The use and QA of biologically related models for treatment planning: Short report of the TG-166 of the therapy physics committee of the AAPM[a)]

X. Allen Li[b)]
Radiation Oncology, Medical College of Wisconsin, Milwaukee, Wisconsin 53226

Markus Alber
Clinic for Radio-oncology, University of Munich, 72076 Tübingen, Germany

Joseph O. Deasy, Andrew Jackson
Medical Physics, Memorial Sloan-Kettering Cancer Center, New York, New York 10065

Kyung-Wook Ken Jee
Radiation Oncology, Massachusetts General Hospital, Boston, Massachusetts 02114

Lawrence B. Marks
Radiation Oncology, University of North Carolina, Chapel Hill, North Carolina 27599

Mary K. Martel
Radiation Physics, UT MD Anderson Cancer Center, Houston, Texas 77030

Charles Mayo
Radiation Oncology, Mayo Clinic, Rochester, Minnesota 55905

Vitali Moiseenko
Physics, Vancouver Cancer Center, Vancouver, British Columbia V5Z 4E6 Canada

Alan E. Nahum
Department of Physics, Clatterbridge Centre for Oncology, Wirral, Merseyside CH63 4JY United Kingdom

Andrzej Niemierko
Radiation Oncology, Massachusetts General Hospital, Boston, Massachusetts 02114

Vladimir A. Semenenko
Radiation Oncology, Medical College of Wisconsin, Milwaukee, Wisconsin 53226

Ellen D. Yorke
Medical Physics, Memorial Sloan-Kettering Cancer Center, New York, New York 10065

(Received 2 November 2011; revised 10 January 2012; accepted for publication 28 January 2012; published online 21 February 2012)

Treatment planning tools that use biologically related models for plan optimization and/or evaluation are being introduced for clinical use. A variety of dose-response models and quantities along with a series of organ-specific model parameters are included in these tools. However, due to various limitations, such as the limitations of models and available model parameters, the incomplete understanding of dose responses, and the inadequate clinical data, the use of biologically based treatment planning system (BBTPS) represents a paradigm shift and can be potentially dangerous. There will be a steep learning curve for most planners. The purpose of this task group is to address some of these relevant issues before the use of BBTPS becomes widely spread. In this report, the authors (1) discuss strategies, limitations, conditions, and cautions for using biologically based models and parameters in clinical treatment planning; (2) demonstrate the practical use of the three most commonly used commercially available BBTPS and potential dosimetric differences between biologically model based and dose-volume based treatment plan optimization and evaluation; (3) identify the desirable features and future directions in developing BBTPS; and (4) provide general guidelines and methodology for the acceptance testing, commissioning, and routine quality assurance (QA) of BBTPS.
[DOI: 10.1118/1.3685447]

Key words: biologically based treatment planning, outcome modeling, planning system QA, TCP, NTCP, EUD

TABLE OF CONTENTS

I. INTRODUCTION

The goal of radiation therapy (RT) is to deliver a therapeutic dose of radiation to target tissues while minimizing the risks of normal tissue complications. Until recently, the quality of a RT plan has been judged by physical quantities, i.e., dose and dose-volume (DV) parameters, thought to correlate with biological response rather than by estimates of the biological outcome itself. It is widely recognized that the DV criteria, which are merely surrogate measures of biological responses, should be replaced by biological indices in order for the treatment process to more closely reflect clinical goals of RT.[1] Developments in our understanding of advantages and limitations of existing dose-response models begin to allow the incorporation of biological concepts into a routine treatment planning process.

I.A. Limitations of dose-volume-based treatment planning

Single or multiple DV constraints used for inverse treatment planning or plan evaluation are based on clinical studies that demonstrate correlation between treatment outcome [e.g., tumor control probability (TCP) and normal tissue complication probability (NTCP)] and particular DV metrics. For example, V_{20} (percentage of volume receiving at least 20 Gy) for lung is used to gauge the probability of a plan causing grade ≥ 2 or grade ≥ 3 radiation pneumonitis.[2] There are a number of limitations associated with this approach. (1) Typically, more than one point on the DV histogram (DVH) (e.g., V_5, V_{40}, mean lung dose) correlates to the complication,[3] indicating that the different portions of the DVH curve may correlate with risk. This correlation is, however, specific to treatment delivery technique, i.e., intensity-modulated RT (IMRT) or 3D conformal RT (3DCRT), beam arrangements, etc.[3] (2) Generally, optimization with DV constraints is indirect, requiring substantial skill in selecting values and relative weights for constraints to achieve optimal TCP and NTCP values. With typically 1–3 constraints, a range of optimized organ-at-risk (OAR) DVHs that satisfy these few constraints, but carry a distinctly different risk of complications, is possible. In biologically based optimization, the DV points are replaced with a function that more efficiently drives the shape of the DVH curves to achieve the

plan leading to the most favorable overall treatment outcome, rather than satisfying the constraints (explained in Sec. II A 3). (3) Specifying multiple DV constraints increases computational complexity. Moreover, the resulted cost function can lead to multiple local minima.[4,5] This implies that a search algorithm designed for global minimum problems is likely to get trapped in a local minimum, potentially leading to less favorable dose distributions.

Most current systems that use DV-based plan optimization also lack tools for routinely evaluating biologically based metrics alongside DVH metrics. Evaluating plans with biologically based metrics is important in developing institutional datasets for outcome correlation and in comparing different RT methods. Since dose distributions driven by biological constraints may differ substantially from those driven by DV constraints, the evaluation tools are helpful in preferentially adopting biologically based optimization.

I.B. Scope of this report

Historically, dose prescription in RT has been performed using population-based knowledge about behavior of a particular type of tumor or normal tissue. Rapid advances in functional imaging, molecular techniques, predictive assays, and RT delivery technology will sooner or later enable implementing truly individualized RT in the clinic.[6] The Task Group (TG) will refer to a framework of RT that takes advantage of information about spatial and temporal distribution of relevant patient-specific biological parameters, such as tumor and normal cell radiosensitivity, oxygenation status, proliferation rate, etc., as biologically guided radiation therapy (BGRT). Other investigators have previously used various alternate terms to describe the same basic concept, e.g., radiobiologically optimized radiation therapy,[7] multidimensional conformal radiotherapy,[8] biologically conformal radiation therapy,[9] biologically based radiation treatment planning,[1] theragnostic imaging,[10] and risk-adaptive optimization.[11]

An integral part of BGRT is the ability to design dose distributions that would produce the desired balance between tumor cure and normal tissue injury based on the knowledge of biological properties of the particular tumor and surrounding normal tissues. Such a multidimensional problem is most appropriately addressed in the framework of inverse treatment planning presently employed for the optimization of IMRT plans and will rely on models to describe relationships between dose distributions and biological outcomes. The TG will refer to any use of biological response models that involves feedback from a model during the treatment planning process as biologically based treatment planning (BBTP). The feedback may be either automated in the case of inverse treatment planning, or with active participation from the planner in the case of forward treatment planning.

Whereas future development of BGRT relies on sufficient advances in methods to obtain individualized biological parameters, BBTP systems (BBTPS) have already started to enter clinical practice. BBTP is viewed as a subset of BGRT. However, BGRT is more than just BBTP based on patient-specific biological parameters; BGRT will employ biological models not only at the time of initial planning but also to adapt treatment based on tumor/OAR response to RT (e.g., Ref. 12). For a detailed discussion on BGRT and BBTP, readers are referred to a vision 20/20 paper by Stewart and Li.[6] The scope of this TG report is limited to BBTP and BBTPS, i.e., the use of biological models for plan optimization and evaluation in external beam radiation therapy. The emphasis is made on the models that are implemented or may potentially be implemented in commercial treatment planning system (TPS). More description on the terminology used in this report is provided in Appendix A.

Presented below is a short and condensed version of the TG report with main focuses on (1) general strategies, advantages, and limitations for using BBTP, (2) practical issues demonstrated with three available representative BBTPS, (3) desirable features and development directions for future BBTPS, and (4) general guidelines for the acceptance testing, commissioning and routine quality assurance (QA) of BBTPS. The full TG report that includes a review of the dose-response models and many other details is available at the website given in Ref. 13.

II. USES OF BIOLOGICALLY BASED MODELS IN TREATMENT PLANNING

A variety of dose-response models for tumor and normal structures are available and can be broadly characterized as either mechanistic or phenomenological. The former attempt to mathematically formulate the underlying biological processes, whereas the latter simply intend to fit the available data empirically. Mechanistic models are often considered preferable, as they may be more rigorous and scientifically sound. However, the underlying biological processes for most tumor and normal tissue responses are fairly complex and often are not fully understood, and it may not be feasible to accurately and/or completely describe these phenomena mathematically. On the other hand, phenomenological models are advantageous since they typically are relatively simple compared to the mechanistic models. Their use obviates the need to fully understand the underlying biological phenomena. Limitations of such empirical approaches are that they strive for mathematical simplicity and thus are limited in their ability to consider more complex phenomena. Further, it may be somewhat risky to extrapolate model predictions beyond the realm within which the model and parameter values were validated. A mechanistic model might be more forgiving in its ability to extrapolate to these more uncertain areas.

The field of study of biophysical models of tumor and normal tissue responses to dose is extensive and beyond the scope of this report to fully explore. As described below (Sec. III), phenomenological models are mostly used in the currently available BBTPS due to their simplicity in implementation. It is a vision of this TG that more mechanistic models will be employed as BBTPS advances (Sec. IV). A review of the models and model parameters that are used or have potential to be used for BBTP is presented in the full

TG report.[13] In this early stage of BBTPS, the equivalent uniform dose (EUD) (Refs. 14 and 15) is the most commonly used phenomenological models. As demonstrated with EUD models, general strategies, advantages and limitations for using biologically based models for treatment planning are described in this short report.

Biologically based figures of merit may be used for both plan optimization and evaluation. Both tasks are closely related as any optimization algorithm continuously evaluates treatment plans and alters them incrementally in order to improve their figures of merit. However, desired properties (e.g., predicting power) of biological models may differ whether they are used for plan optimization or plan evaluation. As pointed out by Choi and Deasy,[16] treatment optimization only requires a model to have the ability to steer the optimization process in the desired direction. In contrast, for an effective use of dose-response models in plan evaluation, especially when absolute TCP/NTCP values are used to guide clinical decisions, the accuracy of model predictions is of paramount importance.

