INSTITUTE OF PHYSICS PUBLISHING PHYSICS IN MEDICINE AND BIOLOGY

Phys. Med. Biol. **49** (2004) 1991–2013 PII: S0031-9155(04)74385-7

A unifying framework for multi-criteria fluence map optimization models

H Edwin Romeijn[1], **James F Dempsey**[2] **and Jonathan G Li**[2]

[1] Department of Industrial and Systems Engineering, University of Florida, Gainesville, Florida 32611-6595, USA
[2] Department of Radiation Oncology, University of Florida, Gainesville, Florida 32610-0385, USA

E-mail: romeijn@ise.ufl.edu, dempsey@ufl.edu and lijg@ufl.edu

Received 6 January 2004
Published 4 May 2004
Online at stacks.iop.org/PMB/49/1991
DOI: 10.1088/0031-9155/49/10/011

Abstract

Models for finding treatment plans for intensity modulated radiation therapy are usually based on a number of structure-based treatment plan evaluation criteria, which are often conflicting. Rather than formulating a model that *a priori* quantifies the trade-offs between these criteria, we consider a multi-criteria optimization approach that aims at finding the so-called *undominated* treatment plans. We present a unifying framework for studying multi-criteria optimization problems for treatment planning that establishes conditions under which treatment plan evaluation criteria can be transformed into convex criteria while preserving the set of undominated treatment plans. Such transformations are identified for many of the criteria that have been proposed to date, establishing equivalences between these criteria. In addition, it is shown that the use of a nonconvex criterion can often be avoided by transformation to an equivalent convex criterion. In particular, we show that models employing criteria such as tumour control probability, normal tissue complication probability, probability of uncomplicated tumour control, as well as sigmoidal transformations of (generalized) equivalent uniform dose are equivalent to models formulated in terms of separable voxel-based criteria that penalize dose in individual voxels.

1. Introduction

Intensity modulated radiation therapy (IMRT) inverse treatment planning necessitates the formulation of a fluence map optimization (FMO) model to guide the planning process. Such models are generally based on sets of conflicting treatment plan evaluation criteria. For example, one may consider measures of (generalized) equivalent uniform dose (EUD and gEUD), tumour control probability (TCP), normal tissue complication probability (NTCP) and dose- or dose-volume-based criteria. The distinguishing feature of different FMO models then lies in the choice of evaluation criteria that are employed. The typical approach has been to formulate a FMO model that makes an *ad hoc* trade-off by imposing bound constraints on criteria and defining an objective function that optimizes the value of a weighted sum of criteria (see Shepard *et al* (1999) for a review).

An important distinction is the difference between the *importance* of a criterion and its *parameterization*. An example to illustrate this point is the structure-based criterion of generalized equivalent uniform dose as proposed by Kutcher and Burman (1989) and Niemierko (1999). This criterion is parametrized by a power parameter that characterizes the radiation response of the underlying tissue of the structure. Due to the biological meaning that is generally attributed to this parameter, it should be estimated from clinical data, rather than determined in a trial-and-error fashion by repeatedly solving FMO models based on this criterion in an attempt to find clinically valuable treatment plans. Put differently, it must be realized that the relative importance of the gEUD criteria between different structures is not implicit in the values of the corresponding gEUD parameters, but should be measured in another way. For example, one might employ the gEUD of the dose distribution in a spinal cord to evaluate the risk of myelitis while the gEUD of the dose distribution in a saliva gland might be employed to evaluate the risk of xerostomia. In both cases, the gEUD parameter depends on the characteristics of the structure. However, the values of these parameters do not define the relative importance of the risk of myelitis versus the risk of xerostomia, while from a clinical standpoint there is an obvious difference between importance of the two risks. Therefore, it is desirable to incorporate an additional parameter measuring the importance or weight associated with each criterion in a FMO model, which can then be used to control the trade-off between disparate criteria.

Unfortunately, there is generally no scientific or fundamental basis for making a trade-off (i.e., choosing the criteria weights and bounds) between criteria in a given model. A more elegant approach that recognizes this fact is to formulate the FMO model as a multi-criteria optimization problem, as recently proposed by Küfer *et al* (2000), Hamacher and Küfer (2002), Thieke *et al* (2003a, 2003b), Bortfeld *et al* (2003) and Lahanas *et al* (2003). In a multi-criteria approach we are only interested in plans with the property that improving a single criterion value is only possible if at least one other criterion value deteriorates. Treatment plans that possess this property are called *undominated* plans. The multi-criteria approach therefore eliminates all treatment plans for which no trade-off is needed to improve at least one criterion value.

Solving a multi-criteria optimization model now entails characterizing its set of undominated treatment plans. The goal of this paper is to provide a unifying mathematical framework that allows for a scientific comparison of different models via the comparison of the corresponding sets of undominated plans. From a practical point of view, such a comparison can only be made if these sets can be characterized efficiently. As we will see below, the set of undominated plans is convex when only convex treatment plan evaluation criteria are considered. In this case, undominated plans can be identified by solving a family of convex FMO models to global optimality. However, many of the criteria that have been proposed in the literature are nonconvex, leading to nonconvex sets of undominated treatment plans. Our unifying framework establishes conditions under which such criteria can be transformed into convex criteria with an equivalent set of undominated treatment plans. We then apply this framework to identify such transformations for many of the treatment plan evaluation criteria that have been proposed to date.

2. A multi-criteria optimization approach to fluence-map optimization

2.1. Introduction and notation

We consider the general FMO problem for IMRT, where every patient has a set of defined targets and critical structures to consider. The dose received by a voxel is generally assumed to be a linear function of the beamlet intensities (or weights):

$$D_{js}(\vec{x}) = \sum_{i=1}^{n} D_{ijs} x_i \qquad j = 1, \ldots, v_s; \quad s = 1, \ldots, S \tag{1}$$

where D_{ijs} is the dose deposited in voxel j in structure s from beamlet i at unit intensity, and the intensity of beamlet i is denoted by x_i. Furthermore, S is the total number of structures, the first T of which are the targets. Finally, the number of voxels contained in structure s is v_s, and the number of beamlets is n. It will be convenient to denote the dose distribution in a structure as a vector function of the vector of beamlet intensities $\vec{x}$ by

$$\vec{D}_s(\vec{x}) = (D_{1s}(\vec{x}), \ldots, D_{v_s s}(\vec{x})) \qquad s = 1, \ldots, S$$

which can be interpreted as a point in the v_s-dimensional space of all possible dose distributions for structure s.

In the IMRT literature, many alternative types of criteria have been proposed for evaluating treatment plans. Usually, evaluation criteria are formulated for individual structures, and can be expressed as a function of the beamlet intensities $\vec{x}$. To adequately evaluate a treatment plan, we are interested in the values of multiple evaluation criteria (usually at least one per structure). In particular, we assume that we are interested in the values of L criterion functions, denoted by $G_\ell(\vec{x})$ ($\ell = 1, \ldots, L$). Without loss of generality, we also assume that lower values are preferred to higher values for each of these L criteria (if this is not the case for some criteria, we simply multiply them by -1). Such criteria are generally conflicting, and there will usually not exist a treatment plan that achieves the smallest value of all criteria simultaneously. We therefore need to make a trade-off between these criteria. As mentioned above, there is generally no fundamental basis for making these trade-offs. We therefore formulate the FMO model as a multi-criteria optimization problem, and develop a unifying framework for such models. From this vantage point, we then explore the relationship between multi-criteria FMO models based on different evaluation criteria.

2.2. Multi-criteria optimization

Multi-criteria optimization is an approach that has been widely applied in many traditional engineering and business situations where a trade-off needs to be made between different conflicting criteria (for example, cost and quality in engineering design, and cost and customer service in business; see the textbooks by Steuer (1986) and Miettinen (1999) for a discussion of this general technique). The nature of the FMO problem makes it very suitable for a multi-criteria approach as well. Using the notation introduced in section 2.1, we obtain the following multi-criteria FMO model formulation (P):

$$\text{minimize}_{\vec{x} \geqslant 0} \{G_1(\vec{x}), \ldots, G_L(\vec{x})\}.$$

The concept of *Pareto efficiency* (sometimes also called *Pareto optimality*; see Pareto (1971)) can then be used to characterize all meaningful candidate solutions that should be considered in making the trade-off between the conflicting goals. In particular, solutions of the multi-criteria optimization problem (P) with the property that improving a single criterion value is only possible if at least one other criterion value deteriorates are called Pareto efficient.

On the other hand, solutions for which it is possible to improve the value of one criterion without making the value of any other criterion worse are *inefficient* and therefore not worth considering. As an illustration, suppose that we are given a treatment plan, i.e., a set of beamlet intensities $\vec{x}'$, with corresponding values $G_\ell(\vec{x}')$ for all criteria ($\ell = 1, \ldots, L$). If it is possible to find another treatment plan, say $\vec{x}''$, for which none of the criteria has a higher value while it has a lower value for at least one of the criteria, then the treatment plan $\vec{x}''$ is said to dominate $\vec{x}'$, and we may discard $\vec{x}'$ from further consideration. All treatment plans with the property that it is impossible to find another treatment plan that dominates it are called *Pareto efficient* (or undominated, as defined above) treatment plans.

Before we characterize the Pareto efficient plans in (P) we first define the set

$$B = \{\vec{U} : \exists \vec{x} \geqslant 0 \text{ such that } G_\ell(\vec{x}) \leqslant U_\ell \text{ for } \ell = 1, \ldots, L\}$$

to be the set of vectors of upper bounds $\vec{U} = (U_1, \ldots, U_L)$ for which there exists at least one treatment plan where none of the criteria have a value that exceeds the corresponding upper bound. Note that the set B has the following obvious yet important property.

Property 2.1. *If $\vec{U} \in B$, then if $\vec{U}' \geqslant \vec{U}$, it follows that $\vec{U}' \in B$ as well.*

Now consider the boundary, ∂B, of the set B. This is the set of vectors $\vec{U} \in B$ with the property that no vector $\vec{U}' \in B$ exists for which $\vec{U}' \leqslant \vec{U}$ and $\vec{U}' \neq \vec{U}$ (i.e., $U'_\ell < U_\ell$ for at least one $\ell = 1, \ldots, L$). In other words, a vector of upper bounds $\vec{U} \in \partial B$ has the property that no treatment plan $\vec{x}$ can be found if the bound on one criterion is tightened without relaxing the bound on at least one other criterion. Pareto efficient treatment plans $\vec{x}$ are therefore precisely the treatment plans for which the criteria satisfy a vector of upper bounds on the boundary ∂B of the set B. We call the boundary ∂B the *Pareto efficient frontier* of the set B.

In the following section, we will study how the sets B and ∂B behave under increasing transformations of the treatment plan evaluation criteria $G_\ell(\vec{x})$ ($\ell = 1, \ldots, L$).

