Convex reformulation of biologically-based multi-crtiteria intensity-modulated radiation therapy optimization including fractionation effects

Hoffmann, A.L.; Siem, A.Y.D.; den Hertog, D.; Kaanders, J.H.A.M.; Huizenga, H.

Published in:
Physics in Medicine and Biology

Document version:
Peer reviewed version

Publication date:
2008

Convex reformulation of biologically-based multi-criteria IMRT optimisation including fractionation effects

Aswin L Hoffmann[1], Dick den Hertog[2], Alex Y D Siem[2], Johannes H A M Kaanders[1], and Henk Huizenga[1]

[1] Department of Radiation Oncology, Radboud University Nijmegen Medical Centre, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands
[2] Department of Econometrics and Operations Research/Center for Economic Research (CentER), Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands

E-mail: a.hoffmann@rther.umcn.nl

Abstract. Finding fluence maps for intensity-modulated radiation therapy (IMRT) can be formulated as a multi-criteria optimisation problem for which Pareto optimal treatment plans exist. To account for the dose-per-fraction effect of fractionated IMRT, it is desirable to exploit radiobiological treatment plan evaluation criteria based on the linear-quadratic (LQ) cell survival model as a means to balance the radiation benefits and risks in terms of biologic response. Unfortunately, the LQ-model based radiobiological criteria are nonconvex functions, which make the optimisation problem hard to solve. We apply the framework proposed by Romeijn, Dempsey & Li (2004) (*Phys. Med. Biol.* 2004, **49** 1991-2013) to find transformations of LQ-model based radiobiological functions and establish conditions under which transformed functions result in equivalent convex criteria that do not change the set of Pareto optimal treatment plans. The functions analysed are: the LQ-Poisson based model for tumour control probability (TCP) with and without inter-patient heterogeneity in radiation sensitivity, the LQ-Poisson based Relative Seriality *s*-model for normal tissue complication probability (NTCP), the equivalent uniform dose (EUD) under the LQ-Poisson model, and the fractionation-corrected Probit-based model for NTCP according to Lyman, Kutcher and Burman. These functions differ from the ones analysed before in that they cannot be decomposed into elementary EUD or generalized-EUD functions. In addition, we show that applying increasing and concave transformations to the convexified functions is beneficial for the piecewise approximation of the Pareto efficient frontier.

Keywords: Radiotherapy, intensity-modulated radiation therapy, fluence map optimisation, convex optimisation, radiobiology

1. Introduction

It is generally acknowledged that designing fluence maps for high-energy photon beams in intensity-modulated radiation therapy (IMRT) can be posed as a constrained optimisation problem (Börgers 1997, Küfer, Scherrer, Monz, Alonso, Trinkaus, Bortfeld & Thieke 2003, Reemtsen & Alber 2006). The aim of the optimisation problem is to find beam fluence maps that deposit a sufficiently high dose to the target volume and simultaneously spare organs at risk (OARs) and other surrounding normal tissue as much as possible. A mathematical fluence map optimisation (FMO) model is required to guide the inverse treatment planning process through the search space of possible solutions. Such models are typically based on a set of conflicting treatment plan evaluation criteria that reflect how well the treatment goals and restrictions are satisfied.

These criteria can either be formulated as physical criteria, i.e. on measurable physical quantities like doses and volumes, or as biological criteria that reflect the responses of the different tissues to dose distributions (Brahme 1995). Physical criteria are often implemented as a dose-dependent function calculating the mean-squared deviation from a prescribed target dose level. The radiobiological rationale of these criteria is questionable since positive and negative deviations from the target dose have different biological consequences, but are weighted equally using a quadratic penalty function. In this paper, we focus on biological criteria only, because they have been shown to predict the response of tumours and healthy tissues more adequately than physical criteria by taking the underlying radiation biology into account. The relevance of biological treatment goals, leading to nonlinear criteria as tumour control probability (TCP), normal tissue complication probability (NTCP), (generalized) equivalent uniform dose ((g)EUD) or dose-volume based criteria, has generally been acknowledged during the past years (Niemierko 2005). Formulation of clinically relevant radiobiological FMO problems often requires the inequality constraints to be formulated in terms of clinically prescribed tolerance bounds, e.g., NTCP $\leq \epsilon$, where ϵ is an acceptable probability of injury.

Direct application of such nonlinear and nonconvex biological criteria as objective and constraint functions can make the FMO problem very hard to solve, especially when the number of optimisation variables (i.e., the beamlet weight intensities) is large. According to Hindi (2004) various sources exist for the difficulties encountered in large-scale nonconvex optimisation. First, the search space may be riddled with multiple local optima. Second, the feasible set could be empty. Third, stopping criteria used in general optimisation algorithms are often arbitrary. Forth, the problem may be degenerate, in which case the same optimal value is attained for multiple solutions (see, e.g., Alber, Meedt, Nusslin & Reemtsen (2002)). Fifth, optimisation algorithms might have poor convergence rates. Sixth, numerical problems could cause the optimisation algorithm to stop or roam.

Fortunately, in case of minimisation, it is known that if all objective functions and upperbound-inequality constraint functions are *convex*, the first three difficulties disappear: any local minimum is necessarily a global minimum; the solution set is convex and therefore comprises either a single solution or is infinite; and very precise stopping criteria are available using *duality* (see, e.g., Bertsekas (1999)). However, convergence rate and numerical stability remain a potential problem. If, in addition to convexity, the objective and constraint functions allow for *self-concordant* barriers, the issues of convergence and numerical sensitivity could be avoided using interior point

methods (Nesterov & Nemirovski 1994). Hence, it is possible to solve a large class of convex optimisation problems with great efficiency using local solvers like gradient-based algorithms. Convexity of the criteria is also sufficient to guarantee that the Pareto efficient frontier is convex in case the FMO model is multi-criteria (Romeijn et al. 2004). This is of particular interest in case piecewise linear approximation techniques like Sandwich type algorithms are used to estimate the Pareto efficient frontier (Craft, Halabi, Shih & Bortfeld 2006, Hoffmann, Siem, den Hertog, Kaanders & Huizenga 2006).

Radiobiological criteria like TCP and NTCP are often sigmoidal functions of dose and hence are inherently nonlinear and nonconvex. Consequently, their direct implementation necessarily leads to nonconvex optimisation problems. Recently, Romeijn et al. (2004) presented a unifying framework for multi-criteria fluence map optimisation problems that establishes conditions under which well-known nonconvex radiobiological treatment plan evaluation criteria can be transformed into convex criteria while preserving the set of Pareto efficient solutions. In particular, they showed that transformations of criteria such as TCP, NTCP and sigmoidal functions of (g)EUD exist that are equivalent to criteria formulated in terms of (g)EUD only, concluding that only two distinct Pareto efficient frontiers exist. Others have explored the convexity properties of transformed radiobiological treatment plan evaluation criteria as well (Deasy 1997, Choi & Deasy 2002). However, to the best of our knowledge the majority of convexity analyses reported so far are limited to simple, single-hit *linear* cell survival models that do not take into account fractionation effects, whereas in clinical practice almost all radiation treatments are delivered over multiple fractions.

As an extension of the work by Romeijn et al. (2004) we explicitly include the dose-per-fraction effect in the convexity analysis of commonly occuring radiobiological criterion functions by using the *linear quadratic* (LQ) cell survival model (see, e.g., Fowler (1989)). We analyse TCP, NTCP and EUD criteria that are not related to the elementary (linear-Poisson) EUD or (power-law) gEUD model, and thus yield different Pareto efficient frontiers from the criteria analysed by Romeijn et al. (2004). More specifically, we establish transformations of the following criterion functions and investigate the conditions under which these criteria are *strictly* convex/concave depending on the criterion to be minimised/maximised, respectively: LQ- Poisson model based TCP function, LQ-Poisson model based Relative Seriality NTCP function, LQ-Poisson model based EUD function, and the fractionation-corrected Probit-model-based NTCP function according to Lyman, Kutcher, and Burman. Strict convexity/concavity of the objective function is an interesting property, as it is a sufficient condition for the existence of a unique solution (i.e., minimiser/maximiser) that rules out the possibility of multiple local optima (i.e., minima/maxima) (see, e.g., Bertsekas (1999)). Furthermore, we present a detailed convexity analysis of the (transformed) population- averaged TCP function that incorporates inter-patient radiosensitivity heterogeneity, for which Choi & Deasy (2002) could not rule out the possibility of multiple local minima.