II.A. Biological models in plan optimization

II.A.1. Advantages of biological cost functions over dose-volume cost functions

Optimization criteria based on biologically related models are potentially more versatile and directly associated with treatment outcome than those based on DV criteria. If biologically related models are constructed to capture the dose response, they would allow some extrapolation beyond the range of clinical evidence. Unfortunately, there is no guarantee that a biologically related model does indeed estimate the consequence of dose distributions if they deviate greatly from the baseline dataset which led to the model parameters. However, for the purpose of dose optimization, it is sufficient as long as the use of the model can guide the optimization toward to favorable dose distributions.

Another aspect of plan optimization is that the figure of merit has to address the inevitable variability of patient geometries, and resultant dose distributions, in a population. In this regard, multiple DV criteria for a single organ quickly may become problematic as they need to be given an individual priority and ideally ought to be combined into a single figure of merit to avoid ambiguities. In contrast, biologically related models have the potential to provide an inherent prioritization of multiple DV criteria incorporated in a single figure of merit.

The optimization with cost functions based on the EUD concept, commonly used in the available BBTPS, is straightforward and numerically expedient.[16–20] In particular, the generalized EUD (gEUD) can be formulated as:[15] $\mathrm{gEUD} = \left(\sum_i v_i D_i^a\right)^{1/a}$, where v_i is the fractional organ volume receiving a dose D_i and a is a tissue-specific parameter that describes the volume effect. This formula allows to consider tissue-specific property into the planning process that cannot be done with DV-based optimization. For example, to mimic the effects of cold spots on TCP, a is taken as negative (often taken as <-10) for tumors. For serial-response complications, which depend strongly on the highest dose to the tissue (see Appendix A) rather than the overall dose distribution, a is large (>10); for parallel-response complications (Appendix A), $a \sim 1$.

II.A.2. Precautions for using biological models in plan optimization

As most of the currently available BBTPSs use EUD/gEUD based optimizations, we mainly discuss the concerns with EUD. With respect to optimization, the DV effect incorporated by a EUD-based model is of paramount importance. For example, the assumption that a normal tissue responds in serial manner leads to lack of control over the low- and mid-dose range, as the risk of complications is predominantly determined by the high doses. Conversely, if a parallel behavior is assumed, hot spots are allowed but large volumes to lower doses are undesirable. In a worst case scenario, a serially responding complication would be assumed to behave in a parallel fashion. Plan optimization may then be steered toward allowing clinically unacceptable hot spots. If in doubt, one should always maneuver intentionally toward a smaller volume effect (serial behavior) as this will put a limit on both the size and the dose of hot spots in normal tissues during plan optimization.

The gEUD models for serial response do not give rise to local minima of the optimization problem.[16] On the other hand, it cannot be ruled out that gEUD models for parallel response create local minima, though due to their generally less pronounced nonlinearity, this risk is less than that for DV objectives. Although gEUD itself is convex for parameter $a \geq 1$,[16] any cost function formulated as a product of nonlinear gEUD/TCP/NTCP models is subject to violating the convex or quasi-convex properties of the component gEUDs.[17] Two examples include the probability of complication-free tumor control, P_+ (Ref. 21) and the product of sigmoid functions based on gEUD.[22] Although the clinical significance of local minima remains to be seen,[5] from a mathematical point of view, inappropriate choice of the cost function for plan optimization may result in multiple local minima, which diminishes the theoretical advantages of using biological-model-based cost functions.

Direct maximization of biological indices for targets (e.g.,TCP or gEUD) is known to produce highly inhomogeneous target dose distributions[22,23] because TCP is increased by the creation of hot spots and using TCP alone does not penalize hot spots. Thus, one must consider limiting planning target volume (PTV) dose inhomogeneity or at least constraining the hot spots to the gross tumor volume (GTV) or clinical target volume (CTV). This can be achieved by adding physical maximum dose cost functions to optimization criteria for target volumes. Alternatively, the hot spots in target volumes can be controlled using biological cost functions assuming serial response by treating the targets as both tumor and "hypothetical" normal tissues.[22]

Another challenging issue is to use a biologically related model for a fractionation scheme that is very different from the scheme under which the model was derived. Applying

model parameters that were derived for a conventional fractionation scheme to the optimization of a hypofractionated treatment [e.g., stereotactic body RT (SBRT)] is especially hazardous. In the absence of clinical data to provide guidance, this TG advises adjusting parameters (DV or biological) to steer critical organ doses into a dose-volume zone that is proven to be safe clinically. Examples of such safety zones are provided in the reports from the recent QUANTEC (QUantitative Analysis of Normal Tissue Effects in the Clinic) initiative[24] and from the TG 101.[25]

II.A.3. Strategies for effective use of biological models in plan optimization

For plan optimization using biologically related models, two rivals, yet complementary, concepts are the EUD and TCP/NTCP models. In one sense, the difference between EUD-based models and TCP/NTCP models is irrelevant, because every TCP/NTCP model can be converted into an EUD model (find the dose which results in the equivalent probability if applied to the whole volume) and vice versa (choose a suitable sigmoid wrapper function which maps EUD onto the interval [0, 1]). Note that this EUD does not need to be the same simple expression as the "gEUD."[17,18]

To optimize a plan based on biological models, multiple EUD-dependent goals can be combined as a weighted combination that defines the composite cost function for the optimization algorithm. These biologically related goals usually need to be supplemented by any number of physical goals which would ensure certain properties deemed clinically desirable by a treating physician. For example, a limit on hot spots in the target volume is usually motivated by established clinical practice rather than rigorous biological considerations. Furthermore, there may be biological goals in normal tissues for which no dose-response model exists, for example the overall conformity of the dose to the target volume. All of these cost functions can be combined in optimization, although how this is implemented is determined by the treatment planning system.

It can be beneficial to treat EUD-based cost functions as hard constraints because they are directly associated with control/complication risks. On the other hand, the definition of EUD allows for a certain freedom in shaping the dose distribution. Therefore, EUD constraints are less restrictive than multiple DV constraints and offer an inherent tradeoff between different dose levels, allowing controlled violations for some DV constraints while over fulfilling other constraints to generate an overall better dose distribution.

Although each biologically motivated cost function incorporates a specific volume effect and thus favors a certain shape of the dose distribution, the result of an optimization depends on a complex interplay of all participating terms of the cost functions. For this reason, it is essential that the treatment planner understands which traits of a dose distribution are controlled by the chosen cost function terms, and which traits are merely coincidental. It is helpful to visualize the action of a EUD-based model on a DVH as a set of connected DV objectives whose weight grows in a specific fashion. For a cost function term assuming serial response, the weight of these virtual DV objectives grows with dose, as demonstrated in Fig. 1(a). The smaller the volume effect, the more rapidly the weight of these objectives grows. For a cost function term assuming parallel response, their weight should tend to zero for very high doses and reach a maximum around the mid-dose range [Fig. 1(b)].

The best measure against the hazards of using biologically motivated cost functions is to understand their effect on the dose distribution and to know the desirable properties of the final dose distribution. The overall dose distribution derived from such an optimization should be carefully inspected—one should not rely purely on DVH metrics. Each desired goal should be reflected by a specific cost function term which should be chosen to be capable of controlling this particular property of the dose distribution sufficiently. Thus, the task of setting up a biologically related optimization problem becomes, in the order of increasing importance: (1) choice of sufficiently many cost functions, (2) choice of right types of cost functions, (3) choice of right volume effect parameters, and (4) clear idea of what features make a dose distribution acceptable or unacceptable in your clinic. For example, an organ like spinal cord, for which the maximum dose is considered to have the highest priority, is ideally modeled by gEUD with $a \gg 1$. This kind of model is very sensitive to high doses while it is very insensitive to low and intermediate doses. Clearly, this kind of behavior is not sought for organs like lung, where the primary objective

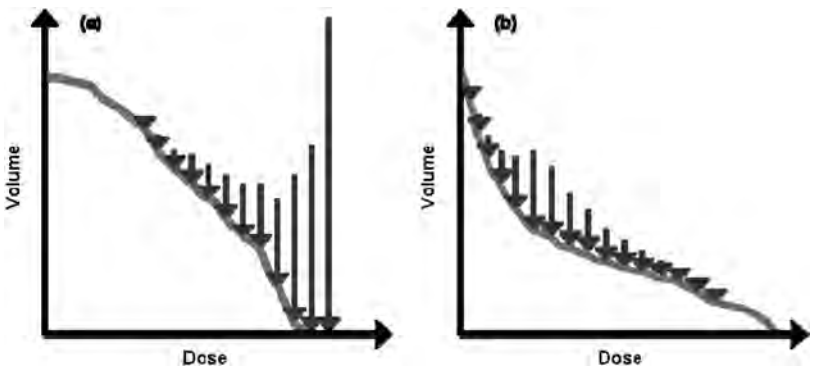

FIG. 1. Weights of "virtual" DV objectives representing the same volume effect as a serial-type cost function (a) or a parallel-type cost function (b).

is to spare sufficient lung volume from intermediate doses while controlling the maximum dose is only of secondary importance. Here, a gEUD model with a smaller a value or a parallel complication model is a better choice, but one has to be aware that this type of model does not control the maximum dose. In order to achieve this, it needs to be complemented with either a second gEUD model with a greater parameter a value or a maximum dose constraint. Notice that, in this example, the two models represent two types of complication control with different volume dependency: one aiming to control volume-related complications like pneumonitis and loss of lung function, while the other trying to manage more local complications like destruction of large blood vessels or even necrosis.

II.A.4. Effects of DVH computation inaccuracies and statistical uncertainties on plan optimization using biological models

Biologically based models that use more pronounced nonlinear functions than DV functions tend to amplify the effects of any uncertainty in the dose and/or DVH computation. For example, the EUD can be calculated directly from the DVH. Depending on the implementation of the TPS, a DVH may be more than just the straightforward statistics of the voxel doses of an organ. Thus, EUD computed directly from the dose calculation grid and from the DVH may differ. Further, all issues associated with the computation of a DVH, such as voxelization, interpolation, binning, and volume normalization, affect the computation of EUD. If the uncertainties are random and not systematic in nature, the EUD error will usually tend to be on the safe side, i.e., normal tissue EUD will be overestimated while target EUD will be underestimated. This is a consequence of the positive curvature of the most common EUD implementations (with the exception of the generalized EUD for parallel complications). These TPS-related sources of error can be taken into account in practice if a number of treatment plans that were considered safe are retrospectively evaluated with the EUD models intended for future dose optimization.