2.3. Invariance under increasing transformations of criteria

In this section, we will show that transforming any or all of the criteria used in the multi-criteria FMO model via increasing functions leads to an *equivalent* Pareto efficient frontier. This has very significant consequences for multi-criteria FMO modelling, since it means that we may wisely choose transformations in such a way that the sets B and ∂B take on a mathematical form that is highly efficient from an algorithmic point of view.

In particular, suppose that we transform criterion function G_ℓ by an increasing function h_ℓ (for $\ell = 1, \ldots, L$). We could then consider the associated multi-criteria optimization problem (P^h):

$$\text{minimize}_{\vec{x} \geqslant 0} \{h_1(G_1(\vec{x})), \ldots, h_L(G_L(\vec{x}))\}.$$

In the remainder of this section, we will show that the sets of Pareto efficient treatment plans corresponding to the multi-criteria FMO models (P) and (P^h) are equivalent.

Analogous to the approach in section 2.2, we define the set

$$B^h = \left\{\vec{U}^h : \exists \vec{x} \geqslant 0 \text{ such that } h_\ell(G_\ell(\vec{x})) \leqslant U^h_\ell \text{ for } \ell = 1, \ldots, L\right\}$$

to be the set of vectors of upper bounds $\vec{U}^h = \left(U^h_1, \ldots, U^h_L\right)$ for which there exists at least one treatment plan where none of the transformed criteria have a value that exceeds the corresponding upper bound. Furthermore, we denote the set of Pareto efficient points in B^h (i.e., its boundary) by ∂B^h.

Now observe that the inequalities

$$G_\ell(\vec{x}) \leqslant U_\ell$$

and

$$h_\ell(G_\ell(\vec{x})) \leqslant h_\ell(U_\ell)$$

are mathematically equivalent due to the fact that the function h_ℓ is increasing. Therefore, it is easy to see that there is a one-to-one correspondence between the sets B and B^h:

$$B = \left\{\left(h_1^{-1}\left(U_1^h\right), \ldots, h_L^{-1}\left(U_L^h\right)\right) : \vec{U}^h \in B^h\right\}$$

and

$$B^h = \{(h_1(U_1), \ldots, h_L(U_L)) : \vec{U} \in B\}. \tag{2}$$

The following theorem contains the first main result of this paper. It shows that there is also a one-to-one correspondence between the set of Pareto efficient points of the sets B and B^h.

Theorem 2.2. *The sets ∂B and ∂B^h are related through*

$$\partial B = \left\{\left(h_1^{-1}\left(U_1^h\right), \ldots, h_L^{-1}\left(U_L^h\right)\right) : \vec{U}^h \in \partial B^h\right\}$$

or, equivalently,

$$\partial B^h = \{(h_1(U_1), \ldots, h_L(U_L)) : \vec{U} \in \partial B\}.$$

Proof. Suppose $\vec{U} \in \partial B$. We will show that $(h_1(U_1), \ldots, h_L(U_L)) \in \partial B^h$. By equation (2), we have that $(h_1(U_1), \ldots, h_L(U_L)) \in B^h$. Now suppose that

$$(h_1(U_1), \ldots, h_L(U_L)) \notin \partial B^h. \tag{3}$$

Then there exists some $\vec{U}^h \in \partial B^h$ such that

$$U_\ell^h \leqslant h_\ell(U_\ell) \qquad \ell = 1, \ldots, L$$

and strict inequality holds for at least one ℓ. Now consider the transformation of this vector $\left(h_1^{-1}\left(U_1^h\right), \ldots, h_L^{-1}\left(U_L^h\right)\right) \in B$. It is easy to see that

$$h_\ell^{-1}\left(U_\ell^h\right) \leqslant U_\ell \qquad \ell = 1, \ldots, L$$

with strict inequality holding for at least one ℓ. This contradicts the fact that $\vec{U} \in \partial B$, so the assumption in equation (3) is false, and we must instead have that

$$(h_1(U_1), \ldots, h_L(U_L)) \in \partial B^h.$$

A similar argument can be used to show that if $\vec{U}^h \in \partial B^h$, we have that $\left(h_1^{-1}\left(U_1^h\right), \ldots, h_L^{-1}\left(U_L^h\right)\right) \in \partial B$. This completes the proof. □

2.4. Convex criteria

In this section, we will consider the case where all criterion functions G_ℓ are continuous convex functions of the beamlet intensities $\vec{x}$. In that case, deciding whether some vector of upper bounds $\vec{U}$ is in the set B can be formulated as a convex optimization problem. This is attractive since efficient algorithms exist to solve such problems to *global* optimality. The following proposition shows that in this case the set B is convex and closed.

Proposition 2.3. *If the criterion functions G_ℓ ($\ell = 1, \ldots, L$) are convex and continuous, then the set B is convex and closed.*

Proof. Let $\vec{U}'$ and $\vec{U}''$ be elements of B, and let $\lambda \in [0, 1]$. Let $\vec{x}' \geqslant 0$ be such that $G_\ell(\vec{x}') \leqslant U'_\ell$ for all $\ell = 1, \ldots, L$, and $\vec{x}'' \geqslant 0$ be such that $G_\ell(\vec{x}'') \leqslant U''_\ell$ for all $\ell = 1, \ldots, L$. Clearly, $\lambda\vec{x}' + (1-\lambda)\vec{x}'' \geqslant 0$. In addition,

$$\begin{aligned} G_\ell(\lambda\vec{x}' + (1-\lambda)\vec{x}'') &\leqslant \lambda G_\ell(\vec{x}') + (1-\lambda)G_\ell(\vec{x}'') \\ &\leqslant \lambda U'_\ell + (1-\lambda)U''_\ell \end{aligned}$$

for $\ell = 1, \ldots, L$ by the convexity of the functions G_ℓ. Therefore, $\lambda\vec{U}' + (1-\lambda)\vec{U}'' \in B$, and we conclude that B is convex. The fact that B is closed follows directly from the continuity of the functions G_ℓ. □

In the case of convex criteria, we may characterize the Pareto efficient frontier ∂B using a method that is different but equivalent to that used in section 2.2. In particular, we will define a family of optimization problems, the optimal solutions of which yield the Pareto efficient frontier ∂B. Due to the convexity, the closedness and property 2.1, any linear function with non-negative coefficients minimized over the set B yields a point on the boundary (i.e., the efficient frontier) ∂B. Using the definition of the set B, this means that the efficient frontier ∂B can be found by solving the following optimization problem (P($\vec{\mu}$)):

$$\text{minimize} \sum_{\ell=1}^{L} \mu_\ell U_\ell$$

subject to

$$\begin{aligned} &G_\ell(\vec{x}) \leqslant U_\ell && \ell = 1, \ldots, L \\ &\vec{x} \geqslant 0 && \\ &U_\ell \text{ free} && \ell = 1, \ldots, L \end{aligned} \tag{4}$$

for all parameter vectors $\vec{\mu} \geqslant 0$. In other words, for each $\vec{U}' \in \partial B$, there exists a treatment plan $\vec{x}' \geqslant 0$ and a parameter vector $\vec{\mu} \geqslant 0$ such that $(\vec{x}', \vec{U}')$ is an optimal solution to the optimization problem (P($\vec{\mu}$)). The following theorem contains the second main result of this paper by providing a simplification of the formulation of this problem.

Theorem 2.4. *The optimization problem* ($P(\vec{\mu})$) *can be equivalently formulated as*

$$\textit{minimize}_{\vec{x}\geqslant 0} \sum_{\ell=1}^{L} \mu_\ell G_\ell(\vec{x})$$

Proof. We will show that, without loss of optimality, we may replace the inequality constraints (4) in (P($\vec{\mu}$)) by equality constraints. Suppose that this statement is untrue, i.e., suppose that $(\vec{x}^*, \vec{U}^*)$ is an optimal solution to (P($\vec{\mu}$)), and one of the criterion bound constraints (4) is satisfied as a strict inequality: $G_{\hat{\ell}}(\vec{x}^*) < U_{\hat{\ell}}$ for some $\hat{\ell}$. Then the solution $(\vec{x}^*, \vec{U}')$ with

$$\begin{aligned} &U'_{\hat{\ell}} = G_{\hat{\ell}}(\vec{x}^*) \\ &U'_\ell = U^*_\ell \qquad \text{for} \quad \ell = 1, \ldots, L; \ell \neq \hat{\ell} \end{aligned}$$

is a feasible solution to (P($\vec{\mu}$)) as well, and its solution value is no worse than the solution value of $(\vec{x}^*, \vec{U}^*)$ since $\mu_{\hat{\ell}} \geqslant 0$. Therefore, $(\vec{x}^*, \vec{U}')$ is an optimal solution to (P($\vec{\mu}$)) as well. Repeating this argument for all bound constraints that are satisfied as a strict inequality we arrive at the desired result. The equivalent formulation of the optimization problem (P($\vec{\mu}$)) then follows by substituting the criterion functions $G_\ell(\vec{x})$ for the upper bounds U_ℓ. □

Clearly, the problem (P($\vec{\mu}$)) is a convex optimization problem by the convexity of the functions G_ℓ ($\ell = 1, \ldots, L$). In summary, we have shown that if the criterion functions are all convex, we can find all solutions on the efficient frontier ∂B with respect to these criteria by solving convex optimization problems of the form (P($\vec{\mu}$)), where the coefficients $\vec{\mu}$ can be viewed as relative importance factors assigned to the individual criteria.

3. Relationships between treatment plan evaluation criteria

3.1. A unifying framework for treatment plan evaluation criteria

In this section, we will show that most structure-based treatment plan evaluation criteria that have been proposed in the literature can be cast in the form

$$(h \circ G)(\vec{x}) \equiv h(G(\vec{x})) \tag{5}$$

where G is a convex criterion for which small values are preferred to large values and where h is an increasing function. Compositions of this type provide a unifying framework for studying the relationships between various treatment plan evaluation criteria.

