 ion at that time is:

$$Q^{BOS} = (Q^{EOU})(1 - n(FR))(e^{-\lambda(T^{TRAN})}) \quad (6)$$

where n is the number of targets, (FR) is the fractional percentage lost, and T^{TRAN} is the time interval between the end of target unload process and the Beginning of Synthesis (i.e., $T^{TRAN} = T^{BOS} - T^{EOU}$).

3.4 Chemistry process

The chemistry module converts the available Fluorine-18 ion into FDG. A percentage of the Fluorine-18 ion is lost during the chemistry process. This percentage is known as the *percent yield*. The time when the chemistry process is complete is defined as the end of synthesis (T^{EOS}) and the quantity of FDG at the EOS is given by:

$$Q^{EOS} = (PY)(Q^{BOS})(e^{-\lambda(T^{CHEM})}) \quad (7)$$

where (PY) represents the percent yield of the process, and T^{CHEM} is time interval needed by the chemistry process (i.e., $T^{CHEM} = T^{EOS} - T^{BOS}$).

3.5 Product sampling and quality control

After the chemistry process is complete, the product is sampled for purposes of quality control (QC) testing. This sample is a percentage of the product resulting from the chemistry process. The remaining quantity of the product is available for dose dispensing and is given by:

$$Q^{DSP} = (1 - (QS))Q^{EOS} \quad (8)$$

where (QS) represents the percentage of the product used in the sample.

3.6 Dose dispensing and total radioactivity demand

The total quantity of FDG will be divided into the individual doses ordered by the imaging centers. More specifically, at time period t, we assume that we have N_t^H imaging centers and the i-th center has ordered D_{it} doses of FDG. The j-th dose has an injection time, τ_{ijt}^{INJ}, and a radioactivity level, Q_{ijt}^{INJ}, at the time of its injection. In order to determine how much radioactivity, Q_{ijt}^{DSP}, should be available at the dispensing time we can use the decay law defined in (3). By solving (3) for Q^{T_0} and then setting $T_0 = \tau_t^{DSP}$, $T_1 = \tau_{ijt}^{INJ}$, $Q^{T_0} = Q_{ijt}^{DSP}$ and $Q^{T_1} = Q_{ijt}^{INJ}$, we have

$$Q_{ijt}^{DSP} = Q_{ijt}^{INJ}\, e^{\lambda(\tau_{ijt}^{INJ} - \tau_t^{DSP})}.$$

Summing up all doses ordered by all imaging centers during period t we can determine the total radioactivity that must be available in order to satisfy the customer demand, that is

$$q_t^{DSP} = \sum_{i=1}^{N_t^H} \sum_{j=1}^{D_{it}} Q_{ijt}^{DSP} \quad (9)$$

4 DISTRIBUTION PROCESS

Once the dispensing process has been completed, the doses are ready to be picked up by transportation vehicles and delivered to the imaging centers. A driver may be allowed to work up to a certain number of hours, denoted by T (e.g., 8 hours). Also there is a maximum number of vehicles (denoted by N). Mathematically the transportation of the doses to the medical imaging centers can be expressed as a general vehicle routing problem with time widows. At time period t, let $\mathcal{G}_t = (\mathcal{V}_t, \mathcal{E}_t)$ be a complete undirected graph where $\mathcal{V}_t = \{0, 1, 2, \ldots, N_t^H\}$ represents the nodes of the graph and $E_t = \{(i, j) : i, j \in \mathcal{V}_t\}$ is the set of edges connecting the nodes. The set of

nodes consists of the radio-pharmacy, denoted by 0, and the imaging centers that have placed orders during period t. For convenience we will denote the set of the imaging centers as $\mathcal{V}_t^H = \{1, 2, \ldots, N_t^H\}$. As mentioned in the previous section, D_{it} represent the number of doses ordered by the i-th imaging center during period t and τ_{ijt}^{INJ} is the injection time of the j-th dose ordered by the i-th imaging center during period t. All doses ordered by the a specific imaging center are delivered by the same vehicle. This means that the delivery vehicle must arrive at the imaging center before the earliest injection time, that is