2. Mathematical definitions

2.1. Multi-criteria optimisation for IMRT

To introduce the problem, we confine to the essential mathematics and refer to (Hoffmann et al. 2006) for a more comprehensive enunciation. In IMRT optimisation, the dose distribution $\mathbf{D} = (D_1, \ldots, D_M)$ is a linear mapping of the bixel weights $\mathbf{w} = (w_1, \ldots, w_N)$ by the dose deposition operator $\mathbf{P}$:

$$\mathbf{D}(\mathbf{w}) = \mathbf{P} \cdot \mathbf{w}.$$

Hence, the dose in voxel i can be denoted as a weighted sum over all N bixel weights

$$D_i(\mathbf{w}) = \sum_{j=1}^{N} P_{ij}\ w_j,$$

where P_{ij} is the dose deposited in voxel i from bixel j at unit intensity. This represents the discretised form of the Fredholm integral equation of the first kind that is commonly encountered in solving the inverse problem of radiation treatment planning (Lind 1990). The elements of $\mathbf{P}$ describe the physics of the beam-tissue interaction, and can be pre-calculated using various dose calculation algorithms.

Given the matrix $\mathbf{P}$, it is the aim of the optimizer to find a suitable vector $\mathbf{w}$ that satisfies the optimisation goals. Typically, the goals are formulated by a set of *objective functions* $F_k : \mathbb{R}_+^M \mapsto \mathbb{R}_+$, $k = 1, \ldots, K$, where K is the number of objective functions. The *constraint functions* $B_l : \mathbb{R}_+^M \mapsto \mathbb{R}_+$, $l = 1, \ldots, L$, where L is the number of constraint functions, define certain equalities and inequalities that the design variable $\mathbf{w}$ and the dose distribution $\mathbf{D}(\mathbf{w})$ must satisfy. Each criterion function quantifies the plan-evaluation score as a function of the dose distribution $\mathbf{D}(\mathbf{w})$. Without loss of generality, it is assumed that lower values are preferred to higher values for each of the criterion functions.

Typically, the constrained multi-criteria optimisation problem in IMRT is formulated as:

$$\begin{array}{ll} \min\limits_{\mathbf{w}} & \mathbf{F}(\mathbf{D}(\mathbf{w})) = \begin{pmatrix} F_1(\mathbf{D}(\mathbf{w})) \\ F_2(\mathbf{D}(\mathbf{w})) \\ \vdots \\ F_K(\mathbf{D}(\mathbf{w})) \end{pmatrix} \\ \text{s.t.} & B_l(\mathbf{D}(\mathbf{w})) \leq 0, \qquad \text{l} = 1, \ldots, \text{L} \\ & \mathbf{D}(\mathbf{w}) = \mathbf{P} \cdot \mathbf{w} \\ & \mathbf{w} \geq 0. \end{array} \tag{1}$$

Since probem (1) is associated with multiple solutions, the concept of *Pareto efficiency* (see, e.g., Miettinen (1999)) is applied to restrict to solutions that have the property that no single objective value can be improved without deteriorating at least one other objective value. Solutions that comply with this definition are called *Pareto optimal* (also called Edgeworth-Pareto optimal, efficient, nondominated or noninferior). In the objective space, the set of Pareto optimal plans is mapped to the *Pareto efficient frontier*. Since the dimension of the objective space is much less than that of the solution space, this frontier is used by the decision maker to navigate efficiently through the set of solutions and select a compromise solution that best meets with his/her approval. Therefore, it is important to have an algorithm that generates a discrete subset of Pareto optimal solutions that forms a representative estimate of the true Pareto efficient frontier (Craft et al. 2006, Hoffmann et al. 2006).

2.2. Commonly used radiobiological treatment plan evaluation criteria

Several mathematical models have been developed to describe the dose-response relationship for tumours and normal tissues. We analyse a commonly used mechanistic TCP model and two well-known NTCP models; a mechanistic and a phenomenological one.

2.2.1. Mechanistic dose-response relationship based on Poisson model To calculate the dose response to a heterogeneous dose distribution, the internal infrastructural organization (i.e., volume dependence) of the irradiated tissue is taken into account.

For eradication of all clonogenic cells of a tumour, every individual tumour element has to be eradicated. Hence, the tumour control probability (TCP) is the product of all individual responses. Based on the well-known linear-quadratic (LQ) Poisson model without cell repopulation in time (Fowler 1989), the TCP function is denoted as:

$$\mathrm{TCP}_{\mathrm{LQ}}(\mathbf{D}) = \exp\left[-N_0 \sum_{k=1}^{N} v_k \ \mathrm{SF}_{\mathrm{LQ}}(D_k)\right], \tag{2}$$

where $\mathrm{SF}_{\mathrm{LQ}}(D_k) = \mathrm{e}^{-\alpha D_k - \beta {D_k}^2/n}$ is the *surviving fraction* of cells exposed to dose D_k in voxel k, v_k is the relative volume of voxel k, N_0 is the total initial number of clonogenic cells, α is the *intrinsic radiosensitivity* representing the nonrepairable radiation damage, β represents a repairable type of injury that is responsible for the *dose-per-fraction effect*, and n is the number of treatment fractions. The α/β ratio of the LQ-model then determines the tissue's sensitivity to alterations in radiation fraction size (Barendsen 1982).

The Relative Seriality s-model is the most well-known mechanistic normal tissue dose-response model that is based on the LQ-Poisson survival function (Källman, Ågren & Brahme 1992). For an OAR irradiated with a heterogeneous dose distribution the model is expressed as:

$$\mathrm{NTCP}_{\mathrm{RS}}(\mathbf{D}) = \left[1 - \prod_{k=1}^{N} \left[1 - \mathrm{P}_{\mathrm{LQ}}(D_k)^s\right]^{v_k}\right]^{\frac{1}{s}}, \tag{3}$$

where $\mathrm{P}_{\mathrm{LQ}}(D_k) = \exp(-N_0 \cdot \mathrm{SF}_{\mathrm{LQ}}(D_k))$ is the Poisson approximation of the binomial probability that no cells survive the dose D_k, and s is the relative seriality parameter that characterizes the internal organization of a tissue. A value of $s \approx 0$ represents a largely parallel organ (e.g., lung, parotid or liver), whereas $s \approx 1$ corresponds to a serial organ (e.g., spinal cord or esophagus).

Response for patient population with inter-tumoural variation of radiation sensitivity. The LQ-Poisson based TCP function of (2) describes the tumour response for an individual patient. However, the clinically measured dose response is a population-averaged dose response, which differs from the individual response due to inter-individual variability in radiobiological characteristics. A population-averaged TCP function, $\mathrm{TCP}_{\mathrm{pop}}$, can be constructed by averaging an individual TCP function, $\mathrm{TCP}_{\mathrm{ind}}$, over the range of parameters found in a population (see, e.g., Roberts & Hendry (1998)). Assuming that only the radiosensitivity α is subject to variations, the population TCP function can be expressed as

$$\mathrm{TCP}_{\mathrm{pop}}(\mathbf{D}) = \int_0^\infty \phi(\alpha)\mathrm{TCP}_{\mathrm{ind}}(\mathbf{D}, \alpha)\mathrm{d}\alpha, \tag{4}$$

where $\phi(\alpha)$ is the probability density function of the parameter α.

For $\phi(\alpha)$ often a normal probability density function has been used, assuming the variation of β to be correlated by a constant value of α/β (e.g., Sanchez-Nieto & Nahum (2000)). Others have applied the log-normal distribution to limit the parameter range of α to biologically meaningful values (Keall & Webb 2007). We analyse both the normal and log-normal averaged $\mathrm{TCP}_{\mathrm{pop}}$ function for log-concavity.

LQ-model-based equivalent uniform dose. The concept of *equivalent uniform dose* (EUD) was first introduced by Niemierko (1997) by equating the TCP of (2) to the TCP of an equivalent homogeneous distribution, i.e., $\mathrm{TCP}_{\mathrm{LQ}}(\mathbf{D}) \equiv \mathrm{TCP}(\mathrm{EUD}_{\mathrm{LQ}}(\mathbf{D}))$, and solving for $\mathrm{EUD}_{\mathrm{LQ}}(\mathbf{D})$. It is easy to show (see, e.g., McGary, Grant & Woo (2000)) that EUD under the LQ-Poisson model can be expressed as:

$$\mathrm{EUD}_{\mathrm{LQ}}(\mathbf{D}) = -\frac{1}{2}\frac{\alpha}{\beta}n\left[1 - \sqrt{1 - \frac{4\beta}{\alpha^2 n}\ln\overline{\mathrm{SF}}_{\mathrm{LQ}}(\mathbf{D})}\right], \tag{5}$$

with

$$\overline{\mathrm{SF}}_{\mathrm{LQ}}(\mathbf{D}) = \sum_{k=1}^{N} v_k \ \mathrm{e}^{-\alpha D_k - \beta {D_k}^2/n}.$$

Others have shown that the EUD under the (single-hit) linear-Poisson model is concave (Choi & Deasy 2002, Romeijn et al. 2004). We analyse the concavity of (5) using a direct approach and compare the result to (Bortfeld, Craft, Dempsey, Halabi & Romeijn 2008), which has been established under an indirect approach using Hardy, Littlewood & Pólya (1952, p. 88, Theorem 106).