If we have specified a set of composite criteria $h_\ell \circ G_\ell$ ($\ell = 1, \ldots, L$) as described above, theorem 2.2 states that the Pareto efficient frontier with respect to the criteria $\{G_\ell\}$ is equivalent to that with respect to the criteria $\{h_\ell \circ G_\ell\}$. *This means that the set of Pareto efficient treatment plans is the same for both choices of criteria.* Moreover, theorem 2.4 states that we can find these Pareto efficient plans by solving convex optimization problems of the form (P($\vec{\mu}$)). Clearly, theorem 2.4 further states that if the criteria $\{h_\ell \circ G_\ell\}$ are convex, we can also find the Pareto efficient plans by solving the following alternative family of convex optimization problems (P^h($\vec{\mu}^h$)):

$$\text{minimize}_{\vec{x} \geqslant 0} \sum_{\ell=1}^{L} \mu_\ell^h h_\ell(G_\ell(\vec{x}))$$

for $\vec{\mu}^h \geqslant 0$. This is relevant if the optimization problems (P^h($\vec{\mu}^h$)) happen to be more efficiently solvable than (P($\vec{\mu}$)). Perhaps more importantly, if the criteria $\{h_\ell \circ G_\ell\}$ are *not* convex, theorem 2.4 then states that the Pareto efficient treatment plans with respect to these criteria can still be found by solving the optimization problems (P($\vec{\mu}$)), which are formulated with respect to the convex criteria $\{G_\ell\}$. *This allows one to replace a nonconvex criterion by an equivalent convex one if a suitable decomposition into a convex criterion G and an increasing function h can be found.*

3.2. Criteria as functions of dose versus beamlet intensities

In the literature, structure-based treatment plan evaluation criteria have usually been cast as a function of the dose distribution in a structure, rather than as a function of the beamlet intensities $\vec{x}$. With a slight abuse of notation, we will use the same symbol, G, for both. That is, if $\vec{d}$ denotes the dose distribution in a given structure, we will use both $G(\vec{x})$ and $G(\vec{d})$ to evaluate a treatment plan. Although the functions $G(\vec{x})$ and $G(\vec{d})$ are formally speaking different, the interpretation of the argument will be clear from the context. Moreover, the two representations are related through

$$G(\vec{x}) = G(\vec{D}(\vec{x}))$$

using the dose model in equation (1) and suppressing, as we will do whenever appropriate in the remainder of this section, the subscript s indicating the particular structure to which the criterion is applied.

In order to study treatment plan evaluation criteria and cast them in terms of our framework (5), it is important to be able to determine whether a given criterion is a convex function of the beamlet intensities. As the following lemma shows, convexity of an evaluation criterion in the dose distribution implies that it is convex in the beamlet intensities as well under the linear dose model (1), so that we may limit ourselves to studying the convexity of treatment plan evaluation criteria as a function of the dose distribution.

Lemma 3.1. *If $G(\vec{d})$ is a convex function of $\vec{d}$, then $G(\vec{x}) = G(\vec{D}(\vec{x}))$ is a convex function of $\vec{x}$.*

Proof. Let $\vec{x}', \vec{x}'' \geqslant 0$ be two vectors of beamlet intensities, and let $\lambda \in [0, 1]$. Since $\vec{D}(\vec{x})$ is a linear function of $\vec{x}$, we have

$$\begin{aligned} G(\lambda\vec{x}' + (1-\lambda)\vec{x}'') &= G(\vec{D}(\lambda\vec{x}' + (1-\lambda)\vec{x}'')) \\ &= G(\lambda\vec{D}(\vec{x}') + (1-\lambda)\vec{D}(\vec{x}'')) \\ &\leqslant \lambda G(\vec{D}(\vec{x}')) + (1-\lambda)G(\vec{D}(\vec{x}'')) \\ &= \lambda G(\vec{x}') + (1-\lambda)G(\vec{x}'') \end{aligned}$$

which proves that $G(\vec{x})$ is a convex function of $\vec{x}$. □

In the following subsection, we will analyse many proposed treatment plan evaluation criteria from the literature and study their relationships via our unifying framework (5).

3.3. Analysis of commonly used criteria

3.3.1. Tumour control probability. The response of a target to irradiation can be characterized by the probability that no clonogen cells remain in the target, often called the *tumour control probability* (see, e.g., Withers and McBride (1998)). Based on a Poisson approximation of a binomial model for the number of remaining clonogens and including cell repopulation effects, the TCP of a target can be expressed as

$$\prod_{j=1}^{v} \exp\left(-\frac{N\,\mathrm{e}^{\lambda(n-1)\Delta T}}{v} \cdot p(d_j)\right) \tag{6}$$

where N is the total number of clonogen cells in the target, $p(d_j)$ is the surviving fraction of clonogen cells in voxel j of the target after receiving dose d_j, n is the number of fractions used for the treatment, ΔT is the time between two consecutive fractions and λ is the net rate of cell birth (see Stavreva *et al* (2003) and Zaider and Minerbo (2000)). The recent work by Stavreva *et al* (2003) shows that a single-hit model with repopulation provides the best empirical fit to tumour response data. In this case, the TCP for a target becomes

$$\mathrm{TCP}(\vec{d}; N, \alpha, \lambda, n, \Delta T) = \prod_{j=1}^{v} \exp\left(-\frac{N\,\mathrm{e}^{\lambda(n-1)\Delta T}}{v} \cdot \mathrm{e}^{-\alpha d_j}\right) \tag{7}$$

where $1/\alpha$ is the mean lethal dose to the target, and describes the radioresistance of cells in that target. The single-hit model without repopulation (see, e.g., Brahme and Agren (1987) and Goitein (1987)) is obtained as a special case by substituting $\lambda = 0$:

$$\mathrm{TCP}(\vec{d}; N, \alpha) = \prod_{j=1}^{v} \exp\left(-\frac{N}{v} \cdot \mathrm{e}^{-\alpha d_j}\right).$$

3.3.2. Equivalent uniform dose. Niemierko (1997) introduced the concept of *equivalent uniform dose* for a target by equating the TCP for an inhomogeneous dose distribution to that for a homogeneous dose distribution and solving for the corresponding homogeneous dose:

$$\mathrm{TCP}(\vec{d}) = \mathrm{TCP}(\mathrm{EUD}\,\vec{\iota})$$

where $\vec{\iota}$ is a vector in which all elements are equal to 1. Applying this approach to the model for the TCP of a target given in equation (7), we obtain Niemierko's original definition of EUD by setting

$$\begin{aligned}\mathrm{TCP}(\vec{d}; N, \alpha, \lambda, n, \Delta T) &= \prod_{j=1}^{v} \exp\left(-\frac{N\,\mathrm{e}^{\lambda(n-1)\Delta T}}{v} \cdot \mathrm{e}^{-\alpha d_j}\right)\\ &= \exp\left(-N\,\mathrm{e}^{\lambda(n-1)\Delta T} \cdot \frac{1}{v}\sum_{j=1}^{v} \mathrm{e}^{-\alpha d_j}\right)\end{aligned}$$

equal to

$$\mathrm{TCP}(\mathrm{EUD}\,\vec{\iota}; N, \alpha, \lambda, n, \Delta T) = \exp(-N\,\mathrm{e}^{\lambda(n-1)\Delta T} \cdot \mathrm{e}^{-\alpha \mathrm{EUD}(\vec{d})})$$

yielding

$$\mathrm{EUD}(\vec{d}; \alpha) \equiv -\frac{1}{\alpha} \ln\left(\frac{1}{v}\sum_{j=1}^{v} \mathrm{e}^{-\alpha d_j}\right)$$

(see also McGary *et al* (2000)). Note that this measure of EUD is independent of the repopulation rate λ, the number of fractions n, and the time between fractions ΔT.

Subsequently, Niemierko (1999) (using a discrete version of a model earlier proposed by Kutcher and Burman (1989)), defined the concept of *generalized EUD* for a target as the generalized mean of the dose distribution in that target, i.e.,

$$\mathrm{gEUD}(\vec{d}; a) = \left(\frac{1}{v}\sum_{j=1}^{v} d_j^{a}\right)^{1/a}$$

(see also Abramowitz and Stegun (1965)), where $-\infty \leqslant a \leqslant 0$ is a structure-dependent parameter that depends on the radiation response of the underlying tissue. For parameter values $1 \leqslant a \leqslant \infty$, this definition extends the concept of gEUD to critical structures. Note that for the limiting cases $a = -\infty$, 0 and ∞, the gEUD becomes the minimum, geometric mean and maximum dose in the structure, respectively. Since its introduction, it has been used by many researchers to formulate apparently different models for FMO; for examples please see Choi and Deasy (2002), Wu *et al* (2002), Thieke *et al* (2003c), and Bortfeld *et al* (2003). In the following section, we will study one of these models in greater detail.

3.3.3. Sigmoidal criteria based on gEUD. Wu *et al* (2002) have proposed a model for FMO that summarizes the dose distribution in each structure by a function of the gEUD. In particular, their model for FMO involves the maximization of the following objective function:

$$\prod_{s=1}^{S} w_s\big(\vec{d}_s; a_s, k_s, \mathrm{gEUD}_s^0\big)$$

or, equivalently, the minimization of

$$-\sum_{s=1}^{S} \ln w_s\big(\vec{d}_s; a_s, k_s, \mathrm{gEUD}_s^0\big)$$

where a_s is the gEUD parameter, gEUD_s^0 is a base gEUD value, and k_s is a positive parameter ($s = 1, \ldots, S$). The functions w_s are chosen to be logistic functions, which have a sigmoidal shape. In particular, for a generic structure

$$w(\vec{d}; a, k, \text{gEUD}^0) = \begin{cases} \dfrac{1}{1 + \left(\frac{\text{gEUD}^0}{\text{gEUD}(\vec{d};a)}\right)^k} & \text{if} \quad -\infty \leqslant a \leqslant 0 \\ \dfrac{1}{1 + \left(\frac{\text{gEUD}(\vec{d};a)}{\text{gEUD}^0}\right)^k} & \text{if} \quad 1 \leqslant a \leqslant \infty. \end{cases}$$

Note that gEUD^0 is in fact the inflection point of the sigmoidal function, and k determines the steepness of the sigmoidal function at the inflection point. In a multi-criteria setting, this suggests the use of the following treatment plan evaluation criteria for structures:

$$\begin{aligned} W(\vec{d}; a, k, \text{gEUD}^0) &= -\ln w(\vec{d}; a, k, \text{gEUD}^0) \\ &= \begin{cases} \ln\left(1 + \left(\dfrac{\text{gEUD}^0}{\text{gEUD}(\vec{d}; a)}\right)^k\right) & \text{if} \quad -\infty \leqslant a \leqslant 0 \\ \ln\left(1 + \left(\dfrac{\text{gEUD}(\vec{d}; a)}{\text{gEUD}^0}\right)^k\right) & \text{if} \quad 1 \leqslant a \leqslant \infty. \end{cases} \end{aligned}$$

As Wu *et al* (2002) mention, they could have chosen another sigmoidal function of gEUD than the logistic one to define their criteria. We will therefore also study what the consequences are of employing Φ, the cumulative distribution function of the standard normal distribution (i.e., $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \mathrm{e}^{-t^2/2}\, \mathrm{d}t$). In that case, we obtain the criteria

$$\begin{aligned} \breve{W}(\vec{d}; a, \sigma, \text{gEUD}^0) &= -\ln \breve{w}(\vec{d}; a, \sigma, \text{gEUD}^0) \\ &= \begin{cases} -\ln\left(1 - \Phi\left(\dfrac{\text{gEUD}^0 - \text{gEUD}(\vec{d}; a)}{\sigma \cdot \text{gEUD}^0}\right)\right) & \text{if} \quad -\infty \leqslant a \leqslant 0 \\ -\ln\left(1 - \Phi\left(\dfrac{\text{gEUD}(\vec{d}; a) - \text{gEUD}^0}{\sigma \cdot \text{gEUD}^0}\right)\right) & \text{if} \quad 1 \leqslant a \leqslant \infty. \end{cases} \end{aligned}$$

Again, gEUD^0 is the inflection point of the sigmoidal function; however, in this case the parameter $\sigma > 0$ determines the steepness of the sigmoidal function at the inflection point.