$$T_{it}^{INJ} = \min\{\tau_{ijt}^{INJ} : j = 1, \ldots, D_{it}\}, \tag{10}$$

Furthermore, every imaging center may require the delivery to arrive during a certain time window, $[E_{it}, L_{it}]$, where E_{it} is the earliest and L_{it} the latest arrival times at the i-th imaging center. The latest arrival time may be a certain number of minutes, p_i, prior to T_{it}^{INJ} defined by (10). Therefore the vehicle must arrive at the i-th imaging center no later than

$$L_{it} = T_{it}^{INJ} - p_i, \forall i \in \mathcal{V}_t^H, \forall t \in \mathcal{T}. \tag{11}$$

The distance and drive-time between node i and node j are denoted by Δ_{ij} and H_{ij}, respectively, and are obtained by using the *geo-coding* services offered by Google[3]. The cost, C_{ij}, of driving from a location i to a location j is defined as

$$C_{ij} = (m + f)\Delta_{ij} + g \tag{12}$$

where the m is the cost of traveling one mile, f is the fuel surcharge for every mile traveled, and g is a flat amount charged by the drivers for every customer site they visit[4]. Also there is a fixed cost, F, associated with every vehicle used. All vehicles have the same capacity, denoted by W^{VEH}. The weight of the containers that carry the FDG doses is denoted by W^{CON}. Hence the total weight of the order of the i-th customer, during period t, is defined by

$$W_{it} = D_{it}\, W^{CON}. \tag{13}$$

We assume that one vehicle will deliver all doses ordered by a customer, that is, we do not consider split orders. This means that all orders weigh less than the vehicle's capacity, i.e., $W_{it} \leq W^{VEH}$. Once a vehicle arrives at a customer location the drivers need to spend some time unloading the containers, signing certain documents and picking up empty boxes. This is called *service time* of customer i, denoted by s_i.

5 OPTIMIZATION MODEL

In this section we discuss the variables and constraints that are used to model the production and transportation processes described in the previous section.

[3]https://developers.google.com/maps/

[4]typical value ranges of m, f and g are \$1.1-\$1.5, \$0.055-\$0.065, and \$10-\$15

We consider a *multiperiod process*, where a day is divided into a small number of periods. Production that takes place in period t is transported to the customers in the following period.

5.1 Decision variables

In this section we describe the main decision variables we use in the model. We first present the variables associated with the production process, followed by those in the transportation process.

5.1.1 Production variables

The *beginning* and *end of bombardment* of target j at cyclotron i in time-period t are denoted by τ_{ijt}^{BOB} and τ_{ijt}^{EOB}, respectively. The *beam current* applied on target j in cyclotron i at period t is denoted by b_{ijt}. We use binary variables, z_{ijt}, to express the fact that a target may be used or not used in a particular time period, that is, z_{ijt} equals 1, if we use target j in cyclotron i during period t, and 0 otherwise. The latest time of the target unload process at period t is denoted by τ_t^{EOU} and the radioactivity level at that time is denoted by q_t^{EOU}. The time when the dispensing of the doses starts is denoted by τ_t^{DSP}, and the total radioactivity at that time is denoted by q_t^{DSP}.

5.1.2 Distribution variables

We use the binary variables y_{ijt} to represent the order by which the customers are visited by a transportation vehicle, that is, $y_{ijt} = 1$ if a vehicle goes from customer i to j during period t, and $y_{ijt} = 0$, otherwise. We define the variable w_{it} to representing the total weight a vehicle has delivered until it has reached customer i in period t. We also define x_{it} to represent the arrival time of a vehicle to customer site i in period t. Note that x_{it} must be within the time window of the corresponding customer.

