INFORMS TutORials in Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Dynamic Optimization in Radiotherapy

Archis Ghate

To cite this entry: Archis Ghate. Dynamic Optimization in Radiotherapy. *In* INFORMS TutORials in Operations Research. Published online: 03 Sep 2014; 60–74.
https://doi.org/10.1287/educ.1110.0088

TUTORIALS IN
OPERATIONS RESEARCH
INFORMS 2011

 | ISBN 978-0-9843378-2-8
doi: http://dx.doi.org/10.1287/educ.1110.0088

Dynamic Optimization in Radiotherapy

Archis Ghate
Industrial and Systems Engineering, University of Washington, Seattle, Washington 98195,
archis@uw.edu

Abstract The goal in external beam radiotherapy for cancer is to maximize tumor damage while limiting toxic effects of radiation on nearby healthy anatomies. This is achieved through spatial localization and temporal dispersion of radiation dose. Once a radiation intensity profile that achieves the maximum possible spatial localization is designed at the beginning of a multiweek treatment course, the total planned dose is split into a series of predetermined *equal-dosage* fractions delivered daily so that healthy cells can recover between sessions. Thus, existing mathematical methods for treatment planning employ *static-deterministic* optimization techniques, and hence cannot adapt to a tumor's uncertain biological response over time. In this tutorial, we review a recently proposed stochastic control framework, where the ultimate objective is to design individualized treatment strategies that dynamically adapt to tumor response to deliver the right dose to the right location at the right time.

Keywords intensity-modulated radiation therapy; stochastic control; approximate dynamic programming

1. Introduction

The 2008 World Cancer Report (Boyle and Levin [8]) estimated that by 2030 there could be 17 million cancer deaths annually worldwide. It also predicted that by 2010, cancer would replace heart disease as the leading cause of deaths. In the United States, about half of all cancer patients undergo some form of radiotherapy.[1] In external beam radiotherapy, a linear accelerator (Figure 1) is used to pass ionizing radiation through the affected part of the patient's body to kill tumor cells. Radiation targeted toward the cancerous regions, however, also passes through other nearby anatomies, thus damaging both tumor and healthy cells. The goal therefore is to maximize the differential between tumor cell kill and damage to healthy cells. Radiotherapy planners attempt a two-pronged approach, which involves spatial localization and temporal dispersion of radiation dose,[2] to reach this objective.

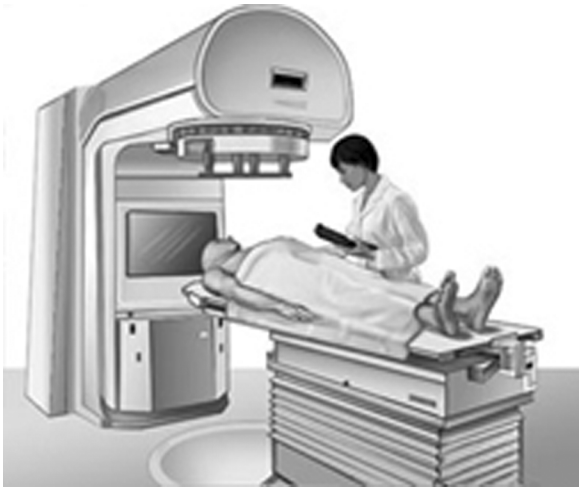
1.1. Spatial Localization of Dose

At the beginning of the treatment planning process, the cancerous targets and nearby unspecified normal tissue and organs at risk (OARs) (collectively, "normal tissue") are often delineated on an anatomical image such as a computed tomography or a magnetic resonance imaging scan. Intensity-modulated radiation therapy (IMRT) technology is then used to tune the radiation field's intensity profile so that the dose delivered to the tumor is much higher than that to the normal tissue (Webb [45]). This basic idea is shown in Figure 2(a). One way to effect precise intensity modulation involves controlling the position and speed of movements of the leaves of a multileaf collimator, shown in Figure 2(b), that is fitted in the

[1] Radiation Therapy for Cancer. National Cancer Institute website: http://www.cancer.gov/cancertopics/factsheet/Therapy/radiation. Accessed August 3, 2011.

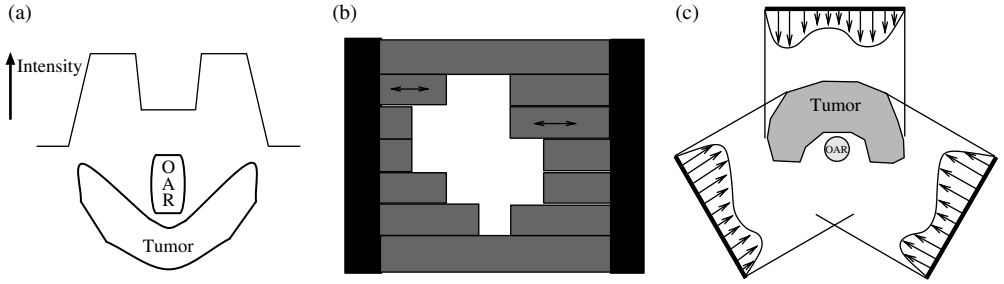
[2] Dose means energy per unit mass; its unit is grays (Gy); 1 Gy = 1 J/kg.

FIGURE 1. A linear accelerator for external beam radiotherapy.



Source. Radiation Therapy for Cancer. The National Cancer Institute: http://www.cancer.gov/cancertopics/factsheet/Therapy/radiation. Accessed August 3, 2011.

FIGURE 2. A schematic of the basic idea in IMRT.



Notes. (a) The radiation intensity profile is modulated so that a high dose is delivered to the tumor, whereas a much lower dose is delivered to the organ at risk. (b) A multileaf collimator is used to modulate intensities. The leaves can be moved in a lateral direction shown by the two arrows. (c) Multiple beams are often used for sufficient tumor coverage while sparing normal tissue. The picture shows three intensity-modulated beams with their intensity profiles marked by arrows.

aperture of the linear accelerator gantry. Radiation fields from multiple directions are often used as shown in Figure 2(c) to avoid tumor cold spots, that is, regions where an insufficient dose would otherwise be delivered in an attempt to spare the normal tissue.

In mathematical models, each beam is discretized into smaller segments called beamlets. A common practice is to perform inverse planning, that is, to start with a treatment protocol that prescribes a uniform high dose to the cancerous targets and sets a tolerance limit on the dose to normal tissue. Beamlet intensities are then optimized to attain a dose distribution that is as close as possible to this shape-conforming prescription. This is called the *fluence-map* problem. Several conformal treatment protocols and corresponding fluence-map optimization models and algorithms have been developed and embedded in commercial software over the years. Many of these are surveyed by Ehrgott et al. [10], Langer et al. [20], and Shepard et al. [33]. Basic background on these models is provided here. For notational brevity, only one tumor and one OAR are considered, ignoring unspecified normal tissue.