3.3.4. Normal tissue complication probability. Several models have been proposed in the literature to describe the *normal tissue complication probability* associated with the dose distribution in a critical structure (see, e.g., Wolbarst (1984), Lyman (1985), Lyman and Wolbarst (1987a, 1987b), Schultheiss *et al* (1983), Niemierko and Goitein (1991)). In this section, we will discuss the most recent work by Stavrev *et al* (2003) and Alber and Nüsslin (2001).

Stavrev *et al* (2003) base the NTCP for inhomogeneous dose distributions on the following model that Lyman (1985) proposed for determining the NTCP for a homogeneous dose distribution in a critical structure:

$$\text{NTCP}(D\vec{\iota}; m, D_{50}) = \Phi\left(\frac{D - D_{50}}{m D_{50}}\right) \tag{8}$$

where D is the homogeneous dose, D_{50} denotes the uniform dose where the structure exhibits a 50% complication probability, and the parameter m determines the mid-point slope of the NTCP curve, sometimes also called γ_{50}. In particular, they determine the NTCP for inhomogeneous dose distributions by evaluating equation (8) at the gEUD for the inhomogeneous dose distribution:

$$\mathrm{NTCP}^{\mathrm{L\&S}}(\vec{d}; a, m, D_{50}) = \mathrm{NTCP}(\mathrm{gEUD}(\vec{d}; a)\vec{\iota}; m, D_{50}) = \Phi\left(\frac{\mathrm{gEUD}(\vec{d}; a) - D_{50}}{m D_{50}}\right)$$

where $1 \leqslant a \leqslant \infty$ is the gEUD parameter associated with the critical structure.

Alber and Nüsslin (2001) have proposed a phenomenological description of an NTCP function derived from mechanistic concepts. They arrive at the expression

$$\mathrm{NTCP}^{\mathrm{A\&N}}(\vec{d}; a, \Delta) = 1 - \exp\left(-\left(\frac{\mathrm{gEUD}(\vec{d}; a)}{\Delta}\right)^{a}\right)$$

where again $1 \leqslant a < \infty$ is the gEUD parameter associated with the critical structure and Δ is a structure-dependent constant denoting the uniform dose where the structure exhibits a $1 - \mathrm{e}^{-1} \approx 63\%$ complication probability.

3.3.5. Separable convex voxel-based criteria. The treatment plan evaluation criteria described thus far are all based on attempts at modelling the biological effect of irradiating the tissues underlying the contoured structures in the FMO problem. However, an approach that has been more widely used to date is one that heuristically measures treatment plan quality by evaluating the dose received by each voxel through a convex function. By considering the average measure received by all voxels in a structure, we can express such models in the context of this paper by defining a treatment plan evaluation criterion for a structure that is *separable* in the doses to individual voxels

$$F(\vec{d}) = \frac{1}{v}\sum_{j=1}^{v} f_j(d_j)$$

where f_j is a convex function. The most common choices found in the literature are of the form

$$f_j(d_j) = c_j |d_j - \delta_j|^{a_j}$$

for targets, and

$$f_j(d_j) = c_j \max\{0, d_j - \delta_j\}^{a_j}$$

for critical structures. Here, δ_j is a dose threshold and a_j and c_j are shape and scale parameters for the convex function corresponding to voxel j in the structure. The shape parameter a_j is usually chosen to be 1 or 2 (where $a_j = 2$ corresponds to the often used least squares model for FMO), and the parameters δ_j and c_j usually depend only on the structure (see Shepard *et al* (1999) for a review of such models). Recently, Tsien *et al* (2003) and Romeijn *et al* (2003) have obtained excellent results by using high powers of dose difference or a piecewise-linear approximation thereof, respectively. Other choices that will prove useful in the remainder of this section are obtained when choosing

$$f_j(d_j; \alpha) = \mathrm{e}^{-\alpha d_j} \qquad \text{for} \quad \alpha > 0$$

or

$$f_j^{\mathrm{g}}(d_j; a) = \begin{cases} d_j^a & \text{for} \quad a \neq 0 \\ -\ln d_j & \text{for} \quad a = 0. \end{cases}$$

We will refer to the corresponding structure-based treatment plan evaluation functions as $F(\vec{d}; \alpha)$ and $F^{\mathrm{g}}(\vec{d}; a)$, respectively.

Table 1. Evaluation criteria for targets.

		Criterion decomposition	
Criterion	*Parameter range*	$G(\vec{d})$	$h(z)$
$\mathrm{TCP}(N, \alpha, \lambda, n, \Delta T)$	$N, \alpha > 0; n \geqslant 1; \lambda, \Delta T \geqslant 0$	$-\ln \mathrm{TCP}(\vec{d}; N, \alpha, n, \Delta T, \lambda)$	$-\mathrm{e}^{-z}$
		$-\mathrm{EUD}(\vec{d}; \alpha)$	$-\exp(-N\mathrm{e}^{\lambda(n-1)\Delta T+\alpha z})$
$\mathrm{EUD}(\alpha)$	$\alpha > 0$	$-\mathrm{EUD}(\vec{d}; \alpha)$	z
		$F(\vec{d}; \alpha)$	$\frac{1}{\alpha} \ln z$
$\mathrm{gEUD}(a)$	$-\infty \leqslant a \leqslant 0$	$-\mathrm{gEUD}(\vec{d}; a)$	z
	$-\infty < a < 0$	$F^{\mathrm{g}}(\vec{d}; a)$	$-z^{1/a}$
	$a = 0$	$F^{\mathrm{g}}(\vec{d}; 0)$	$-\mathrm{e}^{-z}$
$W(a, k, \mathrm{gEUD}^0)$	$-\infty \leqslant a \leqslant 0; k, \mathrm{gEUD}^0 > 0$	$W(\vec{d}; a, k, \mathrm{gEUD}^0)$	z
		$-\mathrm{gEUD}(\vec{d}; a)$	$\ln\left(1 + \left(-\frac{\mathrm{gEUD}^0}{z}\right)^k\right)$
$\breve{W}(a, \sigma, \mathrm{gEUD}^0)$	$-\infty \leqslant a \leqslant 0; \sigma, \mathrm{gEUD}^0 > 0$	$\breve{W}(\vec{d}; a, \sigma, \mathrm{gEUD}^0)$	z
		$-\mathrm{gEUD}(\vec{d}; a)$	$-\ln\left(1 - \Phi\left(\frac{\mathrm{gEUD}^0+z}{\sigma \cdot \mathrm{gEUD}^0}\right)\right)$
$F(\alpha)$	$\alpha > 0$	$F(\vec{d}; \alpha)$	z
	$\alpha > 0$	$-\mathrm{EUD}(\vec{d}; \alpha)$	e^z
$F^{\mathrm{g}}(a)$	$-\infty < a \leqslant 0$	$F^{\mathrm{g}}(\vec{d}; a)$	z
	$-\infty < a < 0$	$-\mathrm{gEUD}(\vec{d}; a)$	$(-z)^a$

Table 2. Evaluation criteria for critical structures.

Criterion	*Parameter range*	$G(\vec{d})$	$h(z)$
$\mathrm{gEUD}(a)$	$1 \leqslant a \leqslant \infty$	$\mathrm{gEUD}(\vec{d}; a)$	z
	$1 \leqslant a < \infty$	$F^{\mathrm{g}}(\vec{d}; a)$	$z^{1/a}$
$W(a, k, \mathrm{gEUD}^0)$	$1 \leqslant a \leqslant \infty; k, \mathrm{gEUD}^0 > 0$	$\mathrm{gEUD}(\vec{d}; a)$	$\ln\left(1 + \left(\frac{z}{\mathrm{gEUD}^0}\right)^k\right)$
$\breve{W}(a, \sigma, \mathrm{gEUD}^0)$	$1 \leqslant a \leqslant \infty; \sigma, \mathrm{gEUD}^0 > 0$	$\mathrm{gEUD}(\vec{d}; a)$	$-\ln\left(1 - \Phi\left(\frac{z-\mathrm{gEUD}^0}{\sigma \cdot \mathrm{gEUD}^0}\right)\right)$
$\mathrm{NTCP}^{\mathrm{L\&S}}(a, m, D_{50})$	$1 \leqslant a \leqslant \infty; m, D_{50} > 0$	$\mathrm{gEUD}(\vec{d}; a)$	$\Phi\left(\frac{z-D_{50}}{mD_{50}}\right)$
$\mathrm{NTCP}^{\mathrm{A\&N}}(a, \Delta)$	$1 \leqslant a < \infty, \Delta > 0$	$\mathrm{gEUD}(\vec{d}; a)$	$1 - \mathrm{e}^{-(z/\Delta)^a}$
$F^{\mathrm{g}}(a)$	$1 \leqslant a \leqslant \infty$	$F^{\mathrm{g}}(\vec{d}; a)$	z
	$1 \leqslant a < \infty$	$\mathrm{gEUD}(\vec{d}; a)$	z^a

3.4. *Application of the unifying framework*

In this section, we apply the unifying framework introduced in section 3.1 to study the relationship between the criteria discussed in section 3.3. To this end, we provide one or more decompositions of each criterion into a convex criterion G and an increasing function h, as required by the framework. Using these decompositions, we can identify the equivalence of Pareto frontiers obtained with different multi-criteria models for FMO.

Some possible decompositions for each of the target criteria are shown in table 1, while table 2 shows decompositions for each of the critical structure criteria. Note that, in section 2 of this paper, we have used the convention that, for all treatment plan evaluation criteria G, smaller values are preferred to larger values. However, for the TCP, EUD and gEUD target criteria in table 1, the opposite is true. Therefore, in each of these cases, we have in fact decomposed the negative of the corresponding criterion.