A special case of EUD estimate bias arises if the dose distribution is calculated with Monte Carlo methods. Here, the statistical uncertainty of the dose translates into a systematic error of EUD (Ref. 26) which grows with the magnitude of the noise in the dose. For low statistical uncertainties this error can be corrected precisely. Thus, EUD calculated without this correction from a Monte Carlo dose grid may differ from the values based on DVH.

II.B. Biological models in plan evaluation

II.B.1. Advantages of biological models over DV evaluation criteria

As with plan optimization, either EUD or TCP/NTCP models can be used for biologically based plan evaluation. Although both concepts can be used interchangeably for plan optimization, the EUD has the advantage of fewer model parameters, as compared to TCP/NTCP models, and allows more clinical flexibility. The proper calibration of a TCP/NTCP model requires monitoring the outcomes for a large number of patients. In contrast, EUD models can rank a number of treatment plans without having to quantify the actual tumor response/complication risk as long as the chosen parameters (a in the case of gEUD) are calibrated to give reasonable results for clinical plans for which the treatment outcomes are known. An EUD model can be calibrated against the past clinical practice of any institution simply by calculating the previously applied distribution of EUD values. This establishes a reference range of EUD values that were considered acceptable in the past.

Thus, a properly calibrated EUD model has the potential to provide a reliable ranking of rival plans and is most useful when a clinician needs to select the best plan from two or more alternatives. Of course, it is essential that the clinician understands the prior calibration process and is willing to consider biological evaluation. The utility of EUD for evaluating a single plan is limited. In contrast, properly calibrated TCP/NTCP models can provide direct estimates of outcome probabilities, which are more clinically meaningful than the EUD. If these estimates are within the clinician's goals, the treatment plan under consideration can be accepted without having to explore other possible plans. The disadvantage of TCP/NTCP models is that they require more parameters (most commonly three) and more effort for their calibration as compared to the EUD (one parameter for gEUD). Similar to an EUD model but with more importance, a TCP/NTCP model derived from the experience of other institutions (different TPS, dose calculation, patient population, dose fractionation, etc.) must be applied with extreme caution.

Use of DV criteria (or EUD alone) for plan evaluation implies a binary outcome, i.e., an effect occurs if a DVH passes above a certain point in DV space, and does not occur in the DVH passes below. Such threshold-like behavior of tumor control/complication risk is a rough approximation of actual biological processes. In contrast, biological evaluation metrics in the form of TCP/NTCP provide continuous estimates of outcome probabilities. Also, consider a case when two or more DV points are used to evaluate a dose distribution in a particular organ. It might happen that the dose distribution passes the evaluation test for some points and fails for others, requiring the treatment planner/radiation oncologist to prioritize different DV criteria. Biological metrics may be advantageous in such situations because they can weigh various DV criteria and can condense them into a single unambiguous estimate of biological outcome.

Some mechanistic biological models directly incorporate terms describing radiosensitivity as a function of dose per fraction. If properly calibrated using the data clinically approved for a range of fraction sizes, these models implicitly take into account the dose per fraction effects and can be used to predict outcomes of different fractionation schemes. DV criteria, on the other hand, apply to a single fraction size for which their efficacy has been tested. If the standard fractionation scheme is significantly altered, DV-based prescription/normal tissue constraints need to be explicitly modified based on clinical experience and/or isoeffect calculations.

II.B.2. Precautions for using biological models in plan evaluation

In contrast to the use of biological models in plan optimization, where biologically based cost functions are only required to capture the correct volume effect and to steer dose distributions in a desired direction, the use of biological models to replace DV criteria in plan evaluation requires clinically realistic correct plan ranking and/or outcome estimates. To evaluate a particular plan (not just plan ranking), accurate TCP/NTCP models and parameter estimates become absolutely essential. It is also essential that the models used be applied retrospectively to make sure that they agree with the treatments that you know to be safe and effective in your practice. Whether the problem lies in the abstract model or its implementation in TPS, such a reality check is necessary before using a model for clinical plan evaluation. As it is desirable to incorporate outcome data in the treatment planning process, two options exist for using biological models in plan evaluation. The users can derive TCP/NTCP model parameters based on their own experience by calibrating selected model(s) against observed clinical outcomes. This approach has the potential to yield the most reliable data directly reflecting the practice adopted at a particular institution. Furthermore, initial parameter estimates can be easily refined as additional follow-up data become available. However, this method may not be feasible for many small and even midsized institutions, as it requires expertise in outcome modeling, sufficient patient throughput, and substantial time commitment.

Another option is to cautiously use published parameter values. Published data are available for many tumor sites and complication types (see the full TG report on-line), affording the user a variety of choices. However, this approach is fraught with significant risks if published parameter sets are applied injudiciously without following the same practices that were used to generate the original data (e.g., Ref. 27). Caution should be exercised if clinical and demographic characteristics of the patient population under evaluation differ substantially from those in the original patient cohort used to derive published parameter estimates. The reason is that additional variables influencing the outcome, which were not present in the original population, may be present in the evaluated patient population (e.g., Ref. 28). When using published parameter estimates for plan evaluation, it has to be carefully verified that they apply to the appropriate endpoints, organ volume definitions and fractionation schemes.

Most NTCP models do not include explicit description of dose per-fraction effects in the attempt to minimize the number of parameters. Given this, whether one is using in-house or published data, parameter estimates can only be used to evaluate treatment plans corresponding to a narrow range of doses per fraction similar to the doses per fraction in the original patient population. If the fraction size in a plan under evaluation is very different from that in the dataset used to derive parameter estimates, both sets of data should be normalized to the same dose per fraction, usually using the linear-quadratic (LQ) formalism.[24] If the dose per fraction varies considerably in the patient cohort of which parameter estimates are being derived, it is reasonable to normalize all doses to some standard fractionation scheme. Examples include Lyman Kutcher Burman (LKB) modeling[29,30] of liver[31] and lung[32] complications. These sets of parameter estimates, obviously, work best for fraction size corrected normal tissue DVHs. However, if these parameter values are used for noncorrected dose delivered with a conventional target fraction size (i.e., $\sim$2 Gy per fraction), the fraction sizes for normal tissues are much less than 2 Gy, and hence, a model produces conservative overestimates of NTCP. This argument is reversed for hypofractionated delivery, and the model can significantly underestimate the risk of a complication.[33] Even if the prescription fraction size is unchanged, the simultaneous use of an increasing number of beams/orientations (e.g., with multifield IMRT) reduces the dose per fraction in the exposed normal tissues away from the target, compared to what they would have seen with a "conventional" plan with a limited number of beam orientations used sequentially (e.g., anterior–posterior/posterior–anterior beams followed by opposed oblique beams).

Whether self-derived or published parameter estimates are used, it is essential to standardize the organ volume relative to which the parameter is computed. For example, the EUD or NTCP for rectum and rectal wall will differ because the dose distributions in each volume differ. The EUD or NTCP will also depend on the delineated length of rectum or rectal wall. Much more subtle is the computation of biological indices for the spinal cord, where either a standardized length has to be segmented (e.g., including all thoracic and cervical vertebrae) or the parameter is computed relative to a normalized volume. Care should be taken that for parallel organs, whose response is correlated with the mean dose, the entire organ is included in the image set and dose calculation grid.

Parameter estimates clearly should be used only with the model for which they were derived. In some cases, fits to more than one model are available for the same dataset. For such situations, it has been observed that different NTCP models often provide different answers to important clinical problems.[34–36] It is generally not possible to determine which model is right based on observing fits to clinical data.[35] To resolve this situation and ensure further progress in the use of biological models for plan evaluation, concerted efforts to select the most practical models and to create databases of parameter estimates are urgently needed. Such sets of data (e.g., the QUANTEC initiative), being supported by experts in TCP/NTCP modeling, will provide a strong basis for TPS manufacturers to include biologically based evaluation tools in their products.

In general, biological figures of merit for target volumes require much less consideration since their utility for outcome prediction is frequently limited by uncertainties of individual tumor biology. Also, current clinical practice demands homogeneous doses to the PTV, which usually includes a large share of normal tissue, while a TCP figure can only be meaningful for the GTV or CTV. It is important to understand which aspects of a target dose distribution

TABLE I. Biological models used for treatment plan optimization in CMS MONACO.

Structure type	Name	Parameters	Objectives/constraints	Comments
Target	Poisson statistics cell kill model	Cell sensitivity ($0.1–1.0\,Gy^{-1}$)	Prescription (1–150 Gy)	Mandatory cost function for targets; no penalty for hot spots
OAR	Serial complication model	Power law exponent (1–20)	Equivalent uniform dose (1–150 Gy)	Effective for controlling maximum organ dose
OAR	Parallel complication model	Reference dose (1–100 Gy) Power law exponent (1–4)	Mean organ damage (1–100%)	Effective for controlling mean organ dose

influence the TCP. Various investigators have demonstrated using the Poisson-based model with interpatient heterogeneity that even very small cold spots may considerably decrease the TCP, whereas the hot spots only affect the TCP to a great extent if the volume of the hot spot is large.[37–40]

Available sets of TCP parameter estimates are less consistent than NTCP parameters in a sense that different analyses use somewhat different assumptions when deriving model parameters, e.g., fixed number of clonogens vs fixed clonogen density, inclusion or exclusion of the time factor, etc. Strictly speaking, it is incorrect to apply parameters derived using one set of assumptions to even a slightly modified model. This poses difficulties because users wishing to integrate TCP calculations into their plan evaluation routine need to implement not only different models that were used to analyze data for different sites but also different variations of the same basic model. Large sets of TCP data compiled using uniform criteria are rare (e.g., Ref. 40). Efforts similar to QUANTEC are needed to summarize TCP data and derive common sets of parameters for one or two models, which could then be built into commercial TPS.

Finally, current clinical standards for acceptable treatment plans in external beam RT include certain dose-volume goals that are not readily transcribed using biological metrics, such as target dose uniformity and overall conformity of high-dose regions. These goals should be considered separately when EUD/TCP/NTCP is used for plan evaluation.