5.2 Model definition

In this section we present the complete production and transportation optimization model, defined by equations (14) through (39). In next section we describe in detail the constraints. The lower case letters denote variables and upper case the parameters.

$$\min \sum_{t\in\mathcal{T}}\sum_{i\in\mathcal{C}}\sum_{j\in\mathcal{R}}((\tau_{ijt}^{EOB} - \tau_{ijt}^{BOB})\kappa + \sum_{i\in\mathcal{V}}\sum_{j\in\mathcal{V}}\sum_{t\in\mathcal{T}} y_{ijt}C_{ij} + F\sum_{j\in\mathcal{V}}\sum_{j\in\mathcal{T}} y_{0jt} \tag{14}$$

s.t.

$$T_{\min}^{BMB}\, z_{ijt} \leq \tau_{ijt}^{EOB} - \tau_{ijt}^{BOB} \leq T_{\max}^{BMB}\, z_{ijt}, \tag{15}$$

$$B_{\min} z_{ijt} \leq b_{ijt} \leq B_{\max} z_{ijt}, \quad (16)$$

$$\tau_{ij1}^{BOB} = T^{INIT}, \quad (17)$$

$$\tau_{i1t}^{EOB} = \tau_{ijt}^{EOB}, \quad (18)$$

$$\tau_{i1t}^{EOB} + \sum_{j \in \mathcal{R}} z_{ijt}\, T^{UNLD} \leq \tau_t^{EOU}, \quad (19)$$

$$\tau_{ijt}^{BOB} \geq \tau_{t-1}^{EOU} + T^{PREP}, \quad (20)$$

$$q_t^{EOU} = \sum_{i \in \mathcal{C}} \sum_{j \in \mathcal{R}} b_{ijt}\, SY\, EN \\ (1 - e^{-\lambda(\tau_{ijt}^{EOB} - \tau_{ijt}^{BOB})}) e^{-\lambda(\tau_t^{EOU} - \tau_{ijt}^{EOB})}, \quad (21)$$

$$q_t^{DSP} = q_t^{EOU}\, (1 - FR)\, PY\, (1 - QS) \\ e^{-\lambda(T^{TRAN} + T^{CHEM} + T^{QS})}, \quad (22)$$

$$\tau_t^{DSP} = \tau_t^{EOU} + T^{TRAN} + T^{CHEM} + T^{QS}, \quad (23)$$

$$q_t^{DSP} \geq \sum_{i=1}^{N_t^H} \sum_{j=1}^{D_{it}} Q_{ijt}^{INJ}\, e^{\lambda(T_{ijt}^{INJ} - \tau_t^{DSP})}, \quad (24)$$

$$\sum_{j \in \mathcal{V}, i \neq j, t \in \mathcal{T}} y_{jit} = 1, \quad (25)$$

$$\sum_{j \in \mathcal{V}, i \neq j} y_{jit} = \sum_{j \in \mathcal{V}, i \neq j} y_{ijt}, \quad (26)$$

$$W_{it} \leq w_{it} \leq W^{VEH}, \quad (27)$$

$$w_{it} \leq W^{VEH} + (W_{it} - W^{VEH})\, y_{0it}, \quad (28)$$

$$w_{jt} \geq w_{it} + W_{jt} - W^{VEH} + W^{VEH}\, y_{ijt} + \\ (W^{VEH} - W_{jt} - W_{it})\, y_{jit}, \quad (29)$$

$$\sum_{j \in \mathcal{V}_t^H} y_{1jt} \leq N_t^{DRV} \quad (30)$$

$$E_{it} \leq x_{it} \leq L_{it}, \quad (31)$$

$$x_{it} \geq \tau_t^{DSP} - T + (T + H_{0i})\, y_{0it}, \quad (32)$$

$$x_{it} \leq y_{i0t}\, (T - H_{i0} - S_i - L_{it}) + L_{it}, \quad (33)$$

$$x_{it} \geq E_{it} + (H_{0i} - E_{it}) y_{0it}, \quad (34)$$

$$x_{jt} \geq x_{it} - L_{it} + (L_{it} + H_{ij} + S_i)\, y_{ijt}, \quad (35)$$

$$\tau_{ijt}^{BOB},\ \tau_{ijt}^{EOB}, b_{ijt} \geq 0, \quad (36)$$

$$\tau_t^{EOU},\ \tau_t^{DSP},\ q_t^{EOU},\ q_t^{DSP} \geq 0, \quad (37)$$

$$w_{it}, x_{it} \geq 0, \quad (38)$$

$$y_{ijt}, z_{ijt} \in \{0, 1\}, \quad (39)$$

5.3 Model details

The objective function is defined in (14). We currently define the production cost as a multiple of the bombardment time needed at every period t. In terms of the transportation costs we define them as the distance traveled by each vehicle multiplied by the cost per mile plus the fuel surcharge (see equation (12)). We have added the fixed cost, F, for using a vehicle.