To verify the validity of the decompositions, we need to establish that the criteria G in the tables are all convex, and that the functions h are all increasing. With respect to the criteria G, we note the following:

- Choi and Deasy (2002) have shown that $\ln \mathrm{TCP}(\vec{d};\, N, \alpha)$ is a concave function of $\vec{d}$ when $N, \alpha > 0$; this immediately implies that
$$\ln \mathrm{TCP}(\vec{d};\, N, \alpha, n, \Delta T, \lambda) = \mathrm{e}^{\lambda(n-1)\Delta T} \cdot \ln \mathrm{TCP}(\vec{d};\, N, \alpha)$$
is a concave function of $\vec{d}$ when $N, \alpha > 0, n \geqslant 1, \Delta T, \lambda \geqslant 0$ as well;
- in appendix A we show that $\mathrm{EUD}(\vec{d};\, \alpha)$ is a concave function of $\vec{d}$ when $\alpha > 0$;
- Choi and Deasy (2002) have shown that $\mathrm{gEUD}(\vec{d};\, a)$ is a convex function of $\vec{d}$ when $1 \leqslant a \leqslant \infty$, and a concave function of $\vec{d}$ when $-\infty \leqslant a \leqslant 0$;
- in appendix B we show that $W(\vec{d};\, a, k, \mathrm{gEUD}^0)$ is a convex function of $\vec{d}$ when $-\infty \leqslant a \leqslant 0$ and $k, \mathrm{gEUD}^0 > 0$.
- in appendix B we also show that $\breve{W}(\vec{d};\, a, \sigma, \mathrm{gEUD}^0)$ is a convex function of $\vec{d}$ when $-\infty \leqslant a \leqslant 0$ or $1 \leqslant a \leqslant \infty$ and $\sigma, \mathrm{gEUD}^0 > 0$.

Recalling that multiplying a concave function by -1 yields a convex function, the sum of convex functions is convex, the functions $\mathrm{e}^{-\alpha z}$ $(\alpha > 0)$ and z^a $(a < 0$ or $a \geqslant 1$ and $z \geqslant 0)$ are convex, and the function $\ln z$ is concave, the convexity of all criteria G in tables 1 and 2 now follows. The fact that all functions h in these tables are increasing is easily verified in most cases; for the functions h used in the decomposition of criteria W and $\breve{W}$ this can be found in appendix B.

The results of section 2 now show that the efficient frontier with respect to *any* combination of criteria listed in the first column of tables 1 and 2 is *equivalent* to the efficient frontier that is obtained when one or more of these are replaced by a corresponding criterion G from the third column of these tables. This important observation can be summarized as follows:

Observation 3.2. *For any combination of criteria listed in the first column of tables 1 and 2, with the exception of the minimum and maximum dose criteria, there exists a set of criteria that are separable in the voxels that yield an equivalent efficient frontier.*

A previous indication that there exists a close relationship between voxel-based penalty functions and biological treatment plan evaluation criteria appeared in the paper by Alber and Nüsslin (1999), who showed how relaxing TCP and NTCP constraints lead to voxel-based penalty formulations of the FMO problem. In the following section, we will illustrate observation 3.2 by applying the unifying framework to the model by Wu *et al* (2002).

3.5. Example of applying the unifying framework

Wu *et al* (2002) have proposed the following model for FMO:

$$\text{minimize}_{\vec{x} \geqslant 0} \sum_{s=1}^{T} W_s\big(\vec{x};\, \tilde{a}_s, \tilde{k}_s, \mathrm{g\tilde{E}UD}_s^0\big) + \sum_{s=1}^{S} W_s\big(\vec{x};\, a_s, k_s, \mathrm{gEUD}_s^0\big)$$

where, in addition to the target criteria, they have also applied critical structure criteria to the targets to achieve homogeneity of the target dose distributions. In particular, for target s $(s = 1, \ldots, T)$, the parameters a_s, k_s and gEUD_s^0 apply to the 'critical structure nature' of the target, while the parameters $\tilde{a}_s$, $\tilde{k}_s$ and $\mathrm{g\tilde{E}UD}_s^0$ apply to the 'target nature' of the target. Note also that we have returned to expressing the criteria as functions of the beamlet intensities $\vec{x}$ rather than the dose distributions $\vec{d}_s$ (see also section 3.2).

The underlying goal of this model is to simultaneously minimize the values of all criteria W_s. However, there appears to be no fundamental basis for the particular trade-off between these conflicting criteria that they have chosen. We may therefore consider the following more general multi-criteria optimization model, which we will call (P^{W}):

$$\begin{aligned}\text{minimize}_{\vec{x}\geqslant 0}\ \big\{&W_1\big(\vec{x};\tilde{a}_1,\tilde{k}_1,\mathrm{g\tilde{E}UD}_1^0\big),\ldots,W_T\big(\vec{x};\tilde{a}_T,\tilde{k}_T,\mathrm{g\tilde{E}UD}_T^0\big),\\&W_1\big(\vec{x};a_1,k_1,\mathrm{gEUD}_1^0\big),\ldots,W_S\big(\vec{x};a_S,k_S,\mathrm{gEUD}_S^0\big)\big\}.\end{aligned}$$

The main disadvantage of this model is the fact that the criterion functions are not all convex, as shown in appendix B. This has two severe consequences. Firstly, this implies that, although the globally optimal solutions to problems ($\mathrm{P}^{\mathrm{W}}(\vec{\mu}^{\mathrm{W}},\vec{\tilde{\mu}}^{\mathrm{W}})$) of the form:

$$\text{minimize}_{\vec{x}\geqslant 0}\ \sum_{s=1}^{T}\tilde{\mu}_s^{\mathrm{W}}W_s\big(\vec{x};\tilde{a}_s,\tilde{k}_s,\mathrm{g\tilde{E}UD}_s^0\big)+\sum_{s=1}^{S}\mu_s^{\mathrm{W}}W_s\big(\vec{x};a_s,k_s,\mathrm{gEUD}_s^0\big)$$

with $\vec{\mu}^{\mathrm{W}},\vec{\tilde{\mu}}^{\mathrm{W}}\geqslant 0$ are Pareto efficient, they do not characterize the entire Pareto efficient frontier of the model. Secondly, the latter optimization problems are nonconvex, so that efficient methods do not exist for solving these problems to global optimality. Moreover, locally optimal solutions (which can be found relatively easily using local search algorithms) are not necessarily Pareto efficient. (Interestingly, note that if Wu *et al* (2002) had chosen to use the criteria $\check{W}$ instead of W, their objective function would have been convex and therefore unimodal! Moreover, convex optimization problems of the form ($\mathrm{P}^{\check{\mathrm{W}}}(\vec{\mu}^{\check{\mathrm{W}}},\vec{\tilde{\mu}}^{\check{\mathrm{W}}})$) could have been used to characterize the entire Pareto efficient frontier of the corresponding model ($\mathrm{P}^{\check{\mathrm{W}}}$).) However, if we now apply our unifying framework to (P^{W}) (or ($\mathrm{P}^{\check{\mathrm{W}}}$)), we can use tables 1 and 2 to conclude that the Pareto efficient frontier of the following multi-criteria model ($\mathrm{P}^{\mathrm{gEUD}}$)

$$\text{minimize}_{\vec{x}\geqslant 0}\ \{-\mathrm{gEUD}_1(\vec{x};\tilde{a}_1),\ldots,-\mathrm{gEUD}_T(\vec{x};\tilde{a}_T),\mathrm{gEUD}_1(\vec{x};a_1),\ldots,\mathrm{gEUD}_S(\vec{x};a_S)\}$$

is equivalent to that of (P^{W}). Since the criteria used in ($\mathrm{P}^{\mathrm{gEUD}}$) (which is in fact the model proposed by Bortfeld *et al* (2003)) are convex, not only do we know that the Pareto efficient frontier is characterized entirely by the globally optimal solutions to problems of the form

$$\text{minimize}_{\vec{x}\geqslant 0}\ -\sum_{s=1}^{T}\tilde{\mu}_s^{\mathrm{gEUD}}\mathrm{gEUD}_s(\vec{x};\tilde{a}_s)+\sum_{s=1}^{S}\mu_s^{\mathrm{gEUD}}\mathrm{gEUD}_s(\vec{x};a_s)$$

with $\vec{\mu}^{\mathrm{gEUD}},\vec{\tilde{\mu}}^{\mathrm{gEUD}}\geqslant 0$, but, in addition, the globally optimal solutions to these optimization problems can be found efficiently due to their convexity. It is interesting to note that the parameters $k_s,\tilde{k}_s,\mathrm{EUD}_s^0$ and $\mathrm{E\tilde{U}D}_s^0$ of (P^{W}) do not appear in the equivalent model ($\mathrm{P}^{\mathrm{gEUD}}$). Clearly, this means that these parameters do not influence the Pareto efficient frontier and are therefore superfluous in a multi-criteria setting. In terms of our unifying framework, any parametrized increasing function h could be applied to the gEUD criteria (for example, the function leading to the criteria $\check{W}$). However, the additional complexity introduced by the corresponding parameters does not change the essence of the model since the set of Pareto efficient solutions is invariant under such transformations.

If all gEUD parameters a_s and $\tilde{a}_s$ are finite, we can now take the application of our unifying framework one step further, and use tables 1 and 2 to apply another set of transformations to arrive at the equivalent multi-criteria optimization model (P^{F}):

$$\text{minimize}_{\vec{x}\geqslant 0}\ \{F_1(\vec{x};\tilde{a}_1),\ldots,F_T(\vec{x};\tilde{a}_T),F_1(\vec{x};a_1),\ldots,F_S(\vec{x};a_S)\}.$$

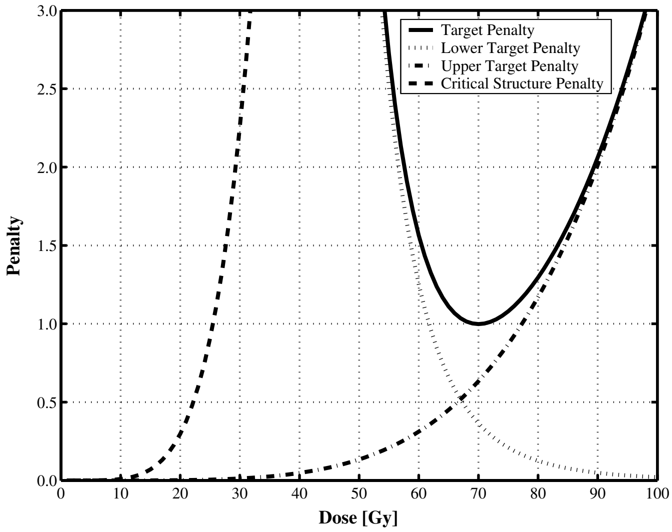

Figure 1. Voxel-based penalty functions for a target and a critical structure.