In a typical fluence-map optimization problem, regions within a patient's body are modeled using small three-dimensional (3D) cubes called voxels. Let n be the number of tumor voxels, and let m the number of OAR voxels. Let k be the number of radiation beamlets, and let A be the $(n+m)\times k$ dose deposition matrix. The tumor rows (assumed without loss

of generality to be the first n rows) of the matrix are denoted by A^τ, whereas the OAR rows are denoted by A^ϕ. Such a matrix is obtained using radiation transport simulations or pencil beam algorithms (Jelen et al. [16]). The tumor–dose vector is given by $d^\tau = A^\tau u \in \Re^n_+$; the dose delivered to tumor voxel i is its ith component, $d^\tau_i = [A^\tau u]_i$. Similarly, $d^\phi = A^\phi u \in \Re^m_+$ is the OAR dose vector, and $d^\phi_j = [A^\phi u]_j$ is the dose to OAR voxel j. Let D_{presc} be the dose prescribed by the radiation oncologist to the tumor, and let D_{tol} be the tolerance dose level for the OAR. Then one of the simplest fluence-map optimization models is given by

$$\min_{u\geq 0} \frac{w^\tau}{n} \sum_{i=1}^{n} ([A^\tau u]_i - D_{\text{presc}})^2 + \frac{w^\varnothing}{m} \sum_{j=1}^{m} (([A^\varnothing u]_j - D_{\text{tol}})_+)^2, \tag{1}$$

where $(x)_+ = \max\{0, x\}$, and the ws are artificial weights assigned to the tumor and the OAR (Ehrgott et al. [10]). The idea in this model is to minimize total squared deviation of tumor voxel dose from the prescription dose and to impose a penalty for overdosing OAR voxels beyond the tolerance level. It is common to include lower and upper bound constraints on dose to tumor voxels and upper bounds on dose to OAR voxels. Linear programming models and column generation algorithms for fluence-map optimization have also been studied (Romeijn et al. [31]). The dose-volume protocol, which states that no more than a fraction η of the OAR should receive a dose of more than Δ Gy, is a common source of nonconvexity in fluence-map optimization. Mixed integer constraints of the form

$$[A^\varnothing u]_j \leq \Delta + M y_j, \quad y_j \in \{0,1\},\ j = 1, 2, \ldots, m, \quad \sum_{j=1}^{m} y_j \leq \eta m, \tag{2}$$

where M is a large enough number, can model this (Shepard et al. [33]). Efforts to optimize the number and angles of radiation beams along with the beamlet intensity levels have also been made (Aleman et al. [2], Lee et al. [22]). Some recent work on fluence-map optimization has used nonuniform treatment protocols that utilize information from *biological* images taken at the outset of the treatment course (Alber et al. [1], Lawrence et al. [21], Ling et al. [24], South et al. [36], Yang and Xing [49]). For instance, a higher dose can be prescribed to regions of high clonogen[3] density within a tumor. Such treatment methods are termed biologically conformal radiotherapy.

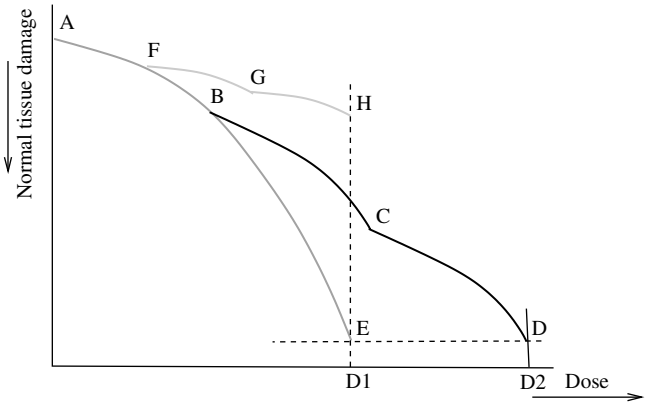
In this subsection we outlined the spatial trade-off in radiotherapy. The temporal aspects of treatment planning are discussed next.

1.2. Temporal Dispersion of Dose

Radiotherapy is typically administered daily over several weeks to give healthy cells time to recover between sessions. This is called fractionation (Figure 3). It is partly motivated by the empirical observation that healthy cells have better damage repair ability than tumors. More generally, repair of sublethal damage in normal tissue, redistribution and reoxygenation of tumor cells into more radiosensitive phases, and repopulation of tumor cells are important determinants of the efficacy of fractionated radiotherapy (Hall and Giaccia [14]). Progression of these complex dynamic biological processes, and hence a patient's biological response to treatment, are difficult to model, are uncertain, and vary depending on how the total dose is broken into fractions. Existing treatment planning methods take a static-deterministic view, where first an intensity profile is obtained by solving a fluence-map problem at the beginning of the treatment course, and then the corresponding total dose is delivered by breaking it into a series of *equal-dosage* fractions. Some representative exceptions to this are outlined next.

[3] A clonogen is a cell that produces its clones. Clonogen density within a tumor can be assessed using biological images listed later in this tutorial.

FIGURE 3. An illustration of a typical curve showing damage to normal tissue and of some benefits of fractionation for normal tissue.



Notes. The curve AE represents the damage to normal tissue as a function of dose administered in one treatment session. The dotted horizontal line marks the normal tissue damage if dose D1 is given in one treatment session. The curve AF-FG-GH, on the other hand, shows the damage inflicted when dose D1 is split into three equal-dose fractions, with curve AF representing the first fraction, FG showing the second fraction, and GH the third. Notice that dose D1 causes less damage when broken into these three fractions. As another "dual" example, curve AB-BC-CD shows that dose D2, in spite of being much higher than D1, causes the same normal tissue damage as D1 because D2 is split into three fractions. See Hall and Giaccia [14] for extensive details on such "cell-survival" curves.

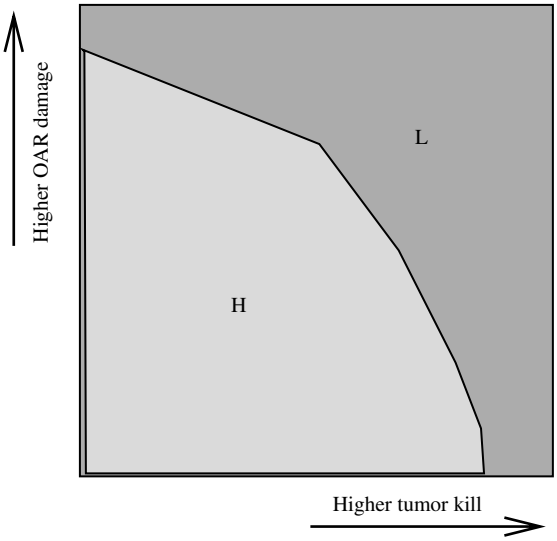
An optimal control approach to fractionation was attempted by Swan [40] with various tumor-response dynamics. But because that research predated the appropriate technological developments, it did not model the fluence-map problem. A similar approach was pursued by Wein et al. [46], although with more sophisticated tumor dynamics. Again, normal tissue and intensity modulation were ignored. These papers also did not consider uncertainties inherent in tumor response. In Kim et al. [18], we attempted to incorporate response uncertainties into the planning process. The fractionation problem was modeled as a finite-horizon Markov decision process, where the states related to the biological condition of the tumor and the OARs, and the decision variable in each session was the tumor prescription dose. For simplicity, we focused on two values for this dose—H for high and L for low. A monotone policy with structure roughly depicted in Figure 4 was observed to be optimal in numerical experiments. However, a simplistic route, somewhat similar to that in Swan [40], was followed in Kim et al. [18] in that the intensity profile was not optimized explicitly—the high-level idea was that a dynamically optimized prescription dose could be employed as an "input" to a stand-alone fluence-map problem solved for each session.