2.2.2. Phenomenological dose-response relationship based on Probit model Assuming a normal distribution of intrinsic radiosensitivities, Lyman (1985) and Kutcher & Burman (1989) applied the Probit model to calculate the response of an OAR to a heterogeneous dose distribution according to:

$$\mathrm{NTCP}_{\mathrm{LKB}}(\mathbf{D}) = \Phi\left(\frac{\mathrm{gEUD}_a(\mathbf{D}) - D_{50}}{m\ D_{50}}\right), \tag{6}$$

where $\Phi(z) = 1/\sqrt{2\pi}\int_{-\infty}^{z}\mathrm{e}^{-t^2/2}\mathrm{d}t$ is the standard normal cumulative distribution function, D_{50} denotes the uniform dose where a 50% complication probability occurs, m determines the slope of the sigmoidal function Φ, and $\mathrm{gEUD}_a(\mathbf{D})$ is the *generalized equivalent uniform dose* (Niemierko 1999) of the dose distribution $\mathbf{D}$:

$$\mathrm{gEUD}_a(\mathbf{D}) = \left[\sum_{k=1}^{N} v_k\ {D_k}^a\right]^{\frac{1}{a}}, \tag{7}$$

and $a \geq 1$ is a tissue-dependent parameter that describes the volume dependence.

To account for the dose-per-fraction effect of a multi-fraction treatment, the dose to each voxel is converted into an iso-effective *biologically effective dose* (BED) (Fowler 1989) prior to calculation of the gEUD:

$$\mathrm{BED}(D_i) = D_i\left(1 + \frac{D_i/n}{\alpha/\beta}\right).$$

Subsequent calculation of the gEUD from the BED-converted dose distribution yields the *generalized equivalent uniform biologically effective dose* (gEUBED) function:

$$\begin{aligned}\mathrm{gEUBED}_a(\mathbf{D}) &= \mathrm{gEUD}_a(\mathbf{BED}(\mathbf{D})) \\ &= \left[\sum_{k=1}^{N} v_k \ \mathrm{BED}(D_k)^a\right]^{\frac{1}{a}}. \end{aligned} \tag{8}$$

The fractionation-corrected LKB model then becomes:

$$\mathrm{NTCP}_{\mathrm{LKB}}(\mathbf{D}) = \Phi\left(\frac{\mathrm{gEUBED}_a(\mathbf{D}) - \mathrm{BED}_{50}}{m\ \mathrm{BED}_{50}}\right), \tag{9}$$

where $\mathrm{BED}_{50} = \mathrm{BED}(D_{50})$.

2.2.3. Composite evaluation functions For multiple targets and vital organs, the overall probability of benefit and injury is respectively given by:

$$\mathrm{TCP}(\mathbf{D}) = \prod_{j=1}^{T} \mathrm{TCP}_j(\mathbf{D})^{\eta_j}, \tag{10}$$

and

$$\mathrm{NTCP}(\mathbf{D}) = 1 - \prod_{j=1}^{S} [1 - \mathrm{NTCP}_j(\mathbf{D})]^{\xi_j}, \tag{11}$$

where T and S are the number of targets and vital organs, respectively, $\mathrm{TCP}_j(\mathbf{D})$ is the control probability for target j, η_j is the fractional volume of the jth target, $\mathrm{NTCP}_j(\mathbf{D})$ is the complication probability in organ j, and ξ_j is the relative weight of the jth complication, such that $\sum_j^T \eta_j = \sum_j^S \xi_j = 1$. Here, $\mathrm{TCP}_j(\mathbf{D})$ and $\mathrm{NTCP}_j(\mathbf{D})$ can be any of the fractionation-corrected dose-response models from the preceding subsections.

Another composite objective function is the probability of uncomplicated tumour control, P_+, which is a measure to *a priori* balance between treatment benefit and injuries (see, e.g., Brahme (1995)). In generalized form it is defined as

$$P_+^\delta(\mathbf{D}) = (1-\delta)P_+^0(\mathbf{D}) + \delta P_+^1(\mathbf{D}),$$

where $P_+^0(\mathbf{D}) \equiv \mathrm{TCP}(\mathbf{D}) - \mathrm{NTCP}(\mathbf{D})$, $P_+^1(\mathbf{D}) \equiv \mathrm{TCP}(\mathbf{D})[1 - \mathrm{NTCP}(\mathbf{D})]$ and δ is the fraction of patients for which tumour and normal tissue responses are statistically independent ($0 \leq \delta \leq 1$). $\delta = 0$ corresponds to the case that the events of tumour control and severe injury are totally correlated (i.e., no complications without achieving tumour control). For $\delta = 1$ it is assumed that tumour control is independent of complications. Note that $P_+^\delta(\mathbf{D})$ is a convex combination of $P_+^0(\mathbf{D})$ and $P_+^1(\mathbf{D})$.

3. Methods

3.1. Convex multi-criteria optimisation

A general convex multi-criteria optimisation problem is formulated in standard form as:

$$\begin{array}{ll} \min\limits_{\mathbf{x}} & \mathbf{G}(\mathbf{x}) = \begin{pmatrix} G_1(\mathbf{x}) \\ G_2(\mathbf{x}) \\ \vdots \\ G_K(\mathbf{x}) \end{pmatrix} \\ \text{s.t.} & C_l(\mathbf{x}) \leq 0, \qquad l = 1, \ldots, L \\ & E_l(\mathbf{x}) = 0, \qquad l \in \mathcal{E}. \end{array} \tag{12}$$

Here $\mathbf{x} \in \mathbb{R}^M$ is the optimisation variable, $G_k : \mathbb{R}^M \mapsto \mathbb{R}$ are convex *objective functions*, $C_l : \mathbb{R}^M \mapsto \mathbb{R}$ are convex *inequality constraint functions*, $E_i : \mathbb{R}^M \mapsto \mathbb{R}$ are affine *equality constraint functions*, and $\mathcal{E}$ is the set of indices for equality constraints.

In case of the FMO problem in IMRT (1), the optimisation variable $\mathbf{x}$ is obtained from the dose deposition operator by a linear mapping of the design variable $\mathbf{w}$, which we control to generate fluence maps. If the only constraint on the treatment plan is the nonnegativity of bixel weights, then the set $\mathcal{A}$ of allowed bixel weights is convex. Therefore, the set $\mathcal{R}$ of realizable dose distributions (being the image of $\mathcal{A}$ under the linear dose deposition operator $\mathbf{P}$) is convex. For the multi-criteria FMO problem to be cast as a convex problem (12), the objective and constraint functions of (1) should all be convex or transformed into equivalent convex functions.

Equivalence of optimisation problem for transformed criterion functions. Direct application of the (nonconvex) criterion functions as objective and constraint functions would result in a nonconvex problem. Romeijn et al. (2004) have shown that transforming any or all criterion functions of the general (nonconvex) minimisation problem (1) via strictly increasing functions leads to an equivalent Pareto efficient frontier. This property can be used to convexify problem (1) such that it can be formulated as a convex problem (12). A sufficient condition is that strictly increasing transformations $h_k : \mathbb{R} \mapsto \mathbb{R}$ $(k = 1, \ldots, K)$ and $\tilde{h}_l : \mathbb{R} \mapsto \mathbb{R}$ $(l = 1, \ldots, L)$ exist, such that compositions

$$G_k(\mathbf{x}) = (h_k \circ F_k)(\mathbf{x}) \equiv h_k(F_k(\mathbf{x}))$$

and

$$C_l(\mathbf{x}) = (\tilde{h}_l \circ B_l)(\mathbf{x}) \equiv \tilde{h}_l(B_l(\mathbf{x}))$$

are convex functions.