III. THE USE OF CURRENTLY AVAILABLE TREATMENT PLANNING SYSTEMS EMPLOYING BIOLOGICAL MODELS

Three commercially available and most commonly used TPSs that employ biologically based models are selected to demonstrate the practical use of BBTPS and some issues discussed above (Sec. II). These three systems, presented in a chronological order based on their releasing times, are MONACO V1.0 (CMS/Elekta, St. Louis, MO), PINNACLE V8.0h (Philips Medical Systems, Andover, MA), and ECLIPSE V10.0 (Varian Medical Systems, Palo Alto, CA) systems. A description including specific features, algorithms, and the models used for each of these systems is provided in Appendix B. Initial experiences for using the MONACO and PINNACLE systems have been reported with a general observation that the biologically based optimization can generate plans with substantially better OAR sparing compared to DV-based optimization.[41,42] Readers are referred to vendor-provided manuals or training for more detailed descriptions about these systems. It should be noted that system upgrades may make some issues discussed below no longer relevant.

III.A. Parameter sensitivity

III.A.1. MONACO system

The biological models and model parameters used for plan optimization in MONACO system is included in Table I. Generally, appropriate selection of a model parameter is essential. The impact of changing model parameters in MONACO on optimized dose distributions is demonstrated using a test head-and-neck (hypopharynx) case irradiated with a 6 MV beam (Fig. 2). Only one parameter or isoconstraint at a time was changed. Figure 2(a) shows the result of changing cell sensitivity for PTV 70 (the PTV prescribed with 70 Gy) from the default value of $0.25\,Gy^{-1}$ to maximum and minimum allowable values of 1.0 and $0.1\,Gy^{-1}$, respectively. This parameter has a small effect on minimum dose in PTV 70, with greater values of cell sensitivity corresponding to larger minimum doses: 60.4 Gy for cell sensitivity of $0.1\,Gy^{-1}$, 62.5 Gy for $0.25\,Gy^{-1}$, and 65.7 Gy for $1.0\,Gy^{-1}$. Because the Poisson cell kill model is only an objective, changing prescription dose for PTV 70 without any changes in dose-limiting constraints does not affect target DVHs (data not shown). Behavior of the Serial complication model has been investigated on an example of spinal cord Planning organ-at-risk Volume (PRV), which was defined as 5 mm expansion around the spinal canal. Increasing power law exponent parameter or decreasing EUD isoconstraint for the Serial complication model applied to the spinal cord PRV results in lower maximum doses for the cord accompanied by some deterioration in target coverage [Figs. 2(b) and 2(c), dashed lines], and vice versa [Figs. 2(b) and 2(c), dotted lines]. Changing power law exponent parameter for the Parallel complication model applied to the parotid gland from a 3.9 value suggested in MONACO reference documentation to the lowest allowable value of 1.0 [Fig. 2(d)] does not result in any discernable trend. Smaller reference dose [Fig. 2(e)] or mean organ damage [Fig. 2(f)] values in the Parallel complication model lead to lower mean doses to the parotid gland. In the example chosen, reducing the reference dose or mean organ damage for the left parotid gland results in substantial sparing of that gland with no deterioration of target coverage. However, mean dose to the other parotid gland somewhat increases (not shown).

III.A.2. PINNACLE system

The Biological models and model parameters used for plan optimization and evaluation in PINNACLE system are

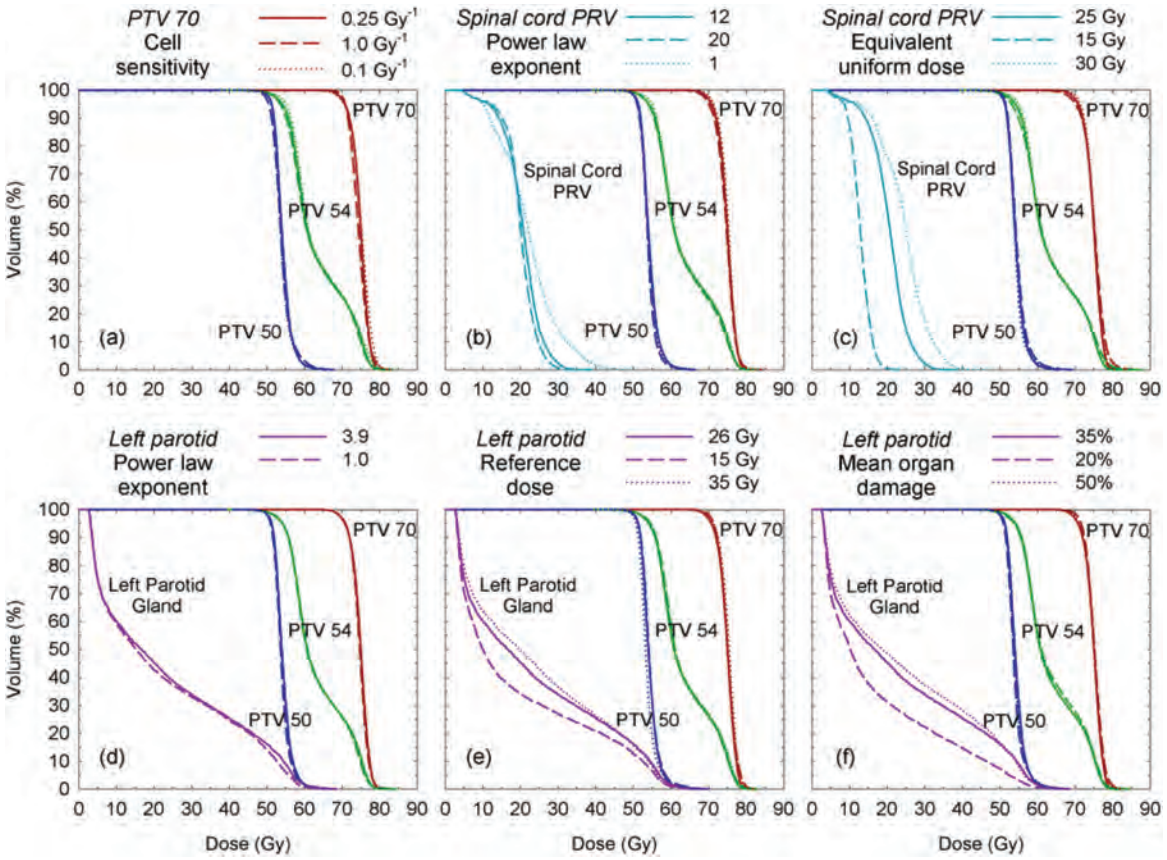

FIG. 2. Sensitivity of dose distributions obtained with CMS MONACO to changes in parameters or isoconstraints. Solid lines show base plan DVHs, dashed and dotted lines show DVHs obtained by varying each parameter/isoconstraint in either direction from its base value. For clarity, only DVHs for an affected OAR (spinal cord PRV or left parotid gland) and for target volumes (PTV 70, PTV 54, and PTV 50) are shown.

tabulated in Tables II and III, respectively. Sensitivity of dose distributions obtained with PINNACLE to changes in volume parameter (same as power law exponent in MONACO) or EUD (Table II) is shown in Fig. 3 for the same head-and-neck case as in Fig. 2. Optimization goals for PTV 70 were created using the Target EUD cost function combined with the Uniformity constraint. Variations in the volume parameter specified with Target EUD cost function have very small effect on PTV 70 DVH [Fig. 3(a)]. Minimum dose to PTV 70 slightly increases as the volume parameter decreases: 62.1 Gy for $a = 50$, 62.4 Gy for $a = -10$, and 64.6 Gy for $a = -50$. Increasing the EUD for PTV 70 has an effect of shifting the DVH toward higher doses [Fig. 3(b)]. Increasing the volume parameter for the Max EUD cost function applied to the spinal cord PRV from 12 to 20 results in the reduction of maximum dose to the cord and some deterioration of target coverage [Fig. 3(c)]. Decreasing the volume parameter from 12 to 1 has little, if any, effect on the dose distributions. Maximum dose to the cord decreases as the EUD value specified for the Max EUD cost function is decreased [Fig. 3(d)]. Minimum dose in a PTV closest to the spinal cord PRV (PTV 54 in this example) also decreases in proportion to the EUD for the cord. Similar trends were observed for the Max EUD cost function applied to the left parotid gland, i.e., increasing the volume parameter or decreasing the EUD reduces dose to the parotid gland and creates cold spots in adjacent target volume (PTV 54). Decreasing the volume parameter or increasing the EUD slightly increases dose to the parotid gland with virtually no effect on target coverage [Figs. 3(e) and 3(f)].

III.A.3. ECLIPSE system

The biological models and model parameters used for plan optimization and evaluation in ECLIPSE system are presented in Tables IV and V, respectively. The sensitivity of

TABLE II. Biological models used for treatment plan optimization in Philips PINNACLE.

Structure type	Name	Parameters	Objectives/constraints (Gy or cGy)	Comments
Target	Min EUD	Volume parameter ($a < 1$)	EUD	Penalizes for too low EUD
Target	Target EUD	Volume parameter ($a < 1$)	EUD	Penalizes for any deviation from the desired EUD
OAR	Max EUD	Volume parameter ($a \geq 1$)	EUD	Penalized for too high EUD; can be used with both serial and parallel structures

TABLE III. Biological models used for treatment plan evaluation in Philips PINNACLE.