This result demonstrates that the equivalent Pareto efficient frontiers of ($\mathrm{P^W}$), ($\mathrm{P^{gEUD}}$) and ($\mathrm{P^F}$) can be found by solving optimization problems of the form

$$\text{minimize}_{\vec{x}\geqslant 0} \sum_{s=1}^{T} \tilde{\mu}_s^{\mathrm{F}} F_s(\vec{x}; \tilde{a}_s) + \sum_{s=1}^{S} \mu_s^{\mathrm{F}} F_s(\vec{x}; a_s)$$

with $\vec{\mu}^{\mathrm{F}}, \vec{\tilde{\mu}}^{\mathrm{F}} \geqslant 0$. Note that the objective function of this model can be rewritten as

$$\text{minimize}_{\vec{x}\geqslant 0} \sum_{s=1}^{T} \left(\tilde{\mu}_s^{\mathrm{F}} F_s(\vec{x}; \tilde{a}_s) + \mu_s^{\mathrm{F}} F_s(\vec{x}; a_s)\right) + \sum_{s=T+1}^{S} \mu_s^{\mathrm{F}} F_s(\vec{x}; a_s)$$

or even

$$\text{minimize}_{\vec{x}\geqslant 0} \sum_{s=1}^{T} \sum_{j=1}^{v_s} \left(\tilde{\mu}_s^{\mathrm{F}} f_{js}(\vec{x}; \tilde{a}_s) + \mu_s^{\mathrm{F}} f_{js}(\vec{x}; a_s)\right) + \sum_{s=T+1}^{S} \sum_{j=1}^{v_s} \mu_s^{\mathrm{F}} f_{js}(\vec{x}; a_s).$$

This is a voxel-based penalty function formulation for FMO where all voxels in a given structure have identical penalty functions. The advantage of this formulation over a model with structure-based criteria is the separability of the objective function in the individual voxel-doses, which is often computationally more efficient.

Figure 1 shows an example of voxel-based penalty functions for a target and a critical structure that lead to a multi-criteria optimization model ($\mathrm{P^F}$) that is equivalent to corresponding models ($\mathrm{P^W}$) and ($\mathrm{P^{gEUD}}$). In particular, we follow the gEUD parameter choices made by Wu *et al* (2002) and show a voxel-based penalty function for a head-and-neck target with parameters $\tilde{a} = -8$ and $a = 4.6$

$$f_{js}(d) = \tilde{\mu}_s^{\mathrm{F}} d^{-8} + \mu_s^{\mathrm{F}} d^{4.6}$$

as well as a voxel-based penalty function for a parotid with parameter $a = 5$

$$f_{js}(d) = \mu_s^{\mathrm{F}} d^5$$

where the coefficients $\tilde{\mu}_s^{\mathrm{F}}$ and μ_s^{F} were chosen for illustrative purposes.

Note that if some of the gEUD parameters are infinite, which means that the minimum or maximum dose in certain structures are in the model ($\mathrm{P}^{\mathrm{gEUD}}$), we can incorporate these in the model (P^{F}) without transforming them. In the corresponding optimization problem ($\mathrm{P}^{\mathrm{F}}(\vec{\underline{\mu}}^{\mathrm{F}}, \vec{\overline{\mu}}^{\mathrm{F}})$), these can be handled very efficiently by introducing auxiliary decision variables and constraints (see, e.g., Winston (2004)).

3.6. Combining criteria

Several other treatment plan evaluation criteria that have been proposed in the literature can be viewed as linear combinations of the elementary criteria that we have discussed in section 3.3. When we compare multi-criteria models for FMO based on linear combinations of elementary convex criteria, it is easy to see that these models can be obtained from multi-criteria models based on the individual elementary criteria by restricting the range of values that is considered for the criterion weights $\vec{\mu}$. Models based on such combined criteria thus impose an *a priori* trade-off between the corresponding elementary criteria, thereby significantly reducing the set of efficient treatment plans. In the absence of a fundamental basis for the quantification of the trade-off between the elementary criteria, it is preferable to consider the elementary criteria explicitly rather than the combined criteria. In the remainder of this section, we will address this issue more explicitly by studying combined criteria that have been proposed in the literature.

3.6.1. Alternative measure of EUD. For critical structures, Dale and Olsen (1997) suggested the use of a linear combination of the maximum and the mean dose as an evaluation criterion of a dose distribution:

$$\begin{aligned}\mathrm{EUD}^{\mathrm{D\&O}}(\vec{d};\alpha) &= \alpha\cdot\mathrm{gEUD}(\vec{d};\infty) + (1-\alpha)\cdot\mathrm{gEUD}(\vec{d};1)\\ &= \alpha\cdot\max_{j=1,\ldots,v} d_j + (1-\alpha)\cdot\frac{1}{v}\sum_{j=1}^{v} d_j\end{aligned}$$

where $\alpha \in [0, 1]$ depends on the critical structure. Thieke *et al* (2002) have further developed this approach and estimated values of the parameter α corresponding to several critical structures based on experimental data. However, due to the difficulties in accurately estimating the parameter α and the lack of an underlying theoretical motivation for using $\mathrm{EUD}^{\mathrm{D\&O}}$, it will be preferable to simply incorporate the maximum dose $\mathrm{gEUD}(\vec{d};\infty)$ and the mean dose $\mathrm{gEUD}(\vec{d};1)$ explicitly in a multi-criteria model for FMO without imposing an *a priori* trade-off between them.

3.6.2. Scaling EUD. Bortfeld *et al* (2003) propose to measure EUD criteria relative to predefined values. In particular, they propose to use

$$\frac{\underline{d} - \mathrm{gEUD}(\vec{d};-\infty)}{\underline{d}} \tag{9}$$

for targets, and

$$\frac{\mathrm{EUD}^{\mathrm{D\&O}}(\vec{d};\alpha) - \overline{d}}{\overline{d}} \tag{10}$$

for critical structures, where $\underline{d}$ and $\overline{d}$ are some pre-defined desirable levels of the criteria $\mathrm{gEUD}(\vec{d};-\infty)$ and $\mathrm{EUD}^{\mathrm{D\&O}}(\vec{d};\alpha)$, respectively. In a multi-criteria setting, this approach is equivalent to using the criteria $\mathrm{gEUD}(\vec{d};-\infty)$ and $\mathrm{EUD}^{\mathrm{D\&O}}(\vec{d};\alpha)$, and the discussion in section 3.6.1 applies.

3.6.3. Probability of uncomplicated tumour control. The probability of uncomplicated tumour control, P_+, is a measure that has been proposed to make a trade-off between TCP values for targets and NTCP values for critical structures. In particular, P_+ is defined as

$$P_+(\vec{d}_1, \ldots, \vec{d}_S) = \prod_{s=1}^{T} \mathrm{TCP}_s(\vec{d}_s) \cdot \prod_{s=T+1}^{S} (1 - \mathrm{NTCP}_s(\vec{d}_s))$$

(see, e.g., Brahme 2001). Since larger values of P_+ are preferred to smaller values, in the remainder of this section we will consider its negative $-P_+$. Now note that we can decompose $-P_+$ as

$$-P_+(\vec{d}_1, \ldots, \vec{d}_S) = h\left(-\sum_{s=1}^{T} \ln \mathrm{TCP}_s(\vec{d}_s) - \sum_{s=T+1}^{S} \ln(1 - \mathrm{NTCP}_s(\vec{d}_s))\right)$$

where $h(z) = -\mathrm{e}^{-z}$. Theorem 2.2 says that the efficient frontier for a multi-criteria formulation of the FMO problem does not change if we replace the criterion $-P_+$ by the combined criterion

$$-\sum_{s=1}^{T} \ln \mathrm{TCP}_s(\vec{d}_s) - \sum_{s=T+1}^{S} \ln(1 - \mathrm{NTCP}_s(\vec{d}_s)).$$

As we have mentioned above, $-\ln \mathrm{TCP}_s(\vec{d}_s)$ is convex. Moreover, note that

$$\begin{aligned} -\ln\left(1 - \mathrm{NTCP}_s^{\mathrm{LS}}(\vec{d}_s; a, m, D_{50})\right) &= -\ln\left(1 - \Phi\left(\frac{\mathrm{gEUD}_s(\vec{d}_s; a) - D_{50}}{m D_{50}}\right)\right) \\ &= \breve{W}(\vec{d}_s; a, m, D_{50}) \end{aligned}$$

and

$$\begin{aligned} -\ln\left(1 - \mathrm{NTCP}_s^{\mathrm{A\&N}}(\vec{d}_s; a, \Delta)\right) &= \left(\frac{\mathrm{gEUD}_s(\vec{d}_s; a)}{\Delta}\right)^a \\ &= \frac{1}{\Delta^a} \cdot F_s^{\mathrm{g}}(\vec{d}_s; a) \end{aligned}$$

which means that $-\ln(1 - \mathrm{NTCP}_s(\vec{d}_s))$ is convex for both the LS and the A&N models. This then implies that we can explore a vastly larger set of efficient treatment plans by considering the individual criteria $-\ln \mathrm{TCP}_s(\vec{d}_s)$ and $-\ln(1 - \mathrm{NTCP}_s(\vec{d}_s))$. By table 1, the former criteria are in fact equivalent to $\mathrm{TCP}_s(\vec{d}_s)$. Similarly, by choosing $h(z) = -\ln(1 - z)$, the latter criteria are equivalent to $\mathrm{NTCP}_s(\vec{d}_s)$. Therefore, the treatment plans that are Pareto efficient with respect to the criterion P_+ can be found by considering the more elementary TCP and NTCP measures. However, the latter measures are preferable, since they do not impose an *a priori* trade-off between these measures. Finally, we note that any multi-criteria model based on P_+, TCP and NTCP measures can be transformed into an equivalent model based on convex separable criteria, leading to voxel-based model formulations for FMO.

As a last remark, the analysis of P_+ reveals a close relationship between NTCP and $\breve{W}$. In fact, if we consider measuring TCP according to the L&S model and define

$$\mathrm{TCP}^{\mathrm{L\&S}}(\vec{d}; a, m, D_{50}) = \Phi\left(\frac{\mathrm{gEUD}(\vec{d}; a) - D_{50}}{m D_{50}}\right)$$

for $-\infty \leqslant a \leqslant 0$ and $m, D_{50} > 0$, we obtain that

$$\begin{aligned} -\ln\left(\mathrm{TCP}_s^{\mathrm{LS}}(\vec{d}_s; a, m, D_{50})\right) &= -\ln\left(\Phi\left(\frac{\mathrm{gEUD}_s(\vec{d}_s; a) - D_{50}}{m D_{50}}\right)\right) \\ &= -\ln\left(1 - \Phi\left(\frac{D_{50} - \mathrm{gEUD}_s(\vec{d}_s; a)}{m D_{50}}\right)\right) \\ &= \breve{W}(\vec{d}_s; a, m, D_{50}). \end{aligned}$$

The corresponding measure of P_+ is then given by

$$P_+^{\mathrm{L\&S}} = \prod_{s=1}^{T} \mathrm{TCP}_s^{\mathrm{L\&S}}(\vec{d}_s) \cdot \prod_{s=T+1}^{S} \left(1 - \mathrm{NTCP}_s^{\mathrm{L\&S}}(\vec{d}_s)\right) = \prod_{s=1}^{S} \breve{w}(\vec{d}_s)$$

establishing a very close relationship between P_+ and the alternative to the criterion by Wu *et al* (2002).