To summarize, state-of-the-art mathematical methods *decouple* the spatial and temporal aspects of radiotherapy, *ignore uncertainty* in dynamic biological processes, and in particular, *cannot adapt*[4] to the patient's actual response over the treatment course.

Recent technological advances in biological imaging have enabled researchers to track tumor response over time. For example, positron emission tomography (PET) radiolabeled with fluorine-18 (^{18}F) and fluorodeoxyglucose (FDG) tracer or magnetic resonance spectroscopic imaging (MRSI) of a choline/citrate ratio can provide spatiotemporal information about tumor clonogen density. PET with iodine-124-iododeoxyuridine can measure tumor

[4] A notable partial exception to this general observation is the recent work on *adaptive radiation therapy*, or *image-guided radiation therapy* (de la Zerda et al. [9], Ferris and Voelker [12], Rehbinder et al. [29], Sir [34], Sir et al. [35], Yan [48]), where the main focus is to adapt dose fractions to compensate for errors in dose delivery as a result of day-to-day and intrafraction variations in patient and equipment setup, patient movements during radiation delivery, and organ displacement. Even though this line of work has relied on control formulations, it is entirely different from the approaches of Kim [17] and Kim et al. [19], also reviewed in this tutorial.

FIGURE 4. A rough schematic of a monotone optimal policy from Kim et al. [18] for setting tumor prescription dose in a typical treatment session.



Notes. At a fixed level of OAR damage, a smaller dose is prescribed as more tumor is killed. For a fixed tumor kill, a smaller dose is prescribed as the OAR damage increases. In Kim et al. [18], some distortions to this intuitive structure were observed near the boundary states, and those are not shown in this picture; see Kim et al. [18] for details.

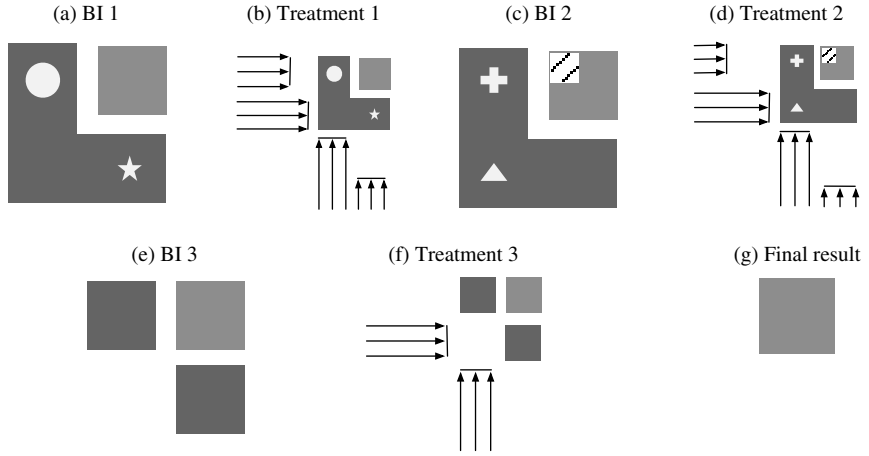
proliferation rates. ^{18}F-Misonidazole PET can map spatiotemporal evolution of hypoxia (lack of oxygen reduces radiosensitivity). See Bading and Shields [3], Eschmann et al. [11], South et al. [36], and Sovik et al. [37] for more details on such imaging methods. At the same time, mathematical models of tumor response are also becoming more sophisticated (Barazzuol et al. [5], Hanin [15], Powathil et al. [27], Rockne et al. [30], Schneider [32], Stavreva et al. [38], Swanson et al. [41]). Potential benefits of adapting treatment based on information obtained from biological images taken over several weeks of treatment, and to mathematical models of tumor response, have recently been documented (South et al. [36], Sovik et al. [37], Stewart and Li [39]).

In this tutorial, we review *dynamic biologically conformal radiation therapy* (DBCRT) (Kim [17], Kim et al. [19]), a stochastic control framework that proposes to utilize such biological images and tumor-response models to design spatiotemporally optimized, patient-specific, adaptive radiation treatment strategies that incorporate response uncertainty and that deliver the right dose to the right location at the right time.

2. Dynamic Biologically Conformal Radiation Therapy

Pictorial intuition behind DBCRT is provided in Figure 5, where a biological image (a), (c), and (e) is taken before each treatment session. Because a high clonogen density is observed in the northwest corner in (a), a very high dose is deposited in that region; hence, only a medium-low density of clonogens remains there. The clonogens proliferate to only a small extent; hence, (c) shows medium clonogen density in the southwest corner. On the negative side, the OAR receives a considerable dose in the first treatment session. This is compensated in the second session by depositing a lower dose in parts of the tumor and in the OAR. Therefore, parts of the tumor survive after the second session, which are eradicated in the third session with high-intensity beamlets that do not deposit dose in the OAR. In this example, DBCRT deposits at least a 33% smaller dose in the OAR compared to a static, equal-fractions approach and eradicates the tumor in three sessions; the static approach would take about six sessions to kill the tumor. Note that even though it might not have

FIGURE 5. Three chronologically ordered treatment sessions illustrating key conceptual ideas in DBCRT.



Notes. The L-shaped gray region is the tumor, whereas the darker square in the northeast corner is the OAR. The white shapes in the tumor mark clonogen density levels seen in biological images (BIs) such as FDG-PET or choline/citrate MRSI. Disk, high density; triangle, medium density; cross, medium-low density; star, low density. The small hatched squares indicate minor damage to the OAR. Treatment is delivered using two perpendicular beams, each discretized into six beamlets (arrows). Arrow height indicates beamlet intensity level: long, high intensity; medium length, medium intensity; short, medium low intensity; very short, low intensity. The patient's condition at the end of three treatment sessions is shown in (g).

been explicit in this schematic example, decisions in one radiotherapy session do affect those in other sessions and hence should not be made independently. DBCRT is a mathematical framework, rooted in discrete-time stochastic control, which formalizes these intuitive ideas.