The constraints defined by inequalities (15) enforce the bombardment time of a target in a cyclotron to be in a specific interval, where $T_{\min}^{BMB}$ and $T_{\max}^{BMB}$ are the minimum and maximum allowed bombardment times. Similarly the constraints (16) deal with the lower and upper bounds of the beam current used.

Constraints (17) state that the bombardment of the first target will take place at T^{INIT}, which is defined as follows

$$T^{INIT} = T^0 + T^{CI} + T^{TL} + T^{TB}$$

where T^0 is the time when the pharmacist turns on a cyclotron for the first time (it represents the beginning of the production cycle of the first time period), T^{CI} is the time needed for cyclotron initialization, T^{TL} defines the time needed to load the targets, and T^{TB} is the time for tuning the beams before the bombardment. The values of T^{CI}, T^{TL} and T^{TB} can be seen on the first three boxes in Figure 1.

Constraints (18) ensure that for every cyclotron the end of the bombardment of all the targets must happen at the same time point. The requirement that at every stage each target in a cyclotron is unloaded sequentially after the end of the bombardment is described by constraint (19). On the other hand constraints (20) model the fact that the beginning of the bombardment in period t must happen after the end of unload (τ_{t-1}^{EOU}) in the previous period $t-1$, plus a fixed time, T^{PREP}, that is required to prepare for the next period (e.g., $T^{PREP} = 10$ minutes). This constraint links the production in consecutive periods.

The constraint that defines the amount of radioactivity that has been produced at the end of unload, in period t, is described in (21). Note that the first exponential term defines the total radioactivity produced during the bombardment of all available targets in all available cyclotrons. Once the bombardment of the targets has been finished the decay process starts. This is modeled by the second exponential in (21).

The formula that measures the amount of radioactivity after the transfer, chemistry and quality control is described in constraints (22). Furthermore, constraints (23) connect the time line between the end of unload and the dispensing times at every period. Constraints (24) connect the total amount of radioactivity that must be available at dispensing time (the left hand side in (24)) and the total demand of radioactivity at dispensing time during period t.

The remaining constraints are related to the distribution of FDG to imaging centers. To ensure that our routes are meaningful and avoid cycling, each customer should be visited once. After visiting a customer, we can only go for one customer next. These requirements are conveyed by constraints (25-26).

Each vehicle contains a number of containers that store the individual doses. Each container has a specific weight. We use the variable w_{it} to measure the

total weights delivered by a vehicle up to customer i along a specific route. As stated by constraint (27) this weight should always be less than or equal to the weight capacity of the vehicle, W^{VEH}. In addition, it should be greater than or equal to W_{it} which represents the total weight of all the order placed by customer i (see (13) for the definition of W_{it}).

Constraint (28) takes care of the case when customer i is the first one to be visited by the vehicle in period t. Indeed, when $y_{1it} = 1$, constraint (28) becomes

$$w_{it} \leq W_{it}, \ \forall i \in \mathcal{V}_t^H, \forall t \in \mathcal{T} \tag{40}$$

Combining constraints (27) and (40) gives us the stronger constraint $w_{it} = W_{it}$, $\forall i \in \mathcal{V}_t^H$. On the other hand, when customer i is not the first one to be visited, we have $y_{1it} = 0$ and constraint (28) becomes $w_{it} \leq W^{VEH}$ which is always satisfied and hence becomes redundant. By combining the two constraints (27) and (28) we are able to strengthen the feasible region of our problem resulting in a faster location of the integer optimal solution.