In the next section, we find transformations for the radiobiological criterion functions described in Section 2.2 and analyse the conditions under which the transformed criteria are (strictly) convex.

Although the criterion functions have been cast as a function of the dose distribution $\mathbf{D}$, rather than as a function of the bixel weights $\mathbf{w}$, the analysis is conducted in terms of $\mathbf{D}$. Formally, the objective functions $G_k(\mathbf{D})$ and $G_k(\mathbf{w}) = G_k(\mathbf{D}(\mathbf{w})) = G_k(\mathbf{P} \cdot \mathbf{w})$ are different, but since $\mathbf{D}$ is a linear mapping of $\mathbf{w}$, the convexity properties of the functions remain invariant. This is due to the fact that the composition of a convex function with an affine mapping is convex (see, e.g., Boyd & Vandenberghe (2004, p. 79)). However, for G_k to be strictly convex with respect to

$\mathbf{w}$, the mapping $\mathbf{P}$ should be full rank (and thus have an empty nullspace). This can be derived from the Hessian of G_k with respect to $\mathbf{w}$:

$$\nabla^2_{ww} G_k(\mathbf{D}(\mathbf{w})) = \mathbf{P}^T \cdot \nabla^2_{DD} G_k(\mathbf{D}) \cdot \mathbf{P},$$

and using proper statements of positive semidefiniteness from linear algebra. Using the properties of positive (semi)definite symmetric products (see, e.g., Horn & Johnson (1987, p. 399)) it can be shown that if $\nabla^2_{DD} G_k(\mathbf{D})$ is positive definite (and thus G_k is strictly convex in $\mathbf{D}$), and $\mathbf{P}$ is any particular dimensionally compatible matrix that has an empty nullspace, then $\nabla^2_{ww} G_k(\mathbf{D}(\mathbf{w}))$ is positive definite (and thus G_k is strictly convex in $\mathbf{w}$). It also follows that, if $\mathbf{P}$ is of compatible dimension but rank-deficient, then $\nabla^2_{ww} G_k(\mathbf{D}(\mathbf{w}))$ is positive semidefinite (and thus G_k is convex in $\mathbf{w}$). Clearly, this also holds for the constraint functions C_l. In practice, the dose grid is finer than the bixel grid, i.e., $M \gg N$ and thus $\mathbf{P}$ has full column rank (Carlsson 2008).

3.2. Convexity analysis of transformed radiobiological criterion functions

In general, convexity of a criterion function $G : \mathbb{R}^M \mapsto \mathbb{R}$ can be analysed via different approaches. A direct approach is to check the conditions on its Hessian, $\nabla^2 G$. A necessary and sufficient condition for convexity of G is that its Hessian is positive semidefinite. Without proof, we state the conditions for convexity and concavity more formally:

Definition 1. *A matrix $\mathbf{A} \in \mathbb{R}^{M \times M}$ is positive (negative) semidefinite on a set $\mathcal{S}$ if $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq (\leq)\ 0$ for all $\mathbf{x} \in \mathcal{S}$, and positive (negative) definite on $\mathcal{S}$ if $\mathbf{x}^T \mathbf{A} \mathbf{x} > (<)\ 0$ for $\mathbf{x} \neq 0$ and $\mathbf{x} \in \mathcal{S}$.*

Lemma 1. *A real-valued, twice differentiable function G is convex (concave) on a set $\mathcal{S}$ if and only if its Hessian, $\nabla^2 G$, is positive (negative) semidefinite on $\mathcal{S}$. The function G is strictly convex (concave) on the set $\mathcal{S}$ if and only if $\nabla^2 G$ is positive (negative) definite on $\mathcal{S}$.*

An indirect approach is to exploit the rules of *composition* to deduce convexity/concavity of differentiable functions by means of the chain rule. The following composition rules for convexity and concavity can be derived (see, e.g. Boyd & Vandenberghe (2004, p. 83−7)).

Lemma 2. *Consider the functions $\mathbf{F} : \mathbb{R}^p \mapsto \mathbb{R}^q$ and $h : \mathbb{R}^q \mapsto \mathbb{R}$ with their composition $G = h \circ \mathbf{F} : \mathbb{R}^p \mapsto \mathbb{R}$, defined by*

$$G(\mathbf{x}) = h(\mathbf{F}(\mathbf{x})) = h(F_1(\mathbf{x}), \ldots, F_q(\mathbf{x})),$$

where $F_i : \mathbb{R}^p \mapsto \mathbb{R}$, $p \geq 1$ and $q \geq 1$. Then,

> *G is convex if h is convex and nondecreasing (nonincreasing) in each argument, and F_i are convex (concave),*
>
> *G is concave if h is concave and nondecreasing (nonincreasing) in each argument, and F_i are concave (convex).*

Depending on the mathematical form of the function, a transformation of the radiobiological criteria functions presented in Section 2.2 is suggested, and one of the two approaches is applied to analyse the convexity of the transformed criterion functions.

4. Results

4.1. TCP function using the LQ-Poisson cell survival model

The concavity of the logarithmically transformed linear-Poisson model-based TCP function has been investigated by others (Choi & Deasy 2002, Romeijn et al. 2004). These authors have applied the strictly increasing transformation $h(z) = \ln z$ to show that the linear-Poisson-based TCP function is log-concave.

Here, the same transformation has been applied to the LQ-Poisson-based TCP function of (2). In Appendix A we show that $-\ln \mathrm{TCP}_{\mathrm{LQ}}(\mathbf{D})$ is strictly convex under the condition that

$$D_k > \sqrt{\frac{1}{2}\frac{(\alpha/\beta)n}{\alpha}} - \frac{1}{2}(\alpha/\beta)n. \tag{13}$$

Since dose is a nonnegative quantity, constraint (13) is satisfied if $\alpha^2 n > 2\beta$. In the opposite case, where $\alpha^2 n < 2\beta$, $\mathrm{SF}_{\mathrm{LQ}}(D)$ is sigmoidal for $D \geq 0$, and its curvature is only strictly positive if (13) is satisfied. For typical clinical values of α, α/β and n, (e.g., for prostate cancer, $\alpha = 0.15 \pm 0.04$ Gy^{-1}, $\alpha/\beta = 3.1 \pm 0.5$ Gy, $n = 30$, (Wang, Guerrero & Li 2003)) the righthand side of (13) is negative, and therefore $-\ln TCP_{\mathrm{LQ}}(\mathbf{D})$ is a strictly convex function. Consequently, $\ln \mathrm{TCP}_{\mathrm{LQ}}(\mathbf{D})$ is a strictly concave function.

4.2. TCP function using the linear-Poisson cell survival model

In case of either a large α/β ratio (i.e., $\beta \approx 0$) or small dose inhomogeneities and a dose-per-fraction close to the reference dose for which the underlying dose-response model was derived, the LQ-Poisson-based TCP function of (2) reduces to the linear-Poisson-based TCP model:

$$\mathrm{TCP}_{\mathrm{L}}(\mathbf{D}) = \exp\left[-N_0 \sum_{k=1}^{N} v_k \ \mathrm{e}^{-\alpha D_k}\right]. \tag{14}$$

Others have shown that $\mathrm{TCP}_{\mathrm{L}}(\mathbf{D})$ is a log-concave function (Choi & Deasy 2002, Romeijn et al. 2004).

In Appendix B we apply a concave and strictly increasing transformation of the (already) convex function $-\ln \mathrm{TCP}_{\mathrm{L}}(\mathbf{D})$ and show that it is still convex. In particular, we conclude that the function $\ln(-\ln \mathrm{TCP}_{\mathrm{L}}(\mathbf{D}))$ is 'less convex' than $-\ln \mathrm{TCP}_{\mathrm{L}}(\mathbf{D})$. Hence, the former double-logarithmic transformation will produce tighter bounds in the piecewise linear approximation of the Pareto efficient frontier than the single-logarithmic transformation of $\mathrm{TCP}_{\mathrm{L}}(\mathbf{D})$ (Siem, den Hertog & Hoffmann 2008).

4.3. Population-averaged TCP function using the LQ-Poisson cell survival model

Assuming that the inter-patient intrinsic radiosensitivity of a certain tumour type is normally distributed over the population, we show in Appendix C that $-\ln \mathrm{TCP}_{\mathrm{pop}}(\mathbf{D})$ is strictly convex for the LQ-Poisson-based TCP function, provided that the dose in each voxel satisfies the condition

$$D_k > \sqrt{\left[\frac{1}{2}(\alpha/\beta)n\right]^2 + \frac{1}{2}\frac{(\alpha/\beta)n}{\alpha}} - \frac{1}{2}(\alpha/\beta)n. \tag{15}$$

For typical clinical values of α, α/β and n, the righthand side of (15) is a small positive value, and thus the inequality is easily satisfied for tumours.