Tool	Structure type	Name/description	Parameters/inputs	Comments
NTCP/TCP editor	Target	Empirical TCP model	D_{50}, m	Sigmoid curve represented by the CDF of the normal distribution
	OAR	Lyman-Kutcher model	D_{50}, m, n	Database of model parameters is provided
Biological response panel	Target	Poisson/LQ-based TCP model	D_{50}, γ, α/β	Database of model parameters is provided
	OAR	Källman s-model	D_{50}, γ, α/β, seriality (s)	Database of model parameters is provided
	Multiple targets	Composite TCP	TCP for individual targets	$\mathrm{TCP} = \prod_i \mathrm{TCP}_i$
	Multiple OARs	Composite NTCP	NTCP for individual OARs	$\mathrm{NTCP} = 1 - \prod_i (1 - \mathrm{NTCP}_i)$
	Targets and OARs	Probability of complication-free tumor control	Composite TCP, composite NTCP	$P_+ = \max(\mathrm{TCP} - \mathrm{NTCP}, 0)$

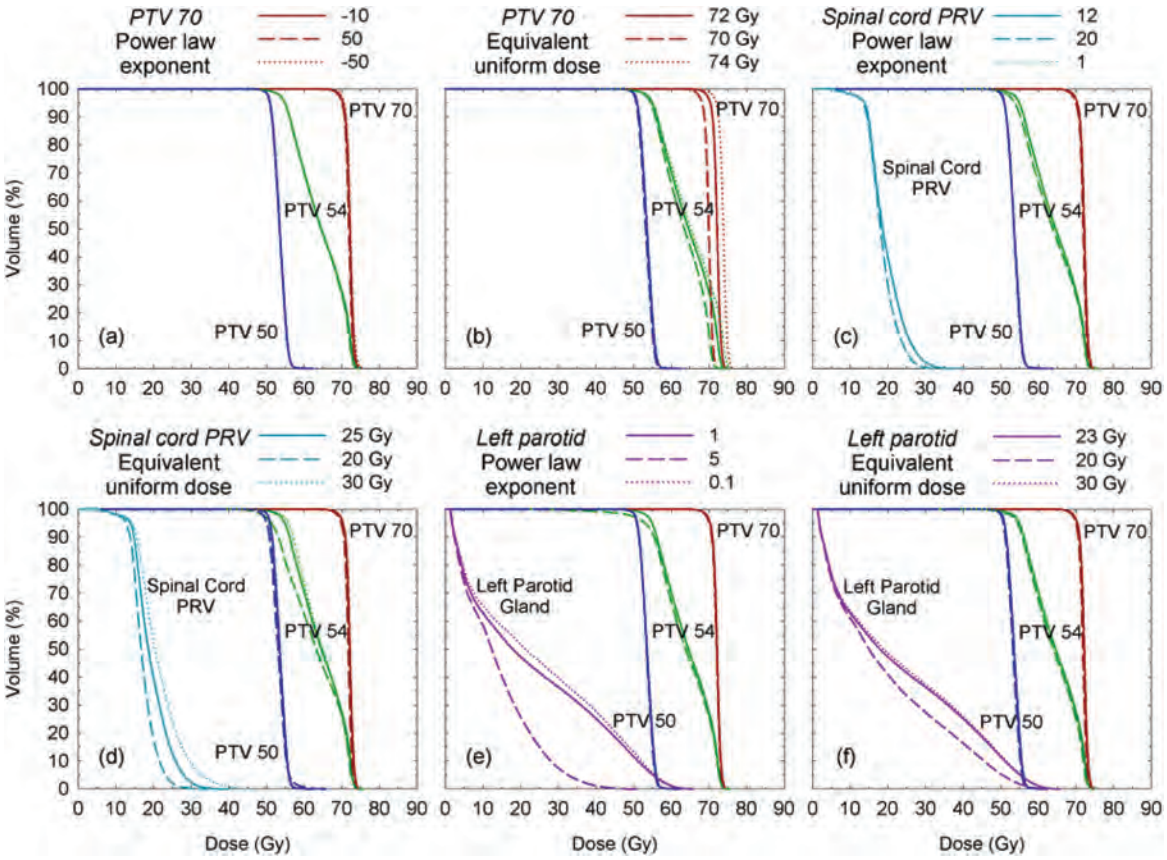

FIG. 3. Sensitivity of dose distributions obtained with Philips PINNACLE to changes in volume parameter (Power law exponent) or EUD. Solid lines show base plan DVHs, dashed and dotted lines show DVHs obtained by varying each parameter in either direction from its base value. For clarity, only DVHs for an affected OAR (spinal cord or left parotid gland) and for target volumes (PTV 70, PTV 54, and PTV 50) are shown.

TABLE IV. Biological models used for treatment plan optimization in Varian ECLIPSE.

Structure type	Name	Parameters	Objectives/constraints	Comments
Target	Min EUD	Volume parameter (a)	EUD (Gy or cGy)	Penalizes for too low values. Cannot be weighted. Listed under physical functions
Target or OAR	Max EUD	Volume parameter (a)	EUD (Gy or cGy)	Penalizes for high values. Cannot be weighted. Listed under physical functions
Target	TCP Poisson-LQ	D_{50}, γ, α/β, *seriality(s)*, $T_{1/2}$ *for short vs long repair time, % with long repair time, repopulation times: Tpot and Tstart*	TCP	Penalizes for small values. Can be weighted
OAR	NTCP Poisson-LQ	D_{50}, γ, α/β, *seriality(s)*, $T_{1/2}$ *for short vs long repair time, % with long repair time*	NTCP	Penalizes for large values. Can be weighted
OAR	NTCP Lyman	D_{50}, m, n,α/β, $T_{1/2}$ *for short vs long repair time, % with long repair time*	NTCP	Penalizes for large values of NTCP. Can be weighted

TABLE V. Biological models used for treatment plan evaluation in Varian ECLIPSE.

Tool	Structure type	Name	Parameters	Comments
Biological evaluation	Target	TCP Poisson-LQ	D_{50}, γ, α/β, *seriality(s)*, $T_{1/2}$ *for short vs long repair time, % with long repair time, repopulation times: Tpot and Tstart*	User selectable parameters or from database of model parameters
	OAR	NTCP Poisson-LQ	D_{50}, γ, α/β, *seriality(s)*, $T_{1/2}$ *for short vs long repair time, % with long repair time*	User selectable parameters or from database of model parameters
	OAR	NTCP Lyman	D_{50}, m, n,α/β, $T_{1/2}$ *for short vslong repair time, % with long repair time*	User selectable parameters or from database of model parameters

DVHs to changes in of model parameter, e.g., gEUD volume parameter (Power law exponent) or to threshold levels for the gEUD specified in the optimizer, were tested in ECLIPSE in a similar way as that discussed for MONACO and PINNACLE. Results of these tests are illustrated in Fig. 4. Optimization goal for PTV 70 were created with a maximum dose constraint and a minimum EUD constraint, instead of a minimum dose constraint. The maximum dose constraint was selected over the use of a uniformity constraint since it provided more reliably consistent results. If only a gEUD constraint is used, without a max dose or uniformity constraint, then the maximum dose in the PTV 70 volume rose to >100 Gy. This is consistent with the discussion in Sec. III A 2, on the lack of sensitivity to hot spots for TCP and gEUD optimization constraints for target tissues. In the creation of the data in Fig. 4, only one parameter was changed at a time. For example, Fig. 4(a) shows the result of changing Power law exponent for PTV 70 from the default value of −10 to values of −1 and −50. It is seen from Fig. 4(a) that variations in DVHs for PTV 70, PTV 54, and PTV 50 was relatively insensitive to the magnitude of the volume parameter. V_{70} was 97.6, 98.5, 99.2, and 99.5% for a values of −1, −5, −10, and −50, respectively. Figure 4(b) presents the result for varying gEUD for PTV 70 that changes the DVHs for PTV 70 and PTV 54. Figure 4(c) demonstrates that larger values of the Power law exponent ($a > 10$) have similar effects on the DVHs, while the D50% for cord PRV was 28, 25, and 25 Gy for $a = 1$, 12, and 20, respectively. Increasing the threshold for maximum gEUD shifted the cord DVH uniformly toward the threshold dose [Fig. 4(d)]. Since the

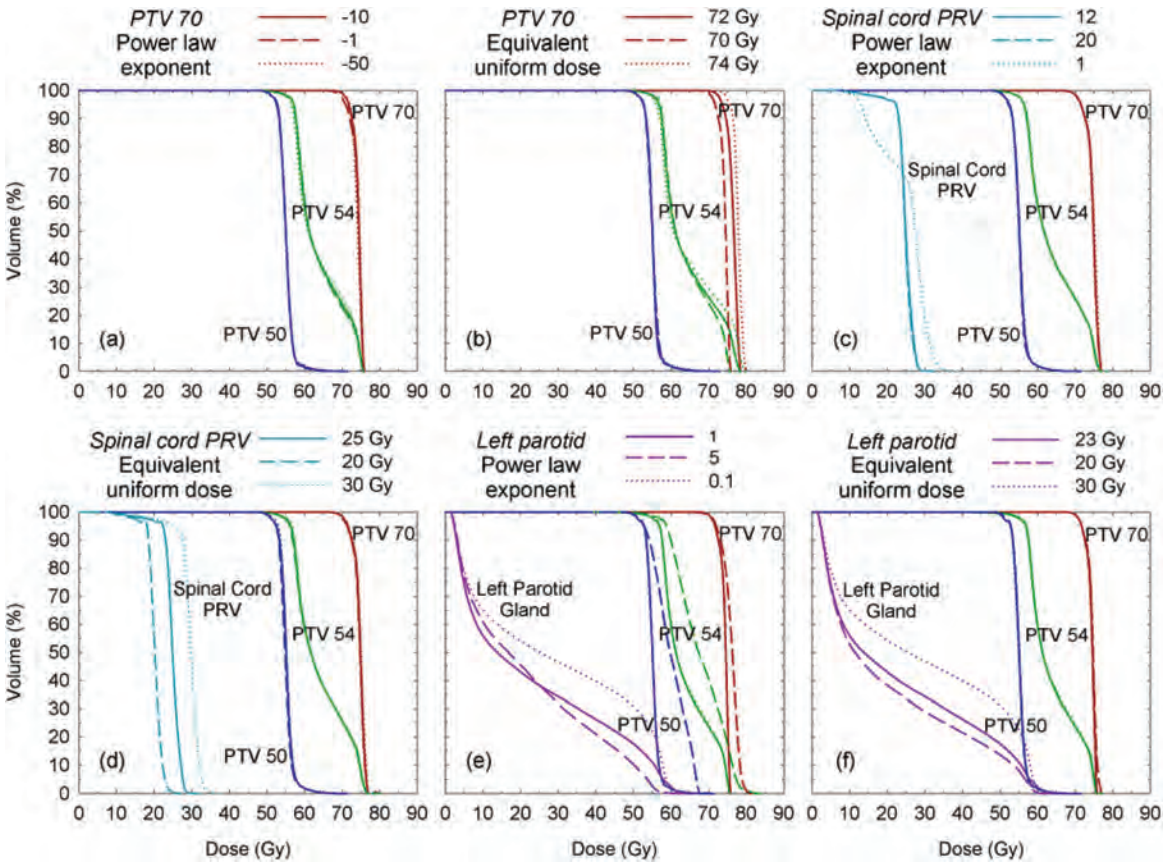

FIG. 4. Sensitivity of dose distributions obtained with Varian ECLIPSE to changes in volume parameter (Power law exponent) or EUD for elected target and OARs. Solid lines show base plan DVHs, dashed and dotted lines show DVHs obtained by varying a parameter from its base value, for example, (a) changing the Power law exponent for PTV 70 as indicated by the headings. For clarity, only DVHs for an affected OAR (spinal cord or left parotid gland) and for all target volumes (PTV 70, PTV 54, and PTV 50) are shown.