The case where customer i is not the first one to be visited deserves special attention. This case is taken care of by constraint (29). In this case the value of the variable w_{it} is equal to the weight of the orders of all the customers that were visited between the pharmacy and customer i itself. For example, if customer j is immediately after customer i in period t, then $y_{ijt} = 1$ and $y_{jit} = 0$. As a result, constraint (29) becomes $w_{jt} \geq w_{it} + W_{jt}$ which means that the weight delivered to customer j is at least equal to that delivered in customer i and the weight of the order of customer j. If, on the other hand, customer j is visited immediately before customer i then $y_{ijt} = 0$, $y_{jit} = 1$ and constraint (29) becomes $w_{jt} \geq w_{it} - W_{jt}$. This constraint states that the quantity delivered between the pharmacy and the customer j is not less than the quantity delivered between the pharmacy and customer i. In addition if customer j is visited immediately before customer i, we can deduce that we should always have $w_{it} \geq w_{jt} + W_{it}$. Combining the last two inequalities we obtain the stronger equality constraint $w_{it} = w_{jt} + W_{it}$. If customers i and j are not visited successively, then constraint (29) becomes $w_{jt} \geq w_{it} + W_{jt} - W^{VEH}$. By noting that the right hand side of the above constraint is always less than zero and by using the fact that $W_{it} \geq 0$ and the constraint (27), we can deduce that (29) becomes redundant.

At each period t we have a certain number of drivers (or equivalently, vehicles) available. This requirement is represented by constraint (30).

A very important requirement in the delivery of radio pharmaceuticals is the time window that a dose must arrive at a customer or imaging center. We use the variable x_i in order to measure the time when a vehicle arrives at customer i. Constraint (31) defines the time window that a vehicle is allowed to arrive at customer i. The upper bound L_{it} defines the latest time that the vehicle must arrive at the customer (see (11) for the definition of L_{it}).

Constraint (32) ensures that the arrival time of a vehicle is placed after the dispensing time, τ_t^{DSP}, during every period t. Usually when the vehicles arrive at a customer, the drivers have to fill in security forms, unload the containers and deliver them to the appropriate floor where the PET scanners are located. The time required to do these tasks is called *service time* and denoted by S_i. Constraints (33) ensures that every driver will not work more than the maximum allowable time T, excluding the time it takes to drive back to the radiopharmacy (H_{i0}) and the service time at the last customer i. In addition, constraints (34) connect the time it takes to drive from the radiopharmacy to the first customer and the early time corresponding to it. On the other hand, constraint (35) connects the arrival time between to consecutive locations. For example, if customer i precedes customer j at period t, then $y_{ijt} = 1$ and (35) becomes $x_{jt} \geq x_{it} + H_{ij} + S_i$ which defines a tighter lower bound than those defined by (31). Otherwise (i.e., $y_{ijt} = 0$), constraint (35) becomes $x_{jt} \geq x_{it} - L_{it}$ which is redundant because its right hand side is always negative due to the constraint (31).

Finally, all the variables are continuous except y_{ijt} and z_{ijt} which are binary. These requirements are described by constraints (36) to (39) in the model.

6 COMPUTATIONAL RESULTS

To illustrate the efficiency and practicality of the proposed optimization model we initially present a case study describing the production and distribution of orders during a typical day in a radio-pharmacy. Due to confidentiality of company and patient data we do not present the names or the locations of the medical imaging centers that placed orders. We solve the model by using the FICO-Xpress optimization package[5], which includes a powerful modeling language (Mosel) and efficient solvers that can solve problems with integer and continuous variables as well as linear and nonlinear constraints.

The current practice in the radio-pharmaceutical industry is to have employees (typically the senior pharmacists) to create the daily schedules for the production and distribution of FDG. Although the pharmacists have great domain knowledge and experience, oftentimes they come up with sub-optimal production or delivery schedules. A manually generated production schedule may produce larger quantities of

[5] http://www.fico.com

Fluorine-18 than those needed, resulting in the unnecessary use of cyclotrons and the waste of the extra radioactivity. In addition, the transportation routes determined by pharmacists are usually suboptimal resulting in routes that cost more to the company and may not guarantee the on-time arrival at a customer site. Another major deficiency is that having pharmacists determining the production schedules and delivery routes takes valuable time away from their main tasks and decreases their productive time by at least 30-60 minutes per day. We expect the optimization model we have developed to serve as a valuable decision support tool which will save them a lot of time and at the same time determine optimal production schedules and delivery routes saving a large amount of money for the pharmaceutical company.