In case the inter-patient radiosensitivity has a log-normal probability density function, $\phi(\alpha)$, the population-averaged LQ-Poisson-based TCP function is not necessarily logarithmic concave since $\phi(\alpha)$ is not logarithmic concave (Bagnoli & Bergstrom 2005).

4.4. NTCP function using the Relative Seriality s-model

For both the (nonfractionation-corrected) gEUD-based reformulation of Lyman's phenomenological NTCP model (6) and the gEUD-based phenomenological NTCP function due to Alber & Nusslin (2001), Romeijn et al. (2004) have proven convexity under the strictly increasing transformation $h(z) = -\ln(1-z)$.

Here, we apply the same transformation to the Relative Seriality s-model of (3), and analyse its convexity properties. Since the Relative Seriality s-model can neither be formulated as a function of EUD nor as a function of gEUD, application of the model in multi-criteria optimisation of IMRT will yield a Pareto frontier that is different from the two aforementioned gEUD-based NTCP models.

The result comprises two steps (see Appendix D). In Theorem 4 it is shown that applying the strictly increasing transformation $h(z) = -\ln(1-z^s)$ to (3) yields a strictly convex function for $s \geq 0$. In Theorem 5, this result is used to prove that $-\ln(1-\mathrm{NTCP}_{\mathrm{RS}}(\mathbf{D}))$ is a strictly convex function, provided that $0 < s < 1$. For most normal tissues this condition is satisfied.

Note that $-\ln(1-\mathrm{NTCP}_{\mathrm{RS}}(\mathbf{D}))$ is a 'more convex' transformation of $\mathrm{NTCP}_{\mathrm{RS}}(\mathbf{D})$ than $-\ln(1-\mathrm{NTCP}_{\mathrm{RS}}(\mathbf{D})^s)$. Therefore, the latter transformation is more efficient for approximation of the Pareto efficient frontier than the former (Siem et al. 2008). Nevertheless, the convexity of $-\ln(1-\mathrm{NTCP}_{\mathrm{RS}}(\mathbf{D}))$ is used in the convexity analysis of the composite NTCP function (11).

4.5. EUD function using the LQ-Poisson cell survival model

Note that $\mathrm{EUD}_{\mathrm{LQ}}(\mathbf{D})$ is concave if and only if its Hessian is negative semidefinite. Unfortunately, the formulas for the second partial derivatives become so complex, that we were unable to analytically derive conditions under which the Hessian is negative semidefinite. Nevertheless, numerical evaluation of the Hessian for the two-dimensional case $\mathbf{D} = (50\ \mathrm{Gy}, 40\ \mathrm{Gy})$ with equally-sized voxel volumes and realistic clinical values $\alpha = 0.30\ \mathrm{Gy}^{-1}, \alpha/\beta = 10\ \mathrm{Gy}, n = 30$ showed that the determinant of the 2×2 Hessian is negative, which implies that the Hessian is indefinite. Hence, $\mathrm{EUD}_{\mathrm{LQ}}(\mathbf{D})$ is neither convex nor concave in this point. The same result is obtained for other values of $\mathbf{D}$. This contradicts the results of Bortfeld et al. (2008), although their concavity criterion $\alpha^2 n > 2\beta$ is fulfilled. Our counter example requires the criterion to be re-evaluated.

4.6. gEUBED function for normal tissues using the LQ-model

By noting that $\mathrm{gEUBED}_a(\mathbf{D})$ can be obtained from (7) via a simple vector transformation, we prove that $\mathrm{gEUBED}_a(\mathbf{D})$ is convex for $a \geq 1$ in Theorem 6 of Appendix E.

4.7. NTCP function using the fractionation-corrected LKB model

In a similar way Romeijn et al. (2004) applied the transformation $h(z) = -\ln(1-z)$ to the (nonfractionation-corrected) gEUD-based LKB model of (6), we apply it to the (fractionation-corrected) gEUBED-based LKB model of (9). Using the result of Section 4.6, we show in Theorem 7 of Appendix E that $-\ln(1-\mathrm{NTCP}_{\mathrm{LKB}}(\mathbf{D}))$ is a convex function.

4.8. Composite TCP and NTCP functions

In Appendix F it is shown that $-\ln \mathrm{TCP}(\mathbf{D})$ and $-\ln(1-\mathrm{NTCP}(\mathbf{D}))$ are convex functions for the relevant fractionation-corrected dose-response models presented in Section 2.2. This extends the results obtained by Romeijn et al. (2004) to include TCP and NTCP models that incorporate fractionation effects.

4.9. Uncomplicated tumour control probability

Without proof, we echo the notion of Romeijn et al. (2004) that from Theorems 8 and 9 (see Appendix F) it immediately follows that $-\ln P_+^1(\mathbf{D})$ is a strictly convex function, provided that the convexity conditions of the constituent TCP and NTCP functions are satisfied. The analysis gets complicated when the transformation $h(z) = -\ln z$ is applied to $P_+^0(\mathbf{D})$. No guarantee for the convexity of $-\ln P_+^0(\mathbf{D})$ can be derived.

Whether this result also rules out convexity of $-\ln P_+^\delta(\mathbf{D})$ for $0 < \delta < 1$ remains to be shown. Little impact of the δ value on $P_+^\delta(\mathbf{D})$ has been reported (Kim & Tomé 2006). Another argument would be that $P_+^1(\mathbf{D}) - P_+^0(\mathbf{D}) = \mathrm{NTCP}(\mathbf{D})[1 - \mathrm{TCP}(\mathbf{D})]$ is small in the optimum. Therefore, it seems reasonable to use $P_+^1(\mathbf{D})$ instead of $P_+^\delta(\mathbf{D})$, since the former criterion is guaranteed to be log-concave.

Nevertheless, for biologically-based multi-criteria optimisation the more elementary composite TCP and NTCP functions are preferred, since they do not impose an *a priori* balance between them.

5. Discussion and conclusions

We have identified transformations of commonly used nonconvex radiobiological treatment plan evaluation criteria that account for the dose-per-fraction effect and analysed conditions under which the transformed criteria are convex functions that do not change the set of Pareto efficient treatment plans of a multi-criteria FMO problem. The dose-per-fraction effect is accounted for by exploiting the well-known log-linear quadratic cell survival model that has often been applied in fractionated radiotherapy. By using the transformations, we were able to reformulate clinically relevant multi-criteria optimization problems being cast in terms of TCP and NTCP endpoints as convex problems.

In contrast to the work conducted by Romeijn et al. (2004), our analysis comprised radiobiological criteria that cannot be cast as a transformation of the elementary EUD or gEUD functions. This implies that the treatment plan evaluation criteria analysed in the current work are truly different from those that were analysed before. Therefore, the criteria analysed will yield different Pareto efficient frontiers.

We have shown that the convexity conditions of transformed criterion functions that are based on the mechanistic LQ-Poisson based dose-response model are directly related to the form of the log-linear-quadratic cell survival function underlying it. In

this respect, it should be mentioned that the log-linear cell survival model underlying the criterion functions analysed by Romeijn et al. (2004) is a limit of the log-linear quadratic model for either a large α/β ratio or for small dose inhomogeneities and a dose-per-fraction close to the reference dose for which the dose-response model was derived. Although Bortfeld et al. (2008) have claimed the LQ-Poisson-based EUD function to be concave under certain conditions, we have provided a numerical counter example that shows that this function is neither convex nor concave.

Another aspect that was introduced as an extension to the work of Romeijn et al. (2004) is the type of transformation applied. Apart from being strictly increasing, we have noticed that transformations of different quality exist. Certain transformations exist that yield transformed criterion functions that are 'less convex' than other transformations. This is particularly useful for the approximation of the Pareto efficient frontier by piecewise linear upper and lower bounds (Siem et al. 2008). These authors have shown that if the criterion function is already convex, an increasing and concave transformation yields a less convex function, for which tighter upper and lower bounds can be obtained than for the original bounds of the Pareto efficient frontier that is to be approximated. We note that for the transformed criteria analysed by Romeijn et al. (2004) increasing and concave transformations may exist that are more efficient.