In DBCRT, the cancerous target and the normal tissue are viewed as parts of a stochastic dynamic system whose states track a spatial map of relevant information observed in biological images. We assume that an image is acquired at the beginning of each treatment session (a discussion on imaging frequency is included in §2.4). Beamlet intensities in each treatment session are the control variables. The objective is to maximize a quantitative measure of the efficacy of adaptive treatment.

In Kim [17] and Kim et al. [19], a general stochastic control formulation of the problem outlined above is presented. The specific structure of this control problem depends on the choices made for defining states of the tumor–normal tissue system, the corresponding system-dynamics equations, the uncertainty model used, and the choice of treatment efficacy measure. Various possibilities for these are described in Kim [17], and we expect that many more will emerge in the future. In this tutorial, we present key ideas in DBCRT through perhaps the simplest yet intuitive and sufficiently insightful choices for these entities—tumor states are defined through clonogen densities in its subregions, OAR states use the concept of biologically effective dose (BED), and the treatment objective is to minimize the total number of tumor cells while limiting OAR damage. The discussion here closely follows Kim [17] and Kim et al. [19], but it is simplified in notation and excludes several details about the algorithms, design of numerical experiments, and computational results.

We use the following notation in the rest of this tutorial:

- $T =$ the number of treatment sessions indexed $t = 1, 2, \ldots, T$;
- $k =$ the number of beamlets;
- $u_t \in \Re_+^k$ is the k-dimensional vector of beamlet intensities used in treatment session t;
- $n =$ the number of tumor subregions indexed $i = 1, 2, \ldots, n$;
- $m =$ the number of OAR subregions indexed $j = 1, 2, \ldots, m$;
- $A =$ the $(n + m) \times k$-dimensional dose deposition matrix as described in §1.

Notice again that we focus on the tumor and OARs and ignore unspecified normal tissue for brevity. Moreover, we assume for simplicity in notation that the m OAR subregions come from the same anatomy. This is easily relaxed by putting subscripts j on various OAR parameters that appear in the discussion below. Finally, we assume that all tumor subregions have the same volume, and similarly for all OAR subregions.

2.1. OAR and Tumor State Definitions and Dynamics

An initial difficulty in designing a dynamic treatment strategy lies in appropriately capturing the effect of a possibly nonconstant sequence of doses on the OAR. In particular, we need a quantitative way to tell whether a dynamic treatment scheme can be tolerated by an OAR. This challenge arises partly because empirical evidence suggests that the biological effect of dose on OAR is not additive across treatment sessions (Hall and Giaccia [14]). For example, the OAR damage induced by two sessions with 2 and 3 Gy, respectively, is not the same as that inflicted by those with 1 and 4 Gy. In addition, traditional dose-tolerance limits for OARs are based on equal-dosage fractionation schemes. For instance, suppose a hypothetical OAR is known to tolerate a total of 20 Gy if delivered in 10 sessions of 2 Gy each (this is called a 2 Gy $\times$ 10 treatment). Owing to the lack of additivity of biological effects, this need not imply that the OAR will be able to tolerate a 10 Gy $\times$ 2 treatment, or three sessions with 10, 7, and 3 Gy, respectively. Figure 3 illustrates these ideas.

One way to tackle this is to use the notion of the BED (Hall and Giaccia [14]). To understand this concept, we need to first introduce the well-known log-linear-quadratic model of OAR cell kill (Hall and Giaccia [14]). According to this model, the fraction of OAR cells that survive a treatment session in which the OAR receives a dose of d Gy is given by

$$\exp(-\alpha d - \beta d^2). \tag{3}$$

Here, α and β are parameters that depend on the OAR under consideration. This implies that the fraction of OAR cells that survive t treatment sessions with doses $d_1, d_2, \ldots, d_t$, respectively, equals

$$\exp\left(-\alpha \sum_{l=1}^{t} d_l - \beta \sum_{l=1}^{t} d_l^2\right). \tag{4}$$

Thus, according to the log-linear-quadratic model, the quantity

$$\alpha \sum_{l=1}^{t} d_l + \beta \sum_{l=1}^{t} d_l^2 \tag{5}$$

determines the OAR damage induced by the sequence $(d_1, d_2, \ldots, d_t)$. In fact, it is more common to work with ratios (α/β), and hence we consider the expression by

$$\sum_{l=1}^{t} d_l + \left(\sum_{l=1}^{t} d_l^2\right) \Big/ (\alpha/\beta) \tag{6}$$

obtained after dividing (5) by α. This expression is called the biologically effective dose of the dose sequence $(d_1, d_2, \ldots, d_t)$. Now suppose an OAR is known to tolerate a conventional equal-dosage fractionation scheme of c Gy per session for N sessions. Note that the BED of this scheme equals

$$cN(1 + c/(\alpha/\beta)). \tag{7}$$

Here, cN is the total dose delivered to the OAR and is denoted by D_{tol} for "tolerance dose." We assume that the sequence $(d_1, d_2, \ldots, d_t)$ can be tolerated by this OAR if its BED is less than the conventional BED, that is, if

$$\left[\sum_{l=1}^{t} d_l + \left(\sum_{l=1}^{t} d_l^2\right) \Big/ (\alpha/\beta)\right] \le D_{\text{tol}}(1 + c/(\alpha/\beta)). \tag{8}$$

This allows us to define the set of feasible beamlet intensities in (11) below. Another advantage of BED is that it yields a simple BED-update formula from one treatment session to the next. To see this, consider the sequence $(d_1, d_2, \ldots, d_t, d_{t+1})$. Let BED_{t+1} denote its BED. Then we have

$$\begin{aligned}\text{BED}_{t+1} &= \sum_{l=1}^{t+1} d_l + \left(\sum_{l=1}^{t+1} d_l^2\right)\Big/(\alpha/\beta) = \sum_{l=1}^{t} d_l + d_{t+1} + \left(\sum_{l=1}^{t} d_l^2\right)\Big/(\alpha/\beta) + d_{t+1}^2/(\alpha/\beta) \\ &= \text{BED}_t + d_{t+1} + d_{t+1}^2/(\alpha/\beta). \end{aligned} \tag{9}$$

Based on this discussion, we use the BED of the dose sequence delivered in treatment sessions $1, 2, \ldots, t-1$ to OAR subregion j as its state at the beginning of session t. We denote this state by $x_{t,j}^{\phi}$. Let x_t^{ϕ} denote the m-dimensional column vector with components $x_{t,j}^{\phi}$ for $j = 1, 2, \ldots, m$. Using other notation set forth earlier, this leads to the OAR state dynamics

$$x_{t+1,j}^{\phi} = x_{t,j}^{\phi} + [A^{\phi} u_t]_j + [([A^{\phi} u_t]_j)^2]/(\alpha/\beta), \tag{10}$$

starting with $x_{1,j}^{\phi} = 0$ for all $j = 1, 2, \ldots, m$, because the initial BED is zero. Thus, the set of feasible beamlet intensities in treatment session t is given by

$$\left\{ u_t \in \Re_+^k \colon x_{t,j}^{\phi} + [A^{\phi} u_t]_j + \frac{1}{\alpha/\beta}([A^{\phi} u_t]_j)^2 \le D_{\text{tol}}(1 + c/(\alpha/\beta)) \right\}. \tag{11}$$

Now we discuss tumor dynamics.