The data for the model's parameters are obtained from the pharmacy's Enterprise Resource Management (ERM) system, which stores information related to the customer orders (e.g., the number of doses, their injection times, the time windows when the doses must arrive, the addresses of the customers, etc). The second source is the Google Geocoding API which can provide the distance and duration matrices of a set of customer locations.

For our case study, we selected a typical day which consists of 16 imaging centers (denoted by C1 to C16), requesting 97 doses in total. The day was split into two time periods. The first period deals with orders that have injection time between 07:00 and 10:30, whereas the second period handles the remaining orders. Table 1 presents more details about the orders. We can observe that there are customers that have ordered many doses and these are placed in different periods based on their injection time. Orders that need to be injected later in the day are placed in the second period. The CPU time needed to determine the optimal solution was 229 seconds.

Cust. ID	Doses in P_1	Earliest inj. in P_1	Doses in P_2	Earliest inj. P_2
C1	5	08:15	2	12:15
C2	4	07:30	3	11:15
C3	5	08:00	1	10:45
C4	3	08:50	3	12:30
C5	2	08:30	4	11:00
C6	6	08:15	-	-
C7	6	08:00	2	10:30
C8	2	08:15	4	11:15
C9	3	08:30	2	10:30
C10	4	08:30	2	12:00
C11	1	08:30	4	11:00
C12	3	09:00	4	12:00
C13	4	08:30	-	-
C14	6	08:00	-	-
C15	3	07:45	5	10:45
C16	3	08:15	3	11:00

Table 1: Details for the orders placed by imaging centers during the two time periods

6.1 Radioactivity production at each period

Tables 2 and 3 summarize the results for the production of radioactivity during Periods 1 and 2. The radiopharmacy we consider has two cyclotrons and each one can use up to two targets. The production for Period 1 started at 01:00 and both cyclotrons and all targets were used to produce enough radioactivity to satisfy the customer demand. The bombardment time for each of the four targets was 60 minutes, which translates to a total of 4 hours of radioactivity produced in just 1 hour. The bombardment on all targets started on 01:30 and ended 02:30[6]. At the end of bombardment the total radioactivity produced (by all targets) is 9,189.47 mCi. After 02:30 the produced radioactive material will start decaying continuously. The unload process of all targets was completed on 02:44[7] and the radioactivity is 8,412.38 mCi. The end of synthesis happens at 03:39[8] and the radioactivity has decayed to 4,952.43 mCi. After the completion of the QC and dispensing processes the radioactivity has reached 3,933.51 mCi. It is that amount of radioactivity that is going to be dispensed to cover the total FDG demand for period 1.

The production for satisfying the demand in Period 2 can start any time after the target unload process of Period 1 has been completed. We assume that production at Period 2 starts at 02:16 (i.e., 2 minutes after the end of target unload of Period 1). The total customer radioactivity demand for Period 2 at dispensing time is 1469.92 mCi. Once we take into account the 30 minutes preparation time preceding a bombardment, the beginning of bombardment in Period 2 is at 03:14. This time the model decided to use only one target to cover the demand for Period 2. The bombardment time was between 03:16 (we assume a 2 minutes gap between each period) and 05.40. The radioactivity at the end of bombardment is 3,614.2 mCi, which is substantially smaller than the corresponding radioactivity produced in Period 1. This is attributed to the following factors: (a) the imaging centers in Period 2 are closer to the radio-pharmacy (b) the injection times are closer to the beginning of the bombardment in Period 2 and (c) there are

[6] we take into consideration the preparation time before the bombardment starts, that is, cyclotron initialization time (15 minutes), target load time (10 minutes) and beam time tuning (5 minutes), as shown in Figure 1

[7] we assume that the targets in each cyclotron were unloaded in parallel, and therefore we need 2×7 minutes to unload all four targets

[8] 50 minutes after the end of unload of all targets

Period #	q_t^{EOB}	q_t^{EOU}	q_t^{EOS} in	q_t^{DSP}	q_t^{DSP-C}
1	9189.47	8412.38	4952.43	3933.51	3575.92
2	3614.2	3458.01	2035.76	1616.92	1469.92