Future research will address the application of convex biologically-based multi-criteria optimisation to assess the effect of different radiobiological models on the Pareto efficient frontier and associated dose distributions for a given treatment technique (i.e., beam modality, beam energy, beam configuration, and machine characteristics).

Acknowledgments

The authors would like to thank Kjell Eriksson and Fredrik Carlsson from RaySearch Laboratories AB, Stockholm, for the valuable discussions. Furthermore, we are indebted to Lieven Vandenberghe from the University of California, Los Angeles, for his help on convexity properties of sigmoidal functions. We also thank the referees for the useful comments that helped to improve this paper.

References

Alber, M., Meedt, G., Nusslin, F. & Reemtsen, R. (2002). On the degeneracy of the IMRT optimization problem, *Med Phys* **29**(11): 2584–9.

Alber, M. & Nusslin, F. (2001). A representation of an NTCP function for local complication mechanisms, *Phys Med Biol* **46**(2): 439–47.

Bagnoli, M. & Bergstrom, T. (2005). Log-concave probability and its apllication, *Econ Theor* **26**: 445–69.

Barendsen, G. W. (1982). Dose fractionation, dose rate and iso-effect relationships for normal tissue responses, *Int J Radiat Oncol Biol Phys* **8**(11): 1981–97.

Bertsekas, D. P. (1999). *Nonlinear Programming*, 2nd edn, Athena Scientific, Nashua.

Börgers, C. (1997). The radiation therapy planning problem, *in* C. Börgers & F. Natterer (eds), *Computational Radiology and Imaging: Therapy and Diagnostics*, The IMA Volumes in Mathematics and its Applications, Springer. Proceedings of IMA Workshop, March 1997.

Bortfeld, T., Craft, D., Dempsey, J. F., Halabi, T. & Romeijn, H. E. (2008). Evaluating target cold spots by the use of tail EUDs, *Int J Radiat Oncol Biol Phys* **71**(3): 880–9.

Boyd, S. & Vandenberghe, L. (2004). *Convex Optimization*, Cambridge University Press, Cambridge.

Brahme, A. (1995). Treatment optimization using physical and radiobiological objective functions, *in* A. R. Smith (ed.), *Radiation Therapy Physics*, Springer, Berlin, pp. 209–46.

Carlsson, F. (2008). *Utilizing problem structure in optimization of radiation therapy*, Doctoral thesis, Royal Institute of Technology. ISBN 978-91-7178-920-4.

Choi, B. & Deasy, J. O. (2002). The generalized equivalent uniform dose function as a basis for intensity-modulated treatment planning, *Phys Med Biol* **47**(20): 3579–89.

Craft, D. L., Halabi, T. F., Shih, H. A. & Bortfeld, T. R. (2006). Approximating convex Pareto surfaces in multiobjective radiotherapy planning, *Med Phys* **33**(9): 3399–407.

Deasy, J. O. (1997). Multiple local minima in radiotherapy optimization problems with dose-volume constraints, *Med Phys* **24**(7): 1157–61.

Fowler, J. F. (1989). The linear-quadratic formula and progress in fractionated radiothyerapy, *Br J Radiol* **62**: 679–94.

Hardy, G. H., Littlewood, J. E. & Pólya, G. (1952). *Inequalities*, Cambridge University Press, Cambridge, England.

Hindi, H. (2004). A tutorial on convex optimization, *Proceedings of the American Control Conference*, Vol. 4, IEEE, Boston, pp. 3252–65.

Hoffmann, A. L., Siem, A. Y., den Hertog, D., Kaanders, J. H. & Huizenga, H. (2006). Derivative-free generation and interpolation of convex pareto optimal IMRT plans, *Phys Med Biol* **51**(24): 6349–69.

Horn, R. A. & Johnson, C. R. (1987). *Matrix Analysis*, Cambridge University Press, Cambridge.

Källman, P., Ågren, A. & Brahme, A. (1992). Tumour and normal tissue responses to fractionated non-uniform dose delivery, *Int J Radiat Biol* **62**(2): 249–62.

Keall, P. J. & Webb, S. (2007). Optimum parameters in a model for tumour control probability, including interpatient heterogeneity: evaluation of the log-normal distribution, *Phys Med Biol* **52**(1): 291–302.

Kim, Y. & Tomé, W. A. (2006). Risk-adaptive optimization: selective boosting of high-risk tumor subvolumes, *Int J Radiat Oncol Biol Phys* **66**(5): 1528–42.

Küfer, K.-H., Scherrer, A., Monz, M., Alonso, F., Trinkaus, H., Bortfeld, T. & Thieke, C. (2003). Intensity-modulated radiotherapy - a large scale multi-criteria programming problem, *OR Spectrum* **25**: 223–49.

Kutcher, G. J. & 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**(6): 1623–30.

Lind, B. (1990). Properties of an algorithm for solving the inverse problem in radiation therapy, *Inverse Probl* **6**(3): 415–26.

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

McGary, J. E., Grant, W. & Woo, S. Y. (2000). Applying the equivalent uniform dose formulation based on the linear-quadratic model to inhomogeneous tumor dose distributions: Caution for analyzing and reporting, *J Appl Clin Med Phys* **1**(4): 126–37.

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

Nesterov, Y. & Nemirovski, A. (1994). Interior-point polynomial methods in convex programming, *SIAM Series in Applied Mathematics*, Vol. 13, Society for Industrial and Applied Mathematics, Philadelphia.

Niemierko, A. (1997). Reporting and analyzing dose distributions: a concept of equivalent uniform dose, *Med Phys* **24**(1): 103–10.

Niemierko, A. (1999). A generalized concept of equivalent uniform dose, *Med Phys* **26**: 1100. Abstract.

Niemierko, A. (2005). Biological optimization, *in* T. Bortfeld, R. Schmidt-Ullrich, W. De Neve & D. E. Wazer (eds), *Image-guided IMRT*, Springer, Berlin, pp. 199–216.

Prékopa, A. (1973). On logarithmic concave measures and functions, *Acta Scientiarum Mathematicarum* **34**: 335–43.

Reemtsen, R. & Alber, M. (2006). Continuous optimization of beamlet intensities for photon and proton radiotherapy. http://www.math.tu-cottbus.de/INSTITUT/lsing1/publications/SurveyRT0306.ps.

Roberts, S. A. & Hendry, J. H. (1998). A realistic closed-form radiobiological model of clinical tumor-control data incorporating intertumor heterogeneity, *Int J Radiat Oncol Biol Phys* **41**(3): 689–99.

Romeijn, H. E., Dempsey, J. F. & Li, J. G. (2004). A unifying framework for multi-criteria fluence map optimization models, *Phys Med Biol* **49**(10): 1991–2013.

Sanchez-Nieto, B. & Nahum, A. E. (2000). BIOPLAN: software for the biological evaluation of radiotherapy treatment plans, *Med Dosim* **25**(2): 71–6.

Siem, A. Y., den Hertog, D. & Hoffmann, A. L. (2008). The effect of transformations on the

approximation of univariate (convex) functions with applications to Pareto curves, *Eur J Oper Res* **189**(2): 347–62.

Wang, J. Z., Guerrero, M. & Li, X. A. (2003). How low is the α/β ratio for prostate cancer?, *Int J Radiat Oncol Biol Phys* **55**(1): 194–203.

Appendix A.

The following theorem shows that the log-transformed LQ-Poisson-based TCP function is concave under certain conditions.

Theorem 1. *Consider the function $TCP_{LQ}(\mathbf{D})$ of (2). Let $h(z) = \ln z$, which is a strictly increasing function. Then the composition*

$$-h(TCP_{LQ}(\mathbf{D})) = -\ln TCP_{LQ}(\mathbf{D}) = N_0 \sum_{k=1}^{N} v_k \; SF_{LQ}(D_k) \tag{A.1}$$

is a strictly convex function on the set of dose distributions

$$\{\mathbf{D} \mid D_k > \sqrt{\frac{1}{2}\frac{(\alpha/\beta)n}{\alpha}} - \frac{1}{2}(\alpha/\beta)n, \quad k = 1, \ldots, M\}.$$

Proof. Since the nonnegative sum of strictly convex functions is strictly convex, it suffices to prove that $\mathrm{SF}_{\mathrm{LQ}}(D)$ is strictly convex for $-\ln \mathrm{TCP}_{\mathrm{LQ}}(\mathbf{D})$ to be strictly convex. The curvature of $\mathrm{SF}_{\mathrm{LQ}}(D)$ is the second derivative with respect to D:

$$\mathrm{SF}''_{\mathrm{LQ}}(D) = \left[(\alpha + 2\beta D/n)^2 - 2\beta/n\right] \mathrm{SF}_{\mathrm{LQ}}(D). \tag{A.2}$$

From (A.2), it is easy to see that the curvature is strictly positive if

$$D > \sqrt{\frac{1}{2}\frac{(\alpha/\beta)n}{\alpha}} - \frac{1}{2}(\alpha/\beta)n.$$

□

Appendix B.