Let $x_{t,i}^{\tau}$ denote the density of tumor cells in tumor subregion i at the beginning of treatment session t as observed, say, in a PET or MRSI scan. The n-dimensional column vector x_t^{τ}, with components $x_{t,i}^{\tau}$ for $i = 1, 2, \ldots, n$, defines the biological state of the tumor. We use a stochastic log-linear tumor cell-kill model here. In particular, the cell density in tumor subregion i is assumed to evolve according to

$$x_{t+1,i}^{\tau} = x_{t,i}^{\tau} \exp(-\theta_{t,i}[A^{\tau} u_t]_i), \quad i = 1, 2, \ldots, n, \tag{12}$$

where $\theta_{t,i}$ are nonnegative random variables. In this tutorial, we make the simplifying assumption that random variables $\theta_{t,i}$ are independent across treatment sessions t as well as across subregions. For instance, $\theta_{t,i}$ can be thought of as $\theta\epsilon_{t,i}$ or as $\theta + \epsilon_{t,i}$ where θ is a deterministic constant, and $\epsilon_{t,i}$ is random noise independent across i and t. This equation is derived from the more well-known deterministic log-linear-quadratic model (Hall and Giaccia [14]) using the following reasoning. Williams et al. [47] surveyed 48 different types of tumors and found that the logarithmic cell kill is relatively linear over the dose range of clinical interest. A random variable is used here as the parameter of the cell-kill model to incorporate uncertainties in tumor response into the planning process.

We use the total number of tumor cells remaining (TNTCR) at the end of the treatment course as the measure of treatment efficacy. Because of our assumption that tumor subregion volumes are equal, the TNTCR is proportional to $\sum_{i=1}^{n} x_{T+1,i}^{\tau}$, which is given by

$$\sum_{i=1}^{n} x_{1,i}^{\tau} \prod_{t=1}^{T} \exp(-\theta_{t,i}[A^{\tau} u_t]_i). \tag{13}$$

In the next subsection, we present a stochastic control formulation based on the above discussion.

2.2. A Stochastic Control Formulation

We denote the state of the tumor–OAR system at the beginning of session t by

$$x_t = \begin{bmatrix} x_t^\tau \\ x_t^\phi \end{bmatrix}. \tag{14}$$

Given that the initial tumor cell density vector is x_1^τ, and the initial BED is the zero vector, the set of all possible tumor–OAR states x_t at the beginning of session t is denoted S_t. We denote the set of feasible intensity vectors $u_t \in \Re_+^k$ as defined in (11) by $U_t(x_t)$. The set of all treatment policies is denoted Π. This Π is the set of all sequences $\pi = (\pi_1, \ldots, \pi_t, \ldots, \pi_T)$, where π_t, for each $t = 1, 2, \ldots, T$, is a mapping that assigns a feasible intensity vector $\pi_t(x_t) \in U_t(x_t)$ to state $x_t \in S_t$. Then, DBCRT calls for finding a treatment policy $\pi \in \Pi$ that solves the stochastic control problem

$$E\left(\sum_{i=1}^{n} x_{1,i}^\tau \prod_{t=1}^{T} \exp(-\theta_{t,i}[A^\tau \pi_t(x_t)]_i) \right) \tag{15}$$

subject to dynamics

$$x_{t+1,i}^\tau = x_{t,i}^\tau \exp(-\theta_{t,i}[A^\tau \pi_t(x_t)]_i), \quad i = 1, 2, \ldots, n,\ t = 1, 2, \ldots, T;$$
$$x_{t+1,j}^\phi = x_{t,j}^\phi + [A^\phi \pi_t(x_t)]_j + [([A^\phi \pi_t(x_t)]_j)^2]/(\alpha/\beta), \quad j = 1, \ldots, m,\ t = 1, 2, \ldots, T.$$

Backward recursive solution of Bellman's equations of dynamic programming is a standard theoretical technique for computing an optimal policy for (15) (Bertsekas [6]). It is not implementable here, however, because it calls for the solution of an *uncountable* number of minimization problems. A discretization is not practical either owing to the curse of dimensionality (Bertsekas [6], Powell [28]). We therefore outline a few other possibilities, used in Kim [17] and Kim et al. [19], for approximately solving the above stochastic control problem. The books by Bertsekas [6] and Powell [28] are excellent resources for more details about these and other algorithms for approximate dynamic programming.

2.3. Some Options for Approximate Dynamic Programming

2.3.1. Deterministic Control. The simplest approach is to ignore all uncertainty assuming the $\theta_{t,i}$s take some nominal values $\bar{\theta}_{t,i}$, say, their expected values, and solve a deterministic version of (15). It suffices to search over sequences of feasible beamlet intensities $u_1, u_2, \ldots, u_T$ based solely on the initial state. This is called an open loop solution. We view this static-deterministic method as the base case because it resembles current practice. BED is monotone increasing over t, and therefore, in a static problem, it suffices to enforce the BED constraints only at $t = T$. Because the initial BED is zero, the problem simplifies to

$$\min_{\substack{u_t \in \Re_+^k \\ t=1,2,\ldots,T}} \left(\sum_{i=1}^{n} x_{1,i}^\tau \prod_{t=1}^{T} \exp(-\bar{\theta}_{t,i}[A^\tau u_t]_i) \right) \tag{16}$$

subject to

$$\left[\sum_{t=1}^{T} [A^\phi u_t]_j + \left(\sum_{t=1}^{T} ([A^\phi u_t]_j)^2 \right) \Big/ (\alpha/\beta) \right] \le D_{\text{tol}}(1 + c/(\alpha/\beta)), \quad j = 1, 2, \ldots, m.$$

It is shown in Kim et al. [19] that problem (16) is convex. With recent advances in convex programming (Boyd and Vandenberghe [7]), we expect that it should typically be possible to solve it efficiently.