Table 2: Results of the radioactivity production during Periods 1 and 2 (the last column records the total customer demand at dispensing time for each period)

Period #	τ_t^{EOB}	τ_t^{EOU}	τ_t^{EOS}	τ_t^{DSP}
1	02:30	02:44	04:00	04:00
2	05:40	05:47	07:07	07:07

Table 3: Results of the different production times during Periods 1 and 2 (note: we assume that the EOS coincides with the dispensing time; we can relax this constraint).

fewer customers and their demand of FDG is smaller. At the end of the unload of the single target the radioactivity has decayed to 3,458.01 mCi and at the end of synthesis and dispensing time it has become 2,035.76 mCi and 1,616.92 mCi, respectively. The total customer radioactivity demand at dispensing time is 1469.92 mCi.

6.2 Transportation of doses in each period

At each period the doses of FDG that have been produced are ready to be picked up by the drivers any time after the dispensing time. Depending on the number of imaging centers, their distance from the radio-pharmacy and the injection times of the individual doses, the transportation problem represents a very complex task for the pharmacists to solve. The optimization model we have developed is able to determine the most cost effective transportation routes that ensure the doses of every customer arrive before they are injected in the patients. We assume that orders can be delivered any time at all customer locations, even when they have not opened yet (i.e., we set E_{it}=00:00). The service time at each customer location was set to 10 minutes. Furthermore, the delivery vehicles must arrive at a customer location at least 30 minutes before the injection time of every dose (i.e., $p_i = 30$).

The total transportation cost for Period 1 was \$1,258.36 and the total distance traveled by all drivers was 950.96 miles. The optimal routes are shown in Figure 2.A. On the left side of every connecting arc there are two values. The values in parentheses denote the drive time from one location to the next, whereas the other values denote the distance (measured in miles). As can be seen a total of 6 vehicles were used to deliver the orders to all imaging centers. In contrast, the routes determined by the pharmacist (actual routes) cost \$1,493.83 and the total number of miles covered was 1,154.83. Figure 2.B shows the graph of the actual routes and Table 4 compares the total cost and distance for the optimal and actual routes. In addition, it summarizes the improvements we get when the optimal routes are used. As can be seen, there was a total of 15.76% reduction in transportation cost and 17.65% reduction in traveled distance. In addition the actual routes needed 8 vehicles, which represents an increase of two vehicles more than those determined by our model. This is quite important since it demonstrates better utilization of the vehicle capacity which is cost effective.

Period 1	Delivery cost	Travel distance	Number of vehicles
Optimal routes	1,258.36	950.96	6
Actual routes	1,493.83	1,154.83	8
Improvements	15.76%	17.65%	33.33%

Table 4: Summary of the improvements in total cost, distance traveled and vehicles used between the optimal and actual routes in Period 1.

Examining Figures 2.A and and 2.B closer we can see that our optimization model determined a much better way of delivering the orders to customers C10 and C11. More specifically, our model used one vehicle to travel to C10 and then to C11, whereas the pharmacist decided to use two vehicles to travel separately to C10 and C11. The total distance traveled by the vehicle of our model was 165+148=313 miles. On the other hand, the two vehicles sent by the pharmacist, traveled a total of 293+165=458 miles. It is these type of route consolidation that can provide substantial savings in traveling distance (in this case 458-313=145 miles), monetary cost and number of vehicles needed. An exactly similar situation arises with customers C9 and C13. Our model used one vehicle and traveled a total of 114+96=210 miles, whereas the pharmacist used two vehicles which traveled a total of 114+207=321 miles.

In Period 2 the distribution routes determined by our optimization model were more economical than those determined by the pharmacists. Table 5 summarizes the transportation cost improvements for Period 2. Due to space limitations we cannot include the graphs that compare the optimal and actual routes. Similar behavior to that shown in Figure 2 was observed.