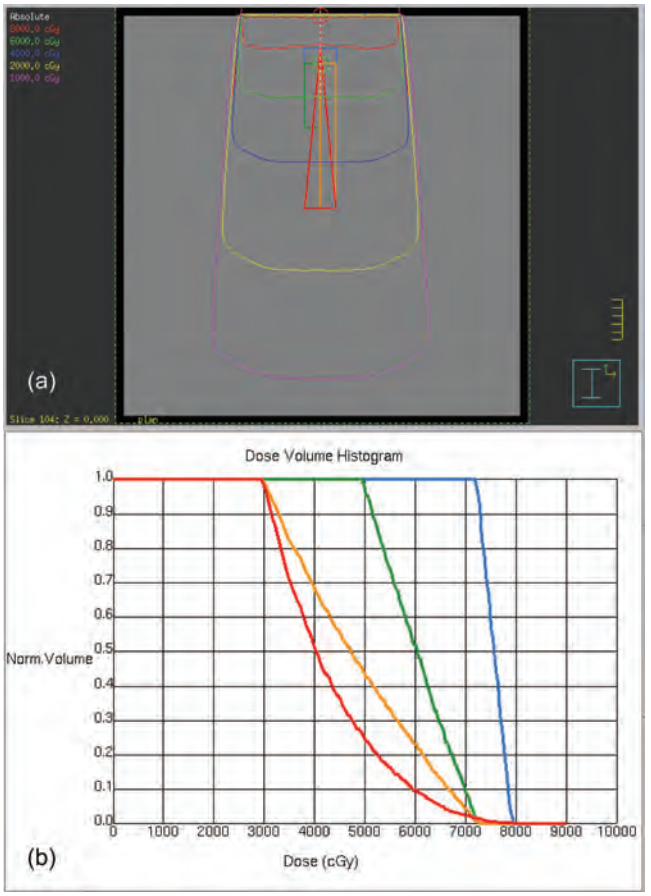

FIG. 5. Benchmark case for verification of EUD, NTCP, TCP, and P_+ calculations in PINNACLE. (a) Beam setup and structures. (b) DVHs.

parotid abuts the target volumes, interaction of constraint on the normal tissue affecting the target dose was noticeable. Sensitivity to the value of the Power law exponent was significant for DVH of the parotid [Fig. 4(e)]. Values of $V_{45} = 36\%$, 22%, and 16% for the parotid were obtained for $a = 0.1$, 1, and 5, respectively. Mean doses were similar for $a = 1$ and 5. The significant changes in the DVHs for the target volumes (PTV 70, PTV 54, and PTV 50) were observed with the large volume parameter value (i.e., $a = 5$). Fixing the volume parameter at $a = 1$ for the parotid and then varying the gEUD threshold increase the mean doses [Fig. 4(f)].

Constraints using NTCP and TCP values may also be used to shape the dose distribution. This is more complex owing to the number of input parameters and, in principal, produces similar results to gEUD constraints as outlined in Sec. II A 3. To obtain a clinically desirable dose distribution using NTCP and TCP based cost functions, it is often necessary to use parameters for these functions quite different from those used to evaluate their value. For example, using a D_{50} much lower for optimizing the parotid dose can produce a better final dose distribution than obtained using the same value of D_{50} used to evaluate the NTCP of the plan.

III.B. Comparison of three systems

III.B.1. Comparison and verification of EUD, NTCP, TCP, or P_+ values obtained with PINNACLE and ECLIPSE systems

Both the PINNACLE and ECLIPSE systems provide tools to calculate EUD, TCP, NTCP, or P_+ as plan evaluation metrics. To verify this calculation, selected metrics have been calculated "manually" in an independent spreadsheet (Microsoft Excel) using the same models and model parameters. Readers are referred to the TG full report for the details of this spreadsheet. The results obtained with the spreadsheet are compared to the same quantities reported by PINNACLE and ECLIPSE. All work in this section has been performed using a benchmark case, which involves a single 6 MV, 20×20 cm^2 photon beam incident on a sufficiently large cubical water phantom at 100 cm source-to-surface distance. A dose of 72 Gy in 40 fractions was prescribed to a point at 6 cm depth along the central axis. Four simple structures (three rectangular, one triangular) were created inside the phantom [Fig. 5(a)]. The DVHs for these structures look similar to those encountered in a typical plan [Fig. 5(b)]. This case was chosen because it could be easily and reproducibly set up in any commercial TPS, providing a simple, nearly identical input dose distribution for comparison of biological evaluation tools between different TPS. Use of a PDD-based dose distribution and simple structures also facilitates spreadsheet based, hand calculations of DVH, TCP, and NTCP values for comparison. Details on the geometries of the structures in the benchmark phantom are provided in Table VI.

In a DVH view page of the PINNACLE system, users can specify parameter a for each structure, and the system will calculate and report gEUD in the ROI Statistics section in addition to physical quantities. By default, $a = 1$, and the reported gEUD is the same as the mean dose. Parameter a was varied in the range between -50 and 50, and gEUD values reported by the TPS for all four structures were recorded. The same gEUD values were then calculated on the spreadsheet based on DVHs exported from PINNACLE. Both sets of data agreed very well (<0.1%) for all combinations of the structure and volume parameter.

TABLE VI. Details of benchmark phantom structures.

Structure	Dimensions	X coordinate range	Y coordinate range (cm)	Depth range (cm)
PTV rectangle	$4 \times 4 \times 2$ cm^3	-2 to 2 cm	-2 to 2	4–6
Rectangle 1	$2 \times 4 \times 8$ cm^3	-2 to 0 cm	-2 to 2	6–14
Rectangle 2	$2 \times 2 \times 18$ cm^3	0–2 cm	-2 to 2	6–24
Triangle 1	4×4 (base) $\times 20$ cm^3	0 cm $\times$ 0 cm at depth = 4 cm -2 to 2 cm at depth = 24 cm	-2 to 2	4–24

TABLE VII. TCP and NTCP values calculated for DVHs obtained in the benchmark phantom. Values for TCP and NTCP were calculated using a dose distribution calculated in the PINNACLE system (Sec. III B 1). Variability among institutions in reproducing the phantom and differences in 6 MV photon beams will produce small interinstitutional differences in the calculated values.

Structure	PTV rectangle	Rectangle 1	PTV rectangle	Rectangle 1	Rectangle 2	Triangle 1
D50 (Gy)	63.3	44.2	80	75.1	55.3	46
γ	5	1.6	3	2.8	3.1	1.8
α/β (Gy)	10	10	3	3	3	3
Seriality	N/A	N/A	0.18	8.4	0.69	1
Function	TCP	TCP	NTCP	NTCP	NTCP	NTCP
Value (%)	94.1	80.3	26.6	18.1	23.5	29.5

To spot-check results reported in the Biological Response panel from PINNACLE, calculations were performed for TCP for two structures and NTCP for all four structures. Input parameters and calculated values are detailed in Table VII. Model parameters were chosen from the included library so that resulting TCP and NTCP estimates are not equal to 0 or 1. All individual TCP and NTCP values and composite metrics (composite TCP, composite NTCP, and P_+) reported by PINNACLE closely matched (<0.5%) the same quantities calculated manually using the spreadsheet. A close agreement (<1.5%) was also obtained between TCP and NTCP estimates reported by NTCP/TCP Editor (another tool in PINNACLE) and corresponding quantities calculated from the DVHs using the spreadsheet.

Similar calculations were performed using the evaluation tool in ECLIPSE with DVHs for the benchmark phantom generated from ECLIPSE. The TCP/NTCP values obtained with ECLIPSE were found to agree, within one percentage point, with those generated by PINNACLE (Biological Response panel). For example, TCP/NTCP values for one of the rectangular structures was 0.81/0.19 and 0.80/0.18 for PINNACLE and ECLIPSE, respectively.

The presence of two different tools to calculate TCP and NTCP in PINNACLE (Biological Response panel and NTCP/TCP Editor) offers the user greater flexibility. The TCP model implemented in the NTCP/TCP Editor is not widely used in the literature, and a database of model parameters is not provided, which is likely to steer users toward the Biological Response panel for all TCP calculations. On the other hand, the LKB model implemented in the Editor is much more widely used than the Poisson-based NTCP model from the Biological Response panel. For both models, the origin of parameters included in the library can be traced to the organ tolerance data of Emami *et al.*[41] It is therefore reasonable to hypothesize that the NTCP models in the Biological Response panel and NTCP/TCP Editor should provide similar estimates for the same endpoint and dose distribution. To test that, structures and endpoints for NTCP calculations in NTCP/TCP Editor were matched to those previously selected in the Biological Response panel. The two tools produce somewhat different values of NTCP, with the Editor giving smaller estimates for all structure/endpoint combinations. In one case, an NTCP estimate provided by the Editor was smaller by a factor of 2 than the corresponding estimate from the Biological Response panel (25% vs 50%). Users should be cautioned to use models and parameter estimates of unverified origin, if predicted TCP/NTCP values are to be used for making clinical decisions. Similar warning statement is provided in vendor's user manual. Alternatively, users can input their own or other validated data using the tools provided in the TPS.