4. Discussion and conclusions

We have a presented a unifying framework for studying multi-criteria models for FMO. Casting FMO models as multi-criteria models has allowed us to establish the equivalence between many sets of treatment plan evaluation criteria that have been proposed in the literature. We have also established conditions under which nonconvex criteria can be transformed into convex criteria without changing the set of Pareto efficient treatment plans.

The framework identifies which of the proposed treatment plan evaluation criteria are truly different. This enables researchers to focus their efforts on comparing non-equivalent models. We have shown that using the ‘biological’ criteria considered in this paper, which include EUD, gEUD, TCP, NTCP, P_+, as well as two sigmoidal transformations of EUD, only two distinct Pareto efficient frontiers exist. We have shown that these frontiers can be obtained using equivalent ‘physical’ convex voxel-based criteria. This result should alleviate concerns on the pros and cons of using biological versus physical criteria that was recently debated in the literature, as they are in essence equivalent (see, e.g., Wu *et al* (2002), Amols and Ling 2002, Thieke *et al* (2003c)). Clearly, many criteria still remain to be analysed. To avoid duplication of efforts, in the pursuit of new treatment plan evaluation criteria one should consider the unifying framework to determine whether potentially new criteria indeed yield different Pareto frontiers and therefore truly new models.

There exist treatment plan evaluation criteria that are fundamentally nonconvex, i.e., transformations to convex criteria using the framework do not exist. Examples of these are traditional dose-volume histogram criteria (see, e.g., Deasy (1997)) and certain measures of delivery efficiency such as the number of beam orientations used (see, e.g., Bortfeld and Schlegel (1993)). The consequence of including nonconvex treatment plan evaluation criteria into a multi-criteria optimization framework is that the Pareto efficient frontier may be nonconvex or even disconnected. This makes the comparison of multi-criteria models based on these criteria with multi-criteria models based on biological or voxel-based physical criteria impractical given the current state-of-the-art in solving the associated optimization problems (see, e.g., Lee *et al* (2003)). In particular, when using heuristic approaches for solving problems that are aimed at identifying points on the Pareto efficient frontier, it cannot be guaranteed that the obtained solutions are indeed Pareto efficient (see, e.g., Schreibmann *et al* (2004)). With respect to dose-volume histograms, Romeijn *et al* (2003) have recently proposed an alternative family of convex criteria that measure the tail mean of a differential dose-volume histogram. Although this means that these criteria could in principle be incorporated into convex multi-criteria optimization models, no equivalent voxel-based criteria exist. However, given the widespread interest in and potential value of traditional dose-volume histogram criteria and beam number and orientation, future research will address nonconvex multi-criteria models for FMO.

Appendix A

The following theorem shows that Niemierko's original EUD (see Niemierko (1997)) is a concave function of the dose distribution.

Theorem A.1. *The function* $\mathrm{EUD}(\vec{d};\alpha)$ *is concave in* $\vec{d}$ *when* $\alpha > 0$.

Proof. For notational simplicity, let us assume that the unit of dose is normalized to make $\alpha = 1$ and define

$$G(\vec{d}) = \mathrm{EUD}(\vec{d};1) = -\ln\left(\frac{1}{v}\sum_{j=1}^{v} \mathrm{e}^{-d_j}\right).$$

The first derivatives of this function are equal to

$$\frac{\partial G(\vec{d})}{\partial d_k} = \frac{\mathrm{e}^{-d_k}}{\sum_{j=1}^{v}\mathrm{e}^{-d_j}} \qquad k = 1,\ldots,v$$

which implies that G is nondecreasing in d_k for all $k = 1,\ldots,v$. The second derivatives of G are given by

$$\frac{\partial^2 G(\vec{d})}{\partial d_k^2} = \frac{\mathrm{e}^{-2d_k} - \mathrm{e}^{-d_k}\left(\sum_{j=1}^{v}\mathrm{e}^{-d_j}\right)}{\left(\sum_{j=1}^{v}\mathrm{e}^{-d_j}\right)^2} \qquad k = 1,\ldots,v$$

$$\frac{\partial^2 G(\vec{d})}{\partial d_k \partial d_{k'}} = \frac{\mathrm{e}^{-d_k-d_{k'}}}{\left(\sum_{j=1}^{v}\mathrm{e}^{-d_j}\right)^2} \qquad k,k' = 1,\ldots,v;\, k \neq k'.$$

Ignoring the common, positive denominator in the partial derivatives and denoting the resulting matrix by $\hat{H}(\vec{d})$, we can study the definiteness of this matrix by studying the sign of the quadratic form $y^\top \hat{H}(\vec{d})y$ for all vectors y:

$$\begin{aligned}
y^\top \hat{H}(\vec{d})y &= \left(\sum_{k=1}^{v}\mathrm{e}^{-d_k}y_k\right)\cdot\left(\sum_{k'=1}^{v}\mathrm{e}^{-d_{k'}}y_{k'}\right) - \left(\sum_{k=1}^{v}\mathrm{e}^{-d_k}\right)\cdot\left(\sum_{k'=1}^{v}\mathrm{e}^{-d_{k'}}y_{k'}^2\right)\\
&= \sum_{k=1}^{v}\sum_{k'=1}^{v}\mathrm{e}^{-d_k-d_{k'}}y_k y_{k'} - \sum_{k=1}^{v}\sum_{k'=1}^{v}\mathrm{e}^{-d_k-d_{k'}}y_{k'}^2\\
&= 2\sum_{k=1}^{v}\sum_{k'=k+1}^{v}\mathrm{e}^{-d_k-d_{k'}}y_k y_{k'} - \sum_{k=1}^{v}\sum_{k'=1}^{k-1}\mathrm{e}^{-d_k-d_{k'}}y_{k'}^2 - \sum_{k=1}^{v}\sum_{k'=k+1}^{v}\mathrm{e}^{-d_k-d_{k'}}y_{k'}^2\\
&= 2\sum_{k=1}^{v}\sum_{k'=k+1}^{v}\mathrm{e}^{-d_k-d_{k'}}y_k y_{k'} - \sum_{k=1}^{v}\sum_{k'=k+1}^{v}\mathrm{e}^{-d_{k'}-d_k}y_k^2 - \sum_{k=1}^{v}\sum_{k'=k+1}^{v}\mathrm{e}^{-d_k-d_{k'}}y_{k'}^2
\end{aligned}$$

$$
\begin{aligned}
&= 2\sum_{k=1}^{v}\sum_{k'=k+1}^{v}\left(\mathrm{e}^{-(d_k+d_{k'})/2}y_k\right)\cdot\left(\mathrm{e}^{-(d_k+d_{k'})/2}y_{k'}\right)\\
&\quad -\sum_{k=1}^{v}\sum_{k'=k+1}^{v}\left(\mathrm{e}^{-(d_k+d_{k'})/2}y_k\right)^2-\sum_{k=1}^{v}\sum_{k'=k+1}^{v}\left(\mathrm{e}^{-(d_k+d_{k'})/2}y_{k'}\right)^2\\
&= -2\sum_{k=1}^{v}\sum_{k'=k+1}^{v}\left(\left(\mathrm{e}^{-(d_k+d_{k'})/2}y_k\right)-\left(\mathrm{e}^{-(d_k+d_{k'})/2}y_{k'}\right)\right)^2\\
&\leqslant 0
\end{aligned}
$$

for all y, which implies that G is concave. □

Appendix B

The following lemma is useful to derive properties of the criterion function W proposed by Wu *et al* (2002).

Lemma B.1. *Define the function*

$$\xi_k(z) = \ln(1+z^k) \qquad z \geqslant 0$$

for $k \in \mathbb{R}\backslash\{0\}$. *If* $k < 0$, ξ_k *is convex on* $\mathbb{R}_+$. *If* $0 < k \leqslant 1$, ξ_k *is concave on* $\mathbb{R}_+$. *If* $k > 1$, ξ_k *is convex on* $[0, \sqrt[k]{k-1}]$ *and concave on* $[\sqrt[k]{k-1}, \infty)$.

Proof. The first derivative of ξ_k is equal to

$$\xi_k'(z) = \frac{kz^{k-1}}{1+z^k}$$

which means that ξ_k is increasing if $k > 0$ and decreasing if $k < 0$. Its second derivative is equal to

$$\xi_k''(z) = \frac{kz^{k-2}(k-1-z^k)}{(1+z^k)^2}.$$

If $k < 0$, we have that ξ_k is convex if

$$z^k \geqslant k-1$$

which means that ξ_k is convex on $\mathbb{R}_+$ whenever $k < 0$. If $k > 0$, we have that ξ_k is convex if

$$z^k \leqslant k-1$$

and concave otherwise. This means that ξ_k is concave on $\mathbb{R}_+$ if $0 < k \leqslant 1$. If $k > 1$, ξ_k is convex on $[0, \sqrt[k]{k-1}]$ and concave on $[\sqrt[k]{k-1}, \infty)$. □

This then leads to the following

Theorem B.2. *The function* $W(\vec{d}; a, k, \mathrm{gEUD}^0)$ *is concave in* $\vec{d}$ *whenever* $-\infty \leqslant a \leqslant 0$ *and* $k, \mathrm{gEUD}^0 > 0$. *The function* $W(\vec{d}; a, k, \mathrm{gEUD}^0)$ *is convex in* $\vec{d}$ *in the region where* $\mathrm{gEUD}(\vec{d}; a) \leqslant \mathrm{gEUD}_0\sqrt[k]{k-1}$ *whenever* $1 \leqslant a \leqslant \infty$, $k > 1$, *and* $\mathrm{gEUD}^0 > 0$.

Proof. If $-\infty \leqslant a \leqslant 0$, we have that

$$W(\vec{d}; a, k, \mathrm{gEUD}_0) = \xi_{-k}\left(\frac{\mathrm{gEUD}(\vec{d}; a)}{\mathrm{gEUD}_0}\right).$$

Since $k > 0$, lemma B.1 says that ξ_{-k} is a decreasing and convex function. Since $a \leqslant 0$, $\mathrm{gEUD}(\vec{d}; a)$ is a concave function of $\vec{d}$, and we conclude that W is a convex function of $\vec{d}$.