2.3.2. Open Loop Control. The open loop control (OLC) approach finds an open loop solution by solving the convex stochastic programming problem

$$\min_{\substack{u_t \in \Re_+^k \\ t=1,2,\ldots,T}} E\left(\sum_{i=1}^{n} x_{1,i}^{\tau} \prod_{t=1}^{T} \exp(-\theta_{t,i}[A^{\tau} u_t]_i) \right) \tag{17}$$

subject to

$$\left[\sum_{t=1}^{T} [A^{\phi} u_t]_j + \left(\sum_{t=1}^{T} ([A^{\phi} u_t]_j)^2 \right) \Big/ (\alpha/\beta) \right] \le D_{\text{tol}}(1 + c/(\alpha/\beta)), \quad j = 1, 2, \ldots, m.$$

2.3.3. Certainty Equivalent Control. The certainty equivalent control (CEC) method begins by replacing future uncertainty with nominal values $\bar{\theta}_{t,i}$, thus leading to a deterministic problem. It is solved to obtain a sequence of controls; only the first control is implemented, discarding the rest. The system then evolves stochastically to the next state, and the process is repeated until termination. Thus, T deterministic problems need to be solved over a treatment course of T sessions; the tth problem, starting in state x_t, is given by

$$\min_{\substack{u_l \in \Re_+^k \\ l=t,\ldots,T}} \left(\sum_{i=1}^{n} x_{t,i}^{\tau} \prod_{l=t}^{T} \exp(-\bar{\theta}_{l,i}[A^{\tau} u_l]_i) \right) \tag{18}$$

subject to

$$\left[x_{t,j}^{\phi} + \sum_{l=t}^{T} [A^{\phi} u_l]_j + \left(\sum_{l=t}^{T} ([A^{\phi} u_l]_j)^2 \right) \Big/ (\alpha/\beta) \right] \le D_{\text{tol}}(1 + c/(\alpha/\beta)), \quad j = 1, 2, \ldots, m.$$

It is shown in Kim et al. [19] that problem (18) is also convex.

2.3.4. Open Loop Feedback Control. The open loop feedback control (OLFC) technique is similar to CEC but future uncertainty is not replaced with nominal values; rather an open loop expected value minimization is performed. Thus, T convex stochastic programming problems of the form (19) are solved:

$$\min_{\substack{u_l \in \Re_+^k \\ l=t,\ldots,T}} E\left(\sum_{i=1}^{n} x_{t,i}^{\tau} \prod_{l=t}^{T} \exp(-\theta_{l,i}[A^{\tau} u_l]_i) \right) \tag{19}$$

subject to

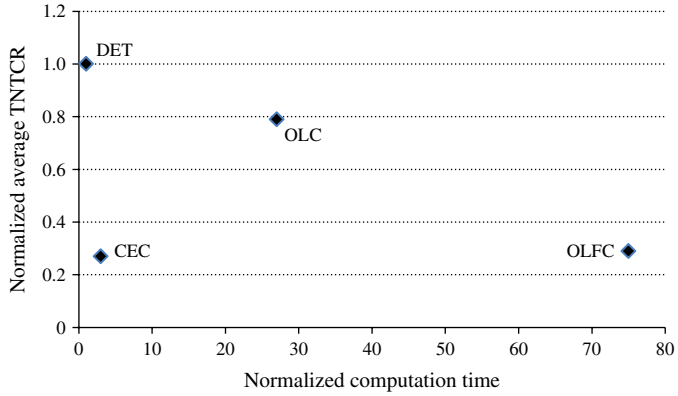
$$\left[x_{t,j}^{\phi} + \sum_{l=t}^{T} [A^{\phi} u_l]_j + \left(\sum_{l=t}^{T} ([A^{\phi} u_l]_j)^2 \right) \Big/ (\alpha/\beta) \right] \le D_{\text{tol}}(1 + c/(\alpha/\beta)), \quad j = 1, 2, \ldots, m.$$

2.4. Numerical Experiments

Extensive numerical experiments and detailed analyses of their results for multiple computer-generated tumor–OAR prototypes are presented in Kim [17] and Kim et al. [19]. Here, we briefly outline the key observations from some of those results. The initial experiments focussed on demonstrating, using prototypes with spatially as well as temporally inhomogeneous[5] tumors, that DBCRT can simultaneously balance spatiotemporal trade-offs in radiotherapy—a capability that existing mathematical optimization methods do not have.

[5] Spatial inhomogeneity here implies varying clonogen density within the tumor, and temporal inhomogeneity implies varying radiosensitivity over treatment sessions.

FIGURE 6. A comparison of four approximate solution methods.



Notes. The x-axis shows computation times normalized with respect to the computation time of deterministic control (DET). The y-axis shows the normalized average TNTCR over independent Monte Carlo simulations. The graph was plotted using numbers reported in Kim [17] and Kim et al. [19].

A comparison of the above four approximate control methods using a computer-generated U-shaped tumor–OAR geometry similar to that in Figure 2(c) was then performed. In these experiments, the tumor-radiosensitivity parameters $\alpha_{t,i}$ were assumed to be independent and identically distributed scaled Beta,[6] and the TNTCR objectives, averaged over independent simulations, were compared. Based on the computation times and the corresponding TNTCR values shown in Figure 6, CEC was seen to be an effective approximation method because it was fast and yet achieved a relatively low TNTCR. The average TNTCR with CEC was about 73% less as compared to deterministic control, which, as stated earlier, is a static-deterministic approach and hence is conceptually analogous to existing fluence-map optimization techniques. Additional numerical experiments with CEC using a simplified 2D head-and-neck cancer geometry showed 64% and 98% improvement in average TNTCR compared to deterministic control, respectively, for low and high degrees of clonogen inhomogeneity within the tumor. The impact of imaging frequency will likely be an important factor in future clinical implementation of DBCRT. How many imaging sessions will be economically and clinically feasible over a treatment course remains to be seen. Therefore, we wanted to quantify the expected value of information obtained through state observations. In Kim [17], the DBCRT framework was extended to the case where an image, i.e., a state observation, is acquired every $K \geq 1$ treatment sessions. Figure 7 shows the decrease in the TNTCR with increasing imaging frequency for one prototype in Kim [17] with 12 treatment sessions. The figure suggests that even one additional observation in the middle of the course may reduce the TNTCR. This is consistent with conjectures made by Stewart and Li [39] regarding patient-specific radiotherapy based on clinical intuition.

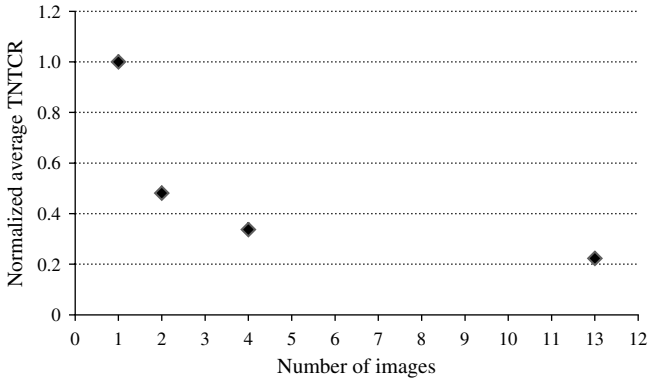
3. Clinical Challenges and Opportunities for Future Research

DBCRT is in its early stages of development. It is currently only a mathematical modeling and optimization framework. Several practical challenges will need to be overcome if it is to become a clinical reality. Some of the hurdles we foresee in this endeavor, possible solutions, and corresponding research opportunities are briefly outlined in this section.