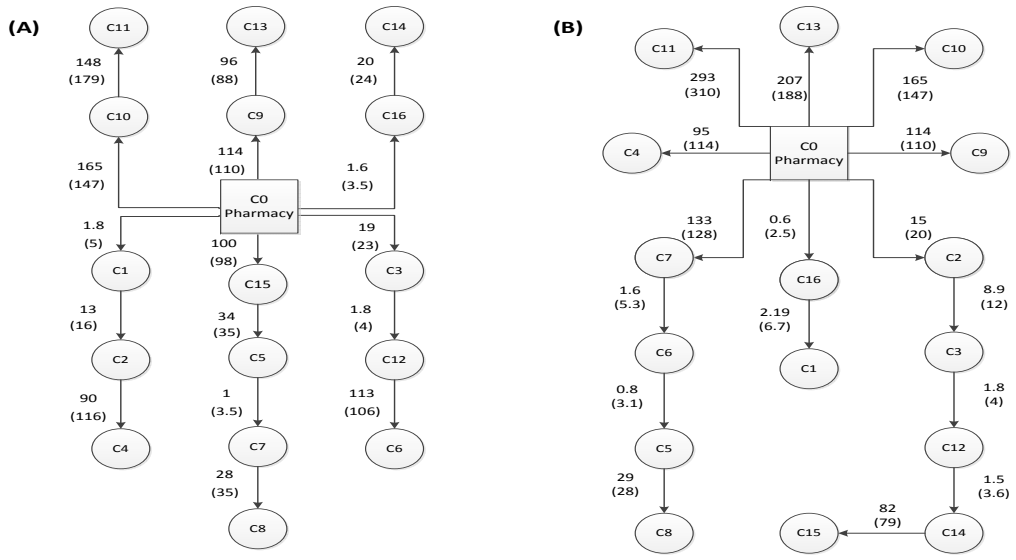

Figure 2: The optimal routes are shown in graph (A) and the actual routes in graph (B). The distances are measured in miles and the drive times (in parentheses) are measured in minutes. The doses of all orders are available for pickup by drivers on 04:00.

Period 2	Delivery cost	Travel distance	Number of vehicles
Optimal routes	954.87	717.2	5
Actual routes	1053.17	802.76	6
Improvements	9.33%	10.65%	16.66%

Table 5: Summary of the improvements in Period 2.

7 CONCLUSION

In this paper we discussed the production and distribution of radioactive tracers. Our focus was on fludeoxyglucose (FDG) which is one of the most widely used tracers in PET. We have developed a mixed integer nonlinear optimization model that efficiently integrates the production and delivery requirements in a multi period framework. The model is flexible enough to handle other types of radioactive tracers with minor modifications. The optimal solution determines the right amount of radioactivity needed to satisfy the daily demand and determines the most cost effective routes for delivering the FDG doses. Finally we presented a case study where we used real data from a typical day in a radiopharmacy and demonstrated that substantial savings can be achieved by our model.

REFERENCES

Burns, M., 2010. Market For PET Radiopharmaceuticals and PET Imaging, Report 320, Bio-Tech Systems Inc., 4167 Pinecrest Circle West, Las Vegas, Nevada 89121.

Floudas C.A., 1995. *Nonlinear and Mixed-Integer Optimization: Fundamentals and Applications*, Oxford University Press.

Jacobson, M.S., R. A. Steichen, and P. J. Peller, 2012. PET Radiochemistry and Radiopharmacy. *In: PET-CT and PET-MRI in Oncology, Peller et al. (Eds)*, Springer-Verlag Berlin Heidelberg.

Nagurney, A. and L.S. Nagurney, 2012. Medical nuclear supply chain design: A tractable network model and computational approach. *International Journal of Production Economics*, 140, p. 865-874.

Shen, Z. J., 2007. Integrated supply chain design models: A survey and future research directions. *Journal of Industrial Management Optimization*, 3(1), p. 1-27.

Simchi-Levi, D., Kaminsky, P., and Simchi-Levi, E., 1999. *Design and managing the supply chain: concepts, strategies and cases*, McGraw-Hill.

Sprawls P., 1995. *Physical principles of medical imaging*, Medical Physics Publishing Corp.

Snyder, L. V. and Z. J. Shen, 2011. *Fundamentals of Supply Chain Theory.* John Wiley & Sons.