The following theorem shows that $\ln(-\ln \mathrm{TCP}_{\mathrm{L}}(\mathbf{D}))$ is a convex function of $\mathbf{D}$.

Theorem 2. *Consider the function $-\ln TCP_{L}(\mathbf{D})$ with $TCP_{L}(\mathbf{D})$ according to (14). Let $h(z) = \ln z$, which is a concave and strictly increasing function. Then the composition*

$$h(-\ln TCP_{L}(\mathbf{D})) = \ln(-\ln TCP_{L}(\mathbf{D})) = \ln N_0 + \ln\left(\sum_{k=1}^{N} v_k \; e^{-\alpha D_k}\right) \tag{B.1}$$

is a convex function on the set of dose distributions $\{\mathbf{D} \mid D_k \geq 0, k = 1, \ldots, M\}$.

Proof. It is easy to verify that $\mathrm{e}^{-\alpha D}$ is a log-convex function of D. Since the sum of log-convex functions is log-convex, $\ln\left(\sum_{k=1}^{N} v_k \, \mathrm{e}^{-\alpha D_k}\right)$ is convex (Boyd & Vandenberghe 2004, p. 105). Hence, $\ln(-\ln \mathrm{TCP}_{\mathrm{L}}(\mathbf{D}))$ is a convex function of $\mathbf{D}$. □

Appendix C.

The following theorem shows that in case the inter-patient distribution of intrinsic radiosensitivities is Gaussian the log-transformed population-averaged LQ-Poisson-based TCP function is concave under certain conditions.

Theorem 3. *Consider the function $TCP_{pop}(\mathbf{D})$ of (4). Let $TCP_{ind}(\mathbf{D},\alpha) = TCP_{LQ}(\mathbf{D},\alpha)$ and let $\phi(\alpha)$ be the standard normal probability density function. Furthermore, let β be correlated to α by a constant value of α/β. Let $h(z) = \ln z$, which is a strictly increasing function. Then the composition*

$$h(\,TCP_{pop}(\mathbf{D})) = \ln\, TCP_{pop}(\mathbf{D}) = \ln\left[\int_0^\infty \phi(\alpha)\, TCP_{ind}(\mathbf{D},\alpha)\, d\alpha\right],$$

is a strictly concave function on the set of dose distributions

$$\{\mathbf{D} \mid D_k > \sqrt{\left[\frac{1}{2}(\alpha/\beta)n\right]^2 + \frac{1}{2}\frac{(\alpha/\beta)n}{\alpha}} - \frac{1}{2}(\alpha/\beta)n, \quad k = 1,\ldots,M\}.$$

Proof. We use Theorem 6 from Prékopa (1973), which states that the function

$$F(\mathbf{D}) = \int_{I_\alpha} f(\mathbf{D},\alpha)\mathrm{d}\alpha,$$

with $I_\alpha \subseteq \mathbb{R}$ a convex set, is logarithmic concave in $\mathbf{D}$ if $f(\mathbf{D},\alpha)$ is a logarithmic concave function of $\mathbf{D}$ and α. In our case, $f(\mathbf{D},\alpha) = \phi(\alpha)\mathrm{TCP}_{\mathrm{ind}}(\mathbf{D},\alpha)$ and $I_\alpha = \mathbb{R}_+$. Note that

$$\ln[\phi(\alpha)\mathrm{TCP}_{\mathrm{ind}}(\mathbf{D},\alpha)] = \ln\phi(\alpha) + \ln\mathrm{TCP}_{\mathrm{ind}}(\mathbf{D},\alpha).$$

It is easy to see that $\phi(\alpha)$ is strictly logarithmic concave (see, e.g., Bagnoli & Bergstrom (2005)). Since the nonnegative sum of strictly concave functions is strictly concave, it remains to prove that $\ln\mathrm{TCP}_{\mathrm{ind}}(\mathbf{D},\alpha)$ is a strictly concave function of $\mathbf{D}$ and α. Note that since

$$-\ln\mathrm{TCP}_{\mathrm{ind}}(\mathbf{D},\alpha) = N_0 \sum_{k=1}^{N} v_k\ \mathrm{SF}_{\mathrm{LQ}}(D_k,\alpha),$$

is separable in the doses to individual voxels, it suffices to show that $\mathrm{SF}_{\mathrm{LQ}}(D,\alpha)$ is a strictly convex function of D and α. Therefore, we analyse the Hessian of $\mathrm{SF}_{\mathrm{LQ}}(D,\alpha)$:

$$\nabla^2\mathrm{SF}_{\mathrm{LQ}}(D,\alpha) = \mathrm{SF}_{\mathrm{LQ}}(D,\alpha)\begin{pmatrix} a & b \\ b & c \end{pmatrix}, \tag{C.1}$$

with $a = (\alpha + 2\beta D/n)^2 - 2\beta/n$, $b = 2\beta D^2/n + \alpha D - 1$ and $c = D^2$. The Hessian (C.1) is positive definite if and only if $a > 0$ and its determinant, $ac - b^2$, is positive. It is easy to show that these two conditions are satisfied if

$$D > \sqrt{\left[\frac{1}{2}(\alpha/\beta)n\right]^2 + \frac{1}{2}\frac{(\alpha/\beta)n}{\alpha}} - \frac{1}{2}(\alpha/\beta)n.$$

□

Appendix D.

The following theorem shows that $-\ln(1-\mathrm{NTCP}_{\mathrm{RS}}(\mathbf{D})^s)$ is a strictly convex function on the set of physically meaningful dose distributions.

Theorem 4. *Consider the function $NTCP_{RS}(\mathbf{D})$ of (3) with $s \geq 0$. Let $h(z) = -\ln(1-z^s)$, which is a strictly increasing function. Then the composition*

$$\begin{aligned} h(NTCP_{RS}(\mathbf{D})) &= -\ln(1-NTCP_{RS}(\mathbf{D})^s) \\ &= -\sum_{k=1}^{N} v_k \ \ln(1-P_{LQ}(D_k)^s) \end{aligned} \tag{D.1}$$

is a strictly convex function on the set of dose distributions $\{\mathbf{D} \mid D_k \geq 0, k = 1,\ldots,M\}$.

Proof. Note that $h(\mathrm{NTCP}_{\mathrm{RS}}(\mathbf{D}))$ is separable in doses to individual voxels. Hence, it suffices to show that $g(D) = -\ln(1-\mathrm{P}_{\mathrm{LQ}}(D)^s)$ is a strictly convex function of D. By calculating the second derivative of g with respect to D, it can be shown that $g''(D)$ is strictly positive if

$$(\alpha + 2\beta D/n)^2 \left(\frac{sN_0\mathrm{SF}_{\mathrm{LQ}}(D)}{1-\mathrm{P}_{\mathrm{LQ}}(D)^s} - 1 \right) + 2\beta/n > 0. \tag{D.2}$$

Note that the lefthand side of (D.2) can only be nonpositive if

$$sN_0\mathrm{SF}_{\mathrm{LQ}}(D) < 1 - \mathrm{P}_{\mathrm{LQ}}(D)^s. \tag{D.3}$$

By substituting $x = sN_0\mathrm{SF}_{\mathrm{LQ}}(D)$ and noting that $\mathrm{P}_{\mathrm{LQ}}(D)^s = \mathrm{e}^{-x}$, it follows that (D.3) is equivalent with

$$x < 1 - \mathrm{e}^{-x},$$

which is not true for $x \in \mathbb{R}$. Hence, it follows that (D.2) holds for any $D \geq 0$, which implies that $-\ln(1-\mathrm{NTCP}_{\mathrm{RS}}(\mathbf{D})^s)$ is a convex function. □

The following theorem shows that $-\ln(1-\mathrm{NTCP}_{\mathrm{RS}}(\mathbf{D}))$ is a strictly convex function of the dose distribution, provided that $0 < s < 1$.