If $1 \leqslant a \leqslant \infty$, we have that

$$W(\vec{d}; a, k, \mathrm{gEUD}_0) = \xi_k\left(\frac{\mathrm{gEUD}(\vec{d}; a)}{\mathrm{gEUD}_0}\right).$$

If $k > 1$, lemma B.1 says that ξ_k is an increasing function that is convex on $[0, \sqrt[k]{k-1}]$. Moreover, since $a \geqslant 1$, $\mathrm{gEUD}(\vec{d}; a)$ is a convex function of $\vec{d}$. Therefore, we conclude that W is a convex function of $\vec{d}$ in the region where $\mathrm{gEUD}(\vec{d}; a) \leqslant \mathrm{gEUD}_0\sqrt[k]{k-1}$. □

In the last theorem of this appendix, we derive properties of the alternative criterion $\breve{W}$.

Theorem B.3. *The function* $\breve{W}(\vec{d}; a, \sigma, \mathrm{gEUD}^0)$ *is convex in* $\vec{d}$ *whenever* $-\infty \leqslant a \leqslant 0$ *or* $1 \leqslant a \leqslant \infty$ *and* $\sigma, \mathrm{gEUD}^0 > 0$.

Proof. Defining

$$\psi(z) = -\ln(1 - \Phi(z))$$

we can write

$$\breve{W}(\vec{d}; a, \sigma, \mathrm{gEUD}^0) = \begin{cases} \psi\left(\dfrac{\mathrm{gEUD}^0 - \mathrm{gEUD}(\vec{d}; a)}{\sigma \cdot \mathrm{gEUD}^0}\right) & \text{if } -\infty \leqslant a \leqslant 0 \\ \psi\left(\dfrac{\mathrm{gEUD}(\vec{d}; a) - \mathrm{gEUD}^0}{\sigma \cdot \mathrm{gEUD}^0}\right) & \text{if } 1 \leqslant a \leqslant \infty. \end{cases}$$

It is straightforward to derive that

$$\psi'(z) = \frac{\Phi'(z)}{1 - \Phi(z)}$$

where $\Phi'(z)$ is the probability density function of a standard normal random variable. Clearly, this implies that ψ is increasing. Moreover, $\psi'(z)$ is the so-called failure rate of the standard normal distribution, an applied probability concept that is often used in the area of reliability. It is known that this failure rate is increasing (see, e.g., Park (1987)), which implies that ψ is convex. Now recall that $-\mathrm{gEUD}(a)$ is convex for $-\infty \leqslant a \leqslant 0$ and $\mathrm{gEUD}(a)$ is convex for $1 \leqslant a \leqslant \infty$; we obtain that $\breve{W}$ is an increasing convex function of a convex function and therefore convex. □

References

Abramowitz M and Stegun I A (ed) 1965 *Handbook of Mathematical Functions* (New York: Dover)

Alber M and Nüsslin F 1999 An objective function for radiation treatment optimization based on local biological measures *Phys. Med. Biol.* **44** 479–93

Alber M and Nüsslin F 2001 A representation of an NTCP function for local complication mechanisms *Phys. Med. Biol.* **46** 439–47

Amols H I and Ling C C 2002 EUD but not QED *Int. J. Radiat. Oncol. Biol. Phys.* **52** 1–2

Bortfeld T, Küfer K-H, Monz M, Scherrer A, Thieke C and Trinkhaus H 2003 Intensity-modulated radiotherapy: a large scale multi-criteria programming problem *Berichte des Fraunhofer ITWM* 43 (Kaiserslautern, Germany: Fraunhofer Insitut Techno- und Wirtschaftsmathematik)

Bortfeld T and Schlegel W 1993 Optimization of beam orientations in radiation therapy: some theoretical considerations *Phys. Med. Biol.* **38** 291–304

Brahme A 2001 Individualizing cancer treatment: biological optimization models in treatment planning and delivery *Int. J. Radiat. Oncol. Biol. Phys.* **49** 327–37

Brahme A and Agren A K 1987 Optimal dose distribution for eradication of heterogeneous tumours *Acta Oncol.* **26** 377–85

Choi B and Deasy J O 2002 The generalized equivalent uniform dose function as a basis for intensity-modulated treatment planning *Phys. Med. Biol.* **47** 3579–89

Dale E and Olsen D R 1997 Specification of the dose to organs at risk in external beam radiotherapy *Acta Oncol.* **36** 129–35

Deasy J O 1997 Multiple local minima in radiotherapy optimization problem with dose-volume constraints *Med. Phys.* **24** 1157–61

Goitein M 1987 Tumour control probability for an inhomogeneously irradiated target volume *Evaluation of Treatment Planning for Particle Beam Radiotherapy* (Bethesda, MD: National Cancer Institute)

Hamacher H W and Küfer K-H 2002 Inverse radiation therapy planning—a multiple objective optimization approach *Discrete Appl. Math.* **118** 145–61

Küfer K-H, Hamacher H W and Bortfeld T R 2000 A multicriteria optimization approach for inverse radiotherapy planning *Proc. XIIIth ICCR* ed T R Bortfeld and W Schlegel *(Heidelberg, Germany)* pp 26–9

Kutcher G J and Burman C 1989 Calculation of complication probability factors for non-uniform normal tissue irradiation: the effective volume method *Int. J. Radiat. Oncol. Biol. Phys.* **16** 1623–30

Lahanas M, Schreibmann E and Baltas D 2003 Multiobjective inverse planning for intensity modulated radiotherapy with constraint-free gradient-based optimization algorithms *Phys. Med. Biol.* **48** 2843–71

Lee E K, Fox T and Crocker I 2003 Integer programming applied to intensity-modulated radiation treatment planning *Ann. Oper. Res.* **119** 165–81

Lyman J T 1985 Complication probability as assessed from dose-volume histograms *Radiat. Res.* **8** S13–9

Lyman J T and Wolbarst A B 1987a Optimization of radiation therapy III: a method of assessing complication probabilities from dose-volume histograms *Int. J. Radiat. Oncol. Biol. Phys.* **13** 103–9

Lyman J T and Wolbarst A B 1987b Optimization of radiation therapy IV: a dose-volume histogram reduction algorithm *Int. J. Radiat. Oncol. Biol. Phys.* **13** 103–9

McGary J E, Grant III W and Woo S Y 2000 Applying the equivalent uniform dose formulation based on the linear-quadratic model to inhomogeneous tumor dose distribution: caution for analyzing and reporting *J. Appl. Clin. Med. Phys.* **1** 126–37

Miettinen K 1999 *Nonlinear Multiobjective Optimization* (Boston, MA: Kluwer)

Niemierko A 1997 Reporting and analyzing dose distributions: a concept of equivalent uniform dose *Med. Phys.* **24** 103–10

Niemierko A 1999 A generalized concept of equivalent uniform dose (EUD) *AAPM Annual Meeting, (Nashville, Tennessee)* Abstract *Med. Phys.* **26** 1100

Niemierko A and Goitein M 1991 Calculation of normal tissue complication probability and dose-volume histogram reduction schemes for tissues with a critical element architecture *Radiother. Oncol.* **20** 166–76

Pareto V 1971 *Manual of Political Economy* (New York: Kelley) (Translation of Manuale di economia politica, 1906)

Park K S 1987 *Human Reliability* (Amsterdam: Elsevier)

Romeijn H E, Ahuja R K, Dempsey J F, Kumar A and Li J G 2003 A novel linear programming approach to fluence map optimization for intensity modulated radiation therapy treatment planning *Phys. Med. Biol.* **48** 3521–42

Schreibmann E, Lahanas M, Xing L and Baltas D 2004 Multiobjective evolutionary optimization of the number of beams, their orientations and weights for intensity-modulated radiation therapy *Phys. Med. Biol.* **49** 747–70

Schultheiss T E, Orton C G and Peck R A 1983 Models in radiotherapy: volume effects *Med. Phys.* **10** 410–5

Shepard D M, Ferris M C, Olivera G H and Mackie T R 1999 Optimizing the delivery of radiation therapy to cancer patients *SIAM Rev.* **41** 721–44

Stavrev P, Hristov D, Warkentin B, Sham E, Stavreva N and Fallone B G 2003 Inverse treatment planning by physically constrained minimization of a biological objective function *Med. Phys.* **30** 2948–58

Stavreva N A, Stavrev P, Warkentin B and Fallone B G 2003 Investigating the effect of cell repopulation on the tumor response to fractionated external radiotherapy *Med. Phys.* **30** 735–42

Steuer R 1986 *Multicriteria Optimization: Theory, Computation and Applications* (New York: Wiley)

Thieke C, Bortfeld T and Küfer K-H 2002 Characterization of dose distributions through the Max and Mean dose concept *Acta Oncol.* **41** 158–61

Thieke C, Bortfeld T and Küfer K H 2003a New optimization concepts in inverse treatment planning *Pre-meeting Workshop on 'Optimization of IMRT', 7th Biennial ESTRO Meeting on Physics and Radiation Technology for Clinical Radiotherapy (Geneva, Switzerland, Sept. 2003)* Abstract *Radiother. Oncol.* **68** (Suppl. 1)S2

Thieke C, Bortfeld T, Niemierko A, Küfer K H and Nill S 2003b Multicriteria optimization in inverse radiotherapy planning *7th Biennial ESTRO Meeting on Physics and Radiation Technology for Clinical Radiotherapy (Geneva, Switzerland, Sept. 2003)* Abstract *Radiother. Oncol.* **68** (Suppl. 1)S44

Thieke C, Bortfeld T, Niemierko A and Nill S 2003c From physical dose constraints to equivalent uniform dose constraints in inverse radiotherapy planning *Med. Phys.* **30** 2332–9

Tsien C, Eisbruch A, McShan D, Kessler M, Marsh R and Fraass B 2003 Intensity-modulated radiation therapy (IMRT) for locally advanced paranasal sinus tumors: incorporating clinical decisions in the optimization process *Int. J. Radiat. Oncol. Biol. Phys.* **55** 776–84

Winston W L 2004 *Operations Research: Applications and Algorithms* 4th edn (Belmont, CA: Thomson Learning Inc.)

Withers H R and McBride W H 1998 Biologic basis of radiation therapy Principles and Practice of Radiotherapy ed C A Perez and L W Brady (Philadelphia, PA: Williams and Wilkins) pp 79–118 chapter 2

Wolbarst A B 1984 Optimization of radiation therapy II: the critical-voxel method *Int. J. Radiat. Oncol. Biol. Phys.* **10** 741–5

Wu Q, Mohan R, Niemierko A and Schmidt-Ullrich R 2002 Optimization of intensity-modulated radiotherapy plans based on the equivalent uniform dose *Int. J. Radiat. Oncol. Biol. Phys.* **52** 224–35

Zaider M and Minerbo G N 2000 Tumour control probability: a formulation applicable to any temporal protocol of dose delivery *Phys. Med. Biol.* **45** 279–93