First and foremost, the aforementioned and other efficient approximate solution algorithms will have to be tested extensively on realistic-scale computer-generated 3D test cases, perhaps using commercial treatment planning systems. This will have to be followed by clinical trials.

[6] We wanted to use a probability density function that was symmetric about its mean, was maximized at the mean, and had a finite nonnegative support. A truncated normal produced qualitatively similar results.

FIGURE 7. Decreasing TNTCR with increasing imaging frequency (normalized with respect to imaging only at the beginning of the treatment course).



Note. The graph was plotted using numbers reported in Kim [17].

One difficulty is related to imaging quality—how much biological information can accurately be gleaned from biological images to effectively guide decisions? Clearly, the better the biological image quality, the larger the potential benefit of an adaptive method such as DBCRT. Significant progress has been made in this direction over the last decade (Bar-Shalom et al. [4], Giesel et al. [13], Levin [23], Nishioka et al. [25], Noz et al. [26], Townsend and Beyer [42], Townsend and Cherry [43]), and medical physicists and doctors now increasingly seem to believe that patient-specific biologically guided radiotherapy will be technologically feasible within a decade (for a detailed discussion and a pictorial representation of a strategy for clinical implementation, see Stewart and Li [39], p. 3747). On the mathematical side, imaging noise could be incorporated using imperfect state information models in stochastic control (Bertsekas [6], van der Heijden et al. [44]).

The simplest radiation–response models were employed in this tutorial. As more sophisticated models become available, they could be incorporated into the general DBCRT framework in Kim [17] and Kim et al. [19]. No matter which response models are used, for truly patient-specific treatment, parameters of the models will have to be estimated for each individual patient based on information gathered from biological images taken over the treatment course. Several statistical estimation methods could be investigated. Such an integrated learning and control framework would fall within the realm of *adaptive control* (Bertsekas [6]) and would provide interesting opportunities for future research.

In Kim [17] and Kim et al. [18, 19], the total number of treatment sessions was fixed a priori. Although this is a common practice in radiotherapy, there is a considerable debate about what the "optimal" number of treatment sessions should be. To address this, a preliminary optimal stopping variation of DBCRT was introduced in Kim [17]. The idea was to decide, at each treatment session, whether it is worthwhile to continue treatment; if treatment is to be continued, then an optimal beamlet intensity profile is computed and implemented in that treatment session. In that approach, treatment is continued if the benefit of an additional session exceeds its "cost," which may include financial, patient inconvenience, and normal tissue-damage components. We foresee interesting opportunities for algorithmic research in this area.

Acknowledgments

The author appreciates financial support from the National Science Foundation through CAREER Award Grant CMMI-1054026. He is grateful to his former doctoral student, Minsun Kim, who taught him almost everything he knows about radiotherapy. The DBCRT framework reviewed in this tutorial was developed in her dissertation. He also thanks Minsun for carefully reading an early draft of this tutorial.

References

[1] M. Alber, F. Paulsen, S. M. Eschmann, and H. J. Machulla. On biologically conformal boost dose optimization. *Physics in Medicine and Biology* 48(2):N31–N35, 2003.

[2] D. Aleman, A. Kumar, R. K. Ahuja, H. E. Romeijn, and J. F. Dempsey. Neighborhood search approaches to beam orientation optimization in intensity modulated radiation therapy treatment planning. *Journal of Global Optimization* 4:587–607, 2008.

[3] J. R. Bading and A. F. Shields. Imaging of cell proliferation: Status and prospects. *Journal of Nuclear Medicine* 49(6):64S–80S, 2008.

[4] R. Bar-Shalom, N. Yefremov, L. Guralnik, D. Gaitini, A. Frenkel, A. Kuten, H. Altman, Z. Keidar, and O. Israel. Clinical performance of PET/CT in evaluation of cancer: Additional value for diagnostic imaging and patient management. *Journal of Nuclear Medicine* 44(8):1200–1209, 2003.

[5] L. Barazzuol, N. G. Burnet, R. Jena, B. Jones, S. J. Jefferies, and N. F. Kirkby. A mathematical model of brain tumour response to radiotherapy and chemotherapy considering radiobiological aspects. *Journal of Theoretical Biology* 262(7):553–565, 2010.

[6] D. P. Bertsekas. *Dynamic Programming and Optimal Control*, Vols. 1 and 2, 3rd ed. Athena Scientific, Nashua, NH, 2007.

[7] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, UK, 2004.

[8] P. Boyle and B. Levin. World cancer report. Technical report, International Agency for Research on Cancer, Lyon, France, 2008.

[9] A. de la Zerda, B. Armbruster, and L. Xing. Formulating adaptive radiation therapy (art) treatment planning into a closed-loop control framework. *Physics in Medicine and Biology* 52(14):4137–4153, 2007.

[10] M. Ehrgott, C. Guler, H. W. Hamacher, and L. Mathematical optimization in intensity modluated radiation therapy. *4OR* 6(3):199–262, 2008.

[11] S. M. Eschmann, F. Paulsen, C. Bedeshem, H.-J. Machulla, T. Hehr, M. Bamberg, and R. Bares. Hypoxia-imaging with 18-F-misonidazole and PET: Changes of kinetics during radiotherapy of head-and-neck cancer. *Radiotherapy and Oncology* 83(3):406–410, 2007.

[12] M. C. Ferris and M. M. Voelker. Fractionation in radiation treatment planning. *Mathematical Programming* 101(2):387–413, 2004.

[13] F. L. Giesel, A. Mehndiratta, J. Locklin, M. J. McAuliffe, S. White, P. L. Choyke, M. V. Knopp, B. J. Wood, U. Haberkorn, and H. von Tengg-Kobligk. Image fusion using CT, MRI and PET for treatment planning, navigation and follow up in percutaneous RFA. *Experimental Oncology* 31(2):106–14, 2009.

[14] E. J. Hall and A. J. Giaccia. *Radiobiology for the Radiologist*, 6th ed. Lippincott Williams & Wilkins, Philadelphia, 2006.

[15] L. G. Hanin. A stochastic model of tumor response to fractionated radiation: Limit theorems and rate of convergence. *Mathematical Biosciences* 191(1):1–17, 2004.

[16] U. Jelen, M. Sohn, and M. Alber. A finite pencil beam for IMRT dose optimization. *Medical Physics* 50(8):1747–1766, 2005.

[17] M. Kim. A mathematical framework for spatiotemporal optimality in radiation therapy. Ph.D. thesis, University of Washington, Seattle, 2010.

[18] M. Kim, A. Ghate, and M. H. Phillips. A Markov decision process approach to temporal mod ulation of dose fractions in radiation therapy planning. *Physics in Medicine and Biology* 54(14):4455–4476, 2009.