Theorem 5. *Let $F(\mathbf{D}) = -\ln(1-NTCP_{RS}(\mathbf{D})^s)$ be the convex function from (D.1) and let $h(z) = -\ln(1-(1-e^{-z})^{1/s})$. Then the composition*

$$h(F(\mathbf{D})) = -\ln(1-NTCP_{RS}(\mathbf{D}))$$

is a strictly convex function on the set of dose distributions $\{\mathbf{D} \mid D_k \geq 0, k = 1,\ldots,M\}$ if $0 < s < 1$.

Proof. From Lemma 2 it is known that $G(\mathbf{D}) = h(F(\mathbf{D}))$ is convex if F is convex and h is convex and nondecreasing. Therefore, we only have to prove that h is convex and nondecreasing.

Let $t = 1/s$, with $0 < s < 1$. Then $t > 1$. Differentiating h with respect to z yields

$$h'(z) = \frac{t(1-\mathrm{e}^{-z})^{t-1}\mathrm{e}^{-z}}{1-(1-\mathrm{e}^{-z})^t}.$$

Since $0 < \mathrm{NTCP}_{\mathrm{RS}}(\mathbf{D}) < 1$, it follows that $z > 0$, from which it is easy to see that $h'(z) > 0$ and thus h is nondecreasing.

For the second derivative of h to be strictly positive, it can be derived that the inequality

$$te^{-z} - 1 + (1 - e^{-z})^t > 0$$

should hold. By substituting $x = e^{-z}$, and noting that $0 < x < 1$, we obtain

$$p(x) = tx - 1 + (1 - x)^t,$$

for which we have to show that $p(x) > 0$ for $0 < x < 1$. Note that $p(0) = 0$ and that the first derivative of p with respect to x is given by $p'(x) = t(1 - (1 - x)^{t-1})$. Since $p'(x) > 0$ for $0 < x < 1$ and $t > 1$, p is an increasing function of x. From $p(0) = 0$ and $p'(x) > 0$ for $0 < x < 1$, it follows that $p(x) > 0$ for $t > 1$ on the interval $0 < x < 1$, which implies that h is strictly convex. □

Using similar arguments, it can be shown that $p(x)$ is negative for $0 < x < 1$ in case $s > 1$, which implies that h is concave. However, this does not imply that $-\ln(1-\mathrm{NTCP}_{\mathrm{RS}}(\mathbf{D}))$ is not convex in $\mathbf{D}$. Unfortunately, the composition rules cannot be used to prove that $-\ln(1-\mathrm{NTCP}_{\mathrm{RS}}(\mathbf{D}))$ is indeed a convex function, provided that $s > 1$.

Appendix E.

The following theorem shows that $\mathrm{gEUBED}_a(\mathbf{D})$ is a convex function of the set of physically meaningful dose distributions.

Theorem 6. *The function $gEUBED_a(\mathbf{D})$ of (8) is convex on the set of dose distributions $\{\mathbf{D} \mid D_k \geq 0, k = 1, \ldots, M\}$ for $a \geq 1$.*

Proof. From its definition, it is clear that (8) is obtained from (7) by a vector transformation of its argument, i.e., $\mathrm{gEUBED}_a(\mathbf{D}) = \mathrm{gEUD}_a(\mathbf{h}(\mathbf{D}))$ with $\mathbf{h}(\mathbf{z}) = (h_1(\mathbf{z}), \ldots, h_N(\mathbf{z}))$ and $h_k(\mathbf{z}) = \mathrm{BED}(z_k) = z_k \left(1 + \frac{z_k/n}{\alpha/\beta}\right)$, where $k = 1, \ldots, N$. The convexity of $\mathrm{gEUBED}_a(\mathbf{D})$ follows from the fact that for $a \geq 1$ the function $\mathrm{gEUD}_a(\mathbf{D})$ is convex and nondecreasing in each argument (Choi & Deasy 2002), and h_k are strictly convex. □

The following theorem shows that $-\ln(1-\mathrm{NTCP}_{\mathrm{LKB}}(\mathbf{D}))$ is a convex function of $\mathbf{D}$.

Theorem 7. *Consider the function $NTCP_{LKB}(\mathbf{D})$ of (9) with $a \geq 1$. Let $h(z) = -\ln(1 - z)$, which is a strictly increasing function. Then the composition*

$$h(NTCP_{LKB}(\mathbf{D})) = -\ln(1 - NTCP_{LKB}(\mathbf{D}))$$

is a convex function on the set of dose distributions $\{\mathbf{D} \mid D_k \geq 0, k = 1, \ldots, M\}$.

Proof. Note that $-\ln(1 - \mathrm{NTCP}_{\mathrm{LKB}}(\mathbf{D}))$ is a transformation of $\mathrm{gEUBED}_a(\mathbf{D})$:

$$-\ln(1 - \mathrm{NTCP}_{\mathrm{LKB}}(\mathbf{D})) = \Psi\left(\frac{\mathrm{gEUBED}_a(\mathbf{D}) - \mathrm{BED}_{50}}{m\ \mathrm{BED}_{50}}\right),$$

where $\Psi(t) = -\ln(1 - \Phi(t))$, and $\Phi(t)$ is the standard normal cumulative distribution function. For Lemma 2 to be applicable, we use Theorem 6 and only have to show that $\Psi(t)$ is an increasing and convex function of t.

Since h and Φ are strictly increasing functions, $\Psi(t)$ is strictly increasing in t. To investigate whether $\Psi(t)$ is convex, the first derivative with respect to t is calculated:

$$\Psi'(t) = \frac{\Phi'(t)}{1 - \Phi(t)}.$$

From Bagnoli & Bergstrom (2005) it is known that if $\Phi'(t)$ is logarithmic concave, $\Psi'(t)$ is increasing. Note that it is easy to show that $\Phi'(t)$ is log-concave, since

$$\ln(\Phi'(t)) = \ln\left(\frac{1}{\sqrt{2\pi\sigma}}\right) - \frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2$$

is a parabola with strictly negative curvature. Since $\Psi'(t)$ is increasing, the second derivative of Ψ with respect to t is positive, which implies that $\Psi(t)$ is convex. □

Appendix F.

The following theorems show that the log-transformed composite TCP and NTCP functions are concave and convex, respectively, on the set of physically meaningful dose distributions for the relevant dose-response models defined in Section 2 of this paper.

Theorem 8. *Consider the function TCP*$(\mathbf{D})$ *of (10), with TCP*$_j(\mathbf{D})$ *according to TCP*$_{LQ}(\mathbf{D})$ *of (2). Let* $h(z) = \ln z$*, which is a strictly increasing function. Then the composition*

$$-h(TCP(\mathbf{D})) = -\ln TCP(\mathbf{D}) = -\sum_{j=1}^{T} \ln TCP_j(\mathbf{D}),$$

is a strictly convex function on the set of dose distributions

$$\{\mathbf{D} \mid D_k > \sqrt{\frac{1}{2}\frac{(\alpha/\beta)n}{\alpha}} - \frac{1}{2}(\alpha/\beta)n, \quad k = 1, \ldots, M\}.$$

Proof. As shown in Theorem 1, the function $-\ln \mathrm{TCP}_{\mathrm{LQ}}(\mathbf{D})$ is strictly convex if (13) is satisfied. By noting that the nonnegative weighted sum of strictly convex functions is strictly convex, it is evident that $-\ln \mathrm{TCP}(\mathbf{D})$ is strictly convex in $\mathbf{D}$. □

Theorem 9. *Consider the function NTCP*$(\mathbf{D})$ *of (11) with NTCP*$_j(\mathbf{D})$ *according to NTCP*$_{RS}(\mathbf{D})$ *of (3) or NTCP*$_{LKB}(\mathbf{D})$ *of (9). Let* $h(z) = -\ln(1-z)$*, which is a strictly increasing function. Then the composition*

$$h(NTCP(\mathbf{D})) = -\ln(1 - NTCP(\mathbf{D})) = -\sum_{j=1}^{T} \xi_j \ \ln(1 - NTCP_j(\mathbf{D})),$$

is a strictly convex function on the set of dose distributions $\{\mathbf{D} \mid D_k \geq 0, k = 1, \ldots, M\}$.

Proof. According to Theorems 5 and 7, the functions $-\ln(1 - \mathrm{NTCP}_{\mathrm{RS}}(\mathbf{D}))$ and $-\ln(1 - \mathrm{NTCP}_{\mathrm{LKB}}(\mathbf{D}))$ are strictly convex, respectively. By noting that the nonnegative weighted sum of strictly convex functions is strictly convex, it is evident that $-\ln(1 - \mathrm{NTCP}(\mathbf{D}))$ is strictly convex in $\mathbf{D}$. □