III.B.2. Comparison of plans generated by MONACO, PINNACLE, and ECLIPSE systems

Treatment plans for three representative test cases designed using MONACO, PINNACLE, and ECLIPSE are compared in Fig. 6. The head-and-neck case shown in Fig. 6 is the same that was used for the parameter sensitivity studies (Sec. III A). The EUD (and gEUD) values computed based on these DVHs for selected organs are tabulated in Table VIII. The EUD/gEUD values were calculated based on the power law and the parameter a used is included in the table. All MONACO plans resulted in substantially less homogeneous target dose distributions compared to the PINNACLE and ECLIPSE plans. This result is a consequence of the compulsory cell killing objective which penalizes small cold spots less drastically than physical minimum dose penalties and the practice of normalizing the treatment plans to the minimum dose in the target, not to equivalent target EUD (Notice that MONACO allows also users to supplement the target EUD cost function with physical constraints). In terms of OAR sparing, the three TPS produced plans of similar quality for the head-and-neck case, with the exception of the spinal cord PRV dose in the ECLIPSE plan. In the prostate case, PINNACLE offers somewhat better sparing of the rectum and pubic bone at low doses. It has been previously demonstrated that one can use MONACO to generate plans with substantially better OAR sparing compared to the plans designed using DV-based TPS XiO.[42] The same trend has been observed for plans created in PINNACLE using biological vs DV-based cost functions.[43]

IV. ACCEPTANCE, COMMISSIONING, AND ROUTINE QA TESTS FOR BIOLOGICALLY BASED PLANNING SYSTEMS

The acceptance, commissioning, and periodic QA tests recommended by previous TG reports (TG-40, TG-53, TG119, and TG-142) (Refs. 44–47) for general TPS features (e.g., data input/output, dose calculations, plan deliverability, clinical software tools) should be performed for BBTPS. These recommendations will not be repeated here. The aim

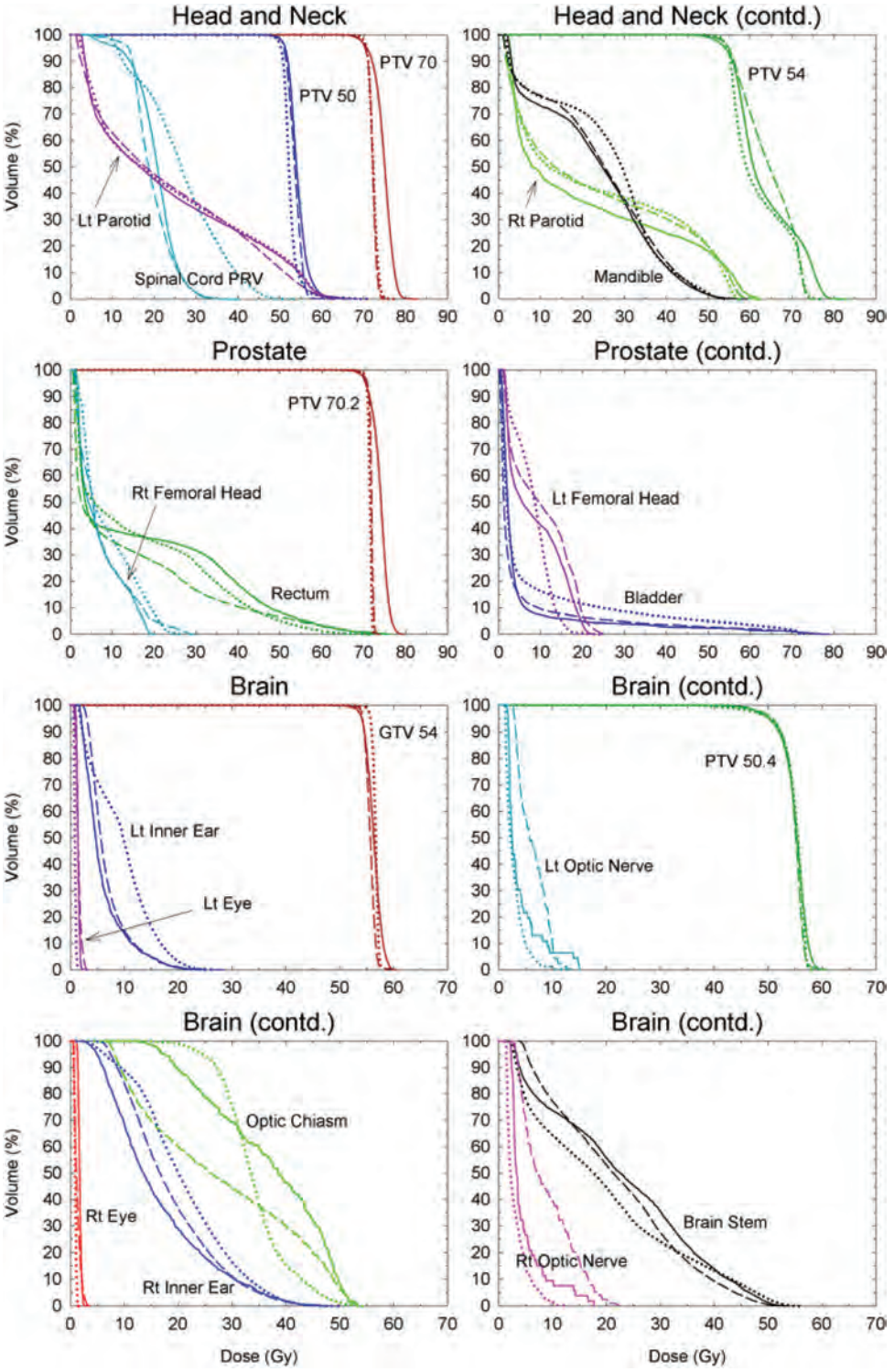

FIG. 6. Dose-volume histograms for head-and-neck, prostate, and brain cases obtained using biologically based optimization features implemented in CMS MONACO (solid lines), Philips PINNACLE (dashed lines), and Varian ECLIPSE (dotted lines).

of this section is to provide examples of the additional QA tests specific to general features pertinent to BBTPS. Physicists are encouraged to develop their own QA tests specific to the BBTPS in question. It is expected that more complete list of such QA tests will become available as more practical experience is gained in the future. All acceptance and commissioning tests should be carried out on the system after it has been installed in the clinic but before it is used clinically.

IV.A. Acceptance tests

Based on the previous TG reports,[44–47] an acceptance test is performed to confirm that the TPS performs according to its specifications. Because the quantifiable and testable specifications are generally lacking for a BBTPS at the present time, the acceptance testing may be limited to verifying functionalities offered by the BBTPS. Examples of the

TABLE VIII. EUD (Gy) values calculated based on the DVHs for three sample plans along with the parameter *a* used.

Case	organs	*a* value	monaco	pinnacle	eclipse
Head-and-Neck	PTV 70	−10	74.40	71.99	72.15
	PTV 54	−10	59.59	61.04	59.01
	PTV 50	−10	53.79	53.25	52.13
	Cord PRV	20	27.19	27.77	39.91
	Lt parotid gland	1	23.12	23.12	23.61
	Rt parotid gland	1	20.57	23.05	23.11
	Mandible	10	38.52	38.66	38.78
Prostate	PTV 70.2	−10	73.61	71.77	71.41
	Rectum	8	46.11	44.61	41.30
	Bladder	8	43.21	43.00	45.79
	Lt femoral head	12	14.76	20.09	18.21
	Rt femoral head	12	16.86	18.43	13.90
	Pubic bone	12	41.25	41.19	49.06
Brain	GTV 54	−10	56.08	55.55	56.56
	PTV 50.4	−10	54.11	53.90	53.26
	Brain stem	16	41.61	40.06	42.80
	Optic chiasm	16	46.11	44.99	41.68
	Lt eye	16	1.94	2.29	1.08
	Rt eye	16	2.04	2.45	1.13
	Lt optic nerve	16	12.33	10.12	8.68
	Rt optic nerve	16	14.07	17.16	8.61
	Lt inner ear	16	16.28	18.60	18.22
	Rt inner ear	16	34.36	34.00	34.57

functionality that might be included in the acceptance testing are: (1) allowing user to update model and/or model parameters, (2) allowing user to specify model-based goals and constraints for optimization, and (3) allowing user to evaluate plan based on biological metrics.

IV.B. Commissioning tests

IV.B.1. Selective verification of biological metrics

Biological metrics, i.e., EUD/TCP/NTCP, calculated within a TPS should be independently verified by the user for selected cases before routine clinical use and after major upgrades. The benchmark phantom used in this report (Sec. III B 1) may be used in this verification. If fully validated by the medical physics community, other research software tools, such as CERR (Refs. 48 and 49), BIOPLAN,[50] or BIOSUITE,[51] may be used as evaluation tools. TPS vendors are urged to provide detailed descriptions of all implemented models and parameter values to make this possible. It is also recommended that TPS vendors provide tools to export dose distributions, DVHs and 3D dose matrices to external programs. Note that results may differ whether EUD/TCP/NTCP is calculated using the dose grid or DVH. TPS documentation should clearly state the technique used to calculate all biological metrics.

IV.B.2. Double planning

It is recommended that for first several cases from each representative tumor site, new users of BBTPS prepare second plans using their standard DV-based TPS. New and traditional plans can be compared to help understand how different aspects of a dose distribution are controlled by biological cost functions. Preparation of backup plans may be discontinued after sufficient expertise in treatment planning and knowledge of advantages/limitations the BBTPS has been gained.

IV.B.3. Compilation of benchmark 3D datasets and DVHs for major sites

The TG has prepared sample plans for the benchmark phantom and three test cases representing major sites often treated using IMRT: head-and-neck, prostate, and brain. Image and structure sets in the DICOM format as well as tabular DVH data for these three cases can be downloaded from the websites in Refs. 52–55. Note that these sample plans do not necessarily represent the best plans possible for the cases. They are provided for comparison purposes. Physicists attempting to commission a BBTPS may compare plans obtained with their system to these sample plans and may explore whether similar or better plans can be obtained with the BBTPS in question. If applicable, the EUD values provided for these sample plans in Table VIII may be used to validate the EUD calculation in the testing BBTPS.

IV.C. Procedures for routine QA

It is suggested that users of biologically based TPS prepare an IMRT sample plan generated by the plan optimization based on either one of the benchmark cases or the user's own case at the time of commissioning. This sample case should be replanned annually or after major upgrades and compared to the baseline data, i.e., DVHs, EUD/TCP/NTCP, to ensure that models, parameters, and algorithms implemented in the TPS plan optimization remain the same. For a BBTPS with a biologically based plan evaluation tool, users should prepare a 3D dose distribution based on one of the benchmark cases or the user's own case and obtain its baseline DV and biological metrics using the evaluation tool at the time of commissioning. These metrics should be re-evaluated annually or after major upgrades. The new metrics should be identical to the baseline data. For TPS capable of Monte Carlo dose calculations, a nonstochastic dose calculation algorithm should be used, if available, for initial and subsequent treatment planning to eliminate statistical uncertainties associated with the Monte Carlo method.

V. VISION OF THE TASK GROUP FOR FUTURE DEVELOPMENT OF BIOLOGICALLY BASED TREATMENT PLANNING

V.A. Evolution of biologically based treatment planning systems

A vision of the TG for current and future developments in BBTP is summarized in Table IX. The majority of existing TPS employ DV-based cost functions for treatment plan optimization. Plan evaluation is also performed using DV criteria, i.e., 3D dose distributions and DVHs. Although some of the existing systems provide tools for calculation of

TABLE IX. Evolution of biologically based treatment planning.