[19] M. Kim, A. Ghate, and M. H. Phillips. A stochastic control formalism for dynamic biologically conformal radiation therapy. *European Journal Of Operational Research*, forthcoming, 2011.

[20] M. Langer, E. K. Lee, J. O. Deasy, R. L. Rardin, and J. A. Deye. Operations research applied to radiotherapy, an NCI-NSF-sponsored workshop, February 7–9, 2002. *International Journal of Radiation Oncology Biology Physics* 57(3):762–768, 2003.

[21] Y. R. Lawrence, M. Werner-Wasik, and A. P. Dicker. Biologically conformal treatment: Biomarkers and functional imaging in radiation oncology. *Future Oncology* 4(5):689–704, 2008.

[22] E. K. Lee, T. Fox, and I. Crocker. Simultaneous beam geometry and intensity map optimization in intensity-modulated radiation therapy treatment planning. *International Journal of Radiation Oncology Biology Physics* 119:165–181, 2003.

[23] C. S. Levin. Primer on molecular imaging technology. *European Journal of Nuclear Medicine and Molecular Imaging* 32(14):S325–S345, 2005.

[24] C. C. Ling, J. Humm, S. Larson, H. Amols, Z. Fuks, S. Leibel, and J. A. Koutcher. Towards multidimensional radiotherapy (MD-CRT): Biological imaging and biological conformity. *International Journal of Radiation Oncology Biology Physics* 47(3):551–560, 2000.

[25] T. Nishioka, T. Shiga, H. Shirato, E. Tsukamoto, K. Tsuchiya, T. Kato, K. Ohmori, et al. Image fusion between 18FDG-PET and MRI/CT for radiotherapy planning of oropharyngeal and nasopharyngeal carcinomas. *International Journal of Radiation Oncology Biology Physics* 53(4):1051–1057, 2002.

[26] M. E. Noz, G. Q. Maguire Jr, M. P. Zeleznik, E. L. Kramer, F. Mahmoud, and J. Crafoord. A versatile functional anatomic image fusion method for volume data sets. *Journal of Medical Systems* 25(5):297–307, 2001.

[27] G. Powathil, M. Kohandel, S. Sivaloganathan, A. Oza, M. Milosevic. Mathematical modeling of brain tumors: Effects of radiotherapy and chemotherapy. *Physics in Medicine and Biology* 52(11):3291–3306, 2007.

[28] W. Powell. *Approximate Dynamic Programming.* John Wiley & Sons, NJ, 2007.

[29] H. Rehbinder, C. Forsgren, and J. Lof. Adaptive radiation therapy for compensation of errors in patient setup and treatment delivery. *Medical Physics* 31(12):3363–3371, 2004.

[30] R. Rockne, E. C. Alvord Jr, J. K. Rockhill, and K. R. Swanson. A mathematical model for brain tumor response to radiation therapy. *Journal of Mathematical Biology* 58(4–5):561–578, 2009.

[31] H. E. Romeijn, R. K. Ahuja, J. F. Dempsey, and A. Kumar. A new linear programming approach to radiation therapy treatment planning problems. *Operations Research* 54(2):201–216, 2006.

[32] U. Schneider. Mechanistic model of radiation-induced cancer after fractionated radiotherapy using the linear-quadratic formula. *Medical Physics* 36(4):1138–1143, 2009.

[33] D. M. Shepard, M. C. Ferris, G. H. Olivera, and T. R. Mackie. Optimizing the delivery of radiation therapy to cancer patients. *SIAM Review* 41(4):721–744, 1999.

[34] M. Y. Sir. Optimization of radiotherapy considering uncertainties caused by daily setup procedures and organ motion. Ph.D. thesis, University of Michigan, Ann Arbor, 2007.

[35] M. Y. Sir, M. A. Epelman, and S. M. Pollock. Stochastic programming for off-line adaptive radiotherapy. *Annals of Opreations Research*, ePub ahead of print, August 26, http://www.springerlink.com/content/48m0n0378447w365/, 2010.

[36] C. P. South, M. Partridge, and P. M. Evans. A theoretical framework for prescribing radiotherapy dose distributions using patient-specific biological information. *Medical Physics* 35(10):4599–4611, 2008.

[37] A. Sovik, E. Malinen, H. K. Skogmo, S. M. Bentzen, O. S. Bruland, and D. R. Olsen. Radiotherapy adapted to spatial and temporal variability in tumor hypoxia. *International Journal Radiation Oncology Biology Physics* 68(5):1496–1504, 2007.

[38] N. Stavreva, P. Stavrev, and B. G. Fallone. Probability dynamics of a repopulating tumor in case of fractionated external radiotherapy. *Physica Medica* 25(4):181–191, 2009.

[39] R. D. Stewart and X. A. Li. BGRT: Biologically guided radiation therapy—The future is fast approaching. *Medical Physics* 34(10):3739–3751, 2007.

[40] G. W. Swan. *Optimization of Human Cancer Radiotherapy.* Springer-Verlag, Berlin, 1981.

[41] K. R. Swanson, R. C. Rostomily, and E. C. Alvord Jr. A mathematical modelling tool for predicting survival of individual patients following resection of glioblastoma: A proof of principle. *British Journal of Cancer* 98(1):113–119, 2008.

[42] D. W. Townsend and T. Beyer. A combined PET/CT scanner: The path to true image fusion. *British Journal of Radiology* 75(Suppl):S24–S25, 2002.

[43] D. W. Townsend and S. R. Cherry. Combining anatomy and function: The path to true image fusion. *European Radiology* 11(10):1968–1974, 2001.

[44] F. van der Heijden, R. P. W. Duin, D. de Ridder, and D. M. J. Tax. *Classiffication, Parameter Estimation and State Estimation.* John Wiley & Sons, Chichester, UK, 2004.

[45] S. Webb. *Contemporary IMRT: Developing Physics and Clinical Implementation.* Institute of Physics, Bristol, UK, 2005.

[46] L. M. Wein, J. E. Cohen, and J. T. Wu. Dynamic optimization of a linear-quadratic model with incomplete repair and volume-dependent sensitivity and repopulation. *International Journal of Radiation Oncology Biology Physics* 47(4):1073–1083, 2000.

[47] M. V. Williams, J. Denekamp, and J. F. Fowler. A review of alpha/beta ratios for experimental tumors: Implications for clinical studies of altered fractionation. *International Journal of Radiation Oncology Biology Physics* 11(1):87–96, 1985.

[48] D. Yan. Image guided/adaptive radiotherapy. W. Schleel, T. Borfeld, and A. L. Grosu, eds. *New Technologies in Radiation Oncology.* Springer-Verlag, Heidelberg, Germany, 321–336, 2006.

[49] Y. Yang and L. Xing. Towards biologically conformal radiation therapy (BCRT): Selective IMRT dose escalation under the guidance of spatial biology distribution. *Medical Physics* 32(6):1473–1484, 2005.