Evolution stage	Plan optimization strategy	Plan evaluation strategy	Representative TPS
0	DV-based cost functions	DVHs	All IMRT TPS
1	EUD for OARs; EUD- and DV-based cost functions for targets	DVHs and relative values of TCP/NTCP	CMS MONACO Philips PINNACLE Varian ECLIPSE
2	EUD-based cost functions for all structures	Absolute values of TCP/NTCP	Future developments
3	Absolute values of TCP/NTCP	Absolute values of TCP/NTCP	Future developments

TCP/NTCP with a purpose of plan evaluation, these tools are not well documented and are not supplied with databases of reliable model parameters, so therefore have not found a widespread use among radiation treatment planners or evaluators. This state of affairs is designated as stage 0 in the proposed BBTP evolution scheme.

Initial transition to stage 1 has occurred with an emergence of the first TPS employing EUD-based cost functions for plan optimization (Sec. III). TPS representative of this stage may provide framework only for biologically based optimization (e.g., CMS MONACO) or may also offer practical tools for biologically based plan evaluation (e.g., Philips PINNACLE). Incorporation of plan comparison tools based on EUD, TCP, and NTCP functions into commercial TPS is a welcome development. Provided that such tools are intuitive and easy-to-use, a greater number of radiation oncology professionals will be willing to gradually integrate biological metrics into their clinical practice, which will facilitate transition to stage 2 in the evolution of BBTP.

The principal difference between stages 1 and 2 is that TCP and NTCP functions will play a more important role in the treatment plan evaluation process and will supersede DV metrics as the major indicators of plan quality. As opposed to the use of TCP and NTCP in stage 1, where these estimates are used primarily in a relative fashion as an ancillary tool to compare alternative treatment plans, in stage 2 the growing confidence in predictive power of dose-response models will allow decisions about plan quality to be based on absolute estimates of TCP/NTCP. Because the effectiveness of a plan will be judged by the predicted biological outcome, dose-based constraints will no longer be required to force target dose to be as uniform as possible so long as the plan results in desired values of TCP and inhomogeneous target doses do not jeopardize intermixed normal tissues and, therefore, treatments could be optimized based on EUD-based cost functions alone. Stage 2 must be accompanied by clear supporting evidence for the reliability of each model used and may be adopted at different times for different disease sites.

Transition to nonuniform target dose distributions represents a major paradigm shift in treatment planning. The requirement for uniform dose delivery was a long-standing tradition in our field. It may have been based on the assumption that tumors behave as "parallel structures," and that the ultimate TCP will be related to the minimum dose. Seeking a uniform target dose that exceeds the desired minimum was a way to reduce the overall integral dose/exposure to the patient. This concept has also been rooted in convention (ICRU 50).[56] However, there is no proof that this construct is valid for tumors. Relaxing target uniformity constraint, and allowing hot spots within the target, may afford the planner increased flexibility in creating a better plan that may lead to better critical structure sparing.[23] Table X summarizes arguments in favor and against using heterogeneous dose distributions. While most of the supporting arguments are theoretical, there are some examples where nonuniform dose delivery has proven to be safe and effective, e.g., brachytherapy, intracranial SRS, simultaneous integrated boost. It is also known that, due to organ motion and daily setup uncertainties, actual delivered dose distributions are less homogeneous than the planned dose distributions, especially if small margins are used.[57] This implies that several generations of radiation oncologists have been treating patients using moderately inhomogeneous dose distributions.

The PINNACLE of BBTP (Table IX, stage 3) is represented by a scenario in which treatment plans will be optimized using objective scores based TCP/NTCP (e.g., Refs. 21 and 58). The optimized values of TCP and NTCP will be used directly to evaluate treatment plans.

V.B. Desired features and functionalities for future biologically based treatment planning systems

It is instructive to speculate as to what constitutes an ideal (optimal) BBTPS. Many preferred general features for TPS (e.g., fast and accurate dose calculation and optimization algorithms, same dose engine for optimization and for forward calculation, accurate DVH generation, robust and effective input and output tools) are also important for future BBTPS. In addition, the following characteristics and functionalities are desirable for a BBTPS:

(1) The system should allow the user to input or to update models and model parameters for both plan evaluation and optimization. A library of model parameters, containing the default parameter values with capability of allowing user to update these parameters based on specific clinical situation or local patient population, should be provided. The system should supply detailed documentation for the models and the default parameter sets (their origination, applicability and provenance). For example, the LKB NTCP model along with a database of parameters for common organs and endpoints, indicating the sources (e.g., QUANTEC), should be provided. For an updated model, inclusion of new biological or medical factors might introduce additional uncertainties. The user should be provided with tools to

TABLE X. Pros and cons of homogeneous vs heterogeneous tumor dose distributions.

	Homogeneous dose	Partial tumor boosts	Heterogeneous dose
Pros	• Proven track record • Consistent with classical radiobiology • Less ambiguity in reporting and analyzing delivered doses • Hot spots are probably useless unless they cover most of the tumor or the most resistant subvolume • Most logical approach if information about tumor heterogeneity is not available	• Predicted to be effective under a wide range of theoretical assumptions • Easy to do using IMRT • The PTV margin provides a natural 'draw-down' region between a GTV boost and critical normal structures	• Adds another degree of freedom to the treatment planning problem and can lead to better critical structure sparing • Hot spots may not be detrimental if they are located inside the GTV • Because tumors are heterogeneous (e.g., PET imaging) there is no reason that tumor dose should be uniform • Would allow biological targeting via the use of PET, etc. • Stereotactic and implant experience is supportive although in different dose/dose rate regimens
Cons	• Old school/tradition-driven • Opportunity to exploit tumor heterogeneities is lost • Opportunity to use the flexibility of IMRT is limited • Mounting data support the notion that heterogeneous dose may be advantageous	• No consistent human or animal data to confirm positive effects although supportive data can be found for certain tumor sites • Due to temporal changes in tumor volumes effect of partial tumor boost may be diminished	• No consistent human or animal data to confirm positive effects although supportive data can be found for certain tumor sites • Theoretical benefits are limited by the accuracy of TCP models • Effect of cold spots may be underestimated, especially if cold spots are located in the GTV • Must be careful that hot spots within the PTV are located within the GTV, and not the normal tissue margin.

do the calculations based on both the latest and previous models. As there may be competing models for a given clinical situation, the user should have a choice regarding which model should be used for the situation. There might be clinical reasons to trust one type of model over others in certain situations. The system should allow the user to consider organ interaction (e.g., interaction between heart and lung, or liver and kidney) by modifying model(s) and/or selecting model parameters. Note that some of these features are included in the existing TPSs (e.g., PINNACLE).

(2) The system should allow combinations of biological model- and DV-based constraints for all structures in the optimization. For example, one might want to use lung EUD (or Lyman model NTCP) but also keep V_{20} below 35%. The maximum dose may be used as hard constraint to limit hot spot. For some organs, the user may choose to use DV constraints only (For example, the uniform dose that approaches the tolerance dose for brainstem may be used in the overlap region of brainstem and glioma PTV). For certain situations, a hybrid approach, e.g., using DV-based optimization followed by biological-based optimization,[59] may be helpful.

(3) The optimizer should "reward" getting lower NTCP than requested, if that is possible without violating the higher priority goals or preventing lower priorities from being achieved, and should allow maximizing TCP for a given NTCP or minimizing NTCP for a given TCP.

(4) The system should allow the user to define "stop values" for iteration during plan optimization. For example, the plan optimization may attempt to lower EUD or NTCP for normal tissue no matter what they are. The user-defined value below which further optimization is unnecessary would improve efficiency of the optimization.

(5) The user should be able to input "dose-modifying factors" to account for the effects of certain medical factors such as the use of chemotherapy, pre-existing conditions (e.g., diabetes), and tobacco use or biological factors such as genetic predispositions to a complication. For example, if chemotherapy is known to be a factor, the user can choose either using a different set of model parameters, or introducing a dose-modifying factor. The system should provide the user with the opportunity to assess the potential variations in outcome with changes in these factors. For example, estimations such as "this plan is anticipated to increase the patient's risk of lung cancer by xx%, and this risk can be reduced to yy% if the patient discontinues smoking tobacco" could be informative.

(6) The system should have connectivity with medical information systems such that clinical parameters that might be used in optimization (e.g., pulmonary function tests) can be accurately and directly incorporated. Similarly, clinical information within the medical information systems may be needed to determine which

predictive model(s) to use. For example, it might be that the predictive models for pneumonitis might be different for patients with cancers of the lung vs breast.

(7) The system should provide an option to assess the outcomes of a tentative plan reflective of anticipated delivery uncertainties including set-up errors, interfractional and intrafractional anatomic variations. The system should allow to build a population-based or patient-specific probability distribution into the evaluation of TCP and NTCP, perhaps with "confidence limited" NTCPs and TCPs for given population-based or patient-specific motion probability distribution.

(8) For a given treatment, the best achievable dose distribution depends on the patient geometry and the physical limitations of the radiation in question. For example, the minimal dose to an OAR adjacent to the PTV mostly depends on the distance between the OAR and PTV and the physical characteristics of the radiation beam and the degradation in target coverage (TCP) that the physician is willing to accept. It is desirable that the treatment planning system can predict the best possible plan, thus can estimate best possible NTCP conditional on a chosen TCP, avoiding the unnecessary effort in search for nonachievable plans.

(9) Models should allow for spatial radiobiological variations (e.g., clonogenic cell density, radiosensitivity, hypoxia) linking to biological/functional images. The planning system should provide the "painting-by-number" feature. That is, user could generate highly nonuniform dose distribution required based on the spatially varied biological/function information. Plan evaluation tools, such as the functional DVH (fDVH) (Refs. 60 and 61) that can take into account the spatial distribution of the functional importance, should be provided.

(10) Outcome models should be able to consider time effects (e.g., treatment breaks, fractionation, tumor growth, and delivery time). The system should have the capability to optimize around a prior dose distribution either from external beams or from brachytherapy. Validated deformable registration for calculating delivered dose from previous treatment courses should be available.

(11) The system should have user friendly graphic tools to show information during the optimization process. This information includes tables with constraints and priorities, graphs displaying components of cost function (biological-based and/or DV-based constraints), DV and/or biological indices as in the current iteration.

(12) For plan evaluation, the calculation of biological indices should be accompanied by parameter sensitivity analysis. Ability to renormalize/adjust a plan to achieve biological goals is also useful. For example, in addition to being able to renormalize the plan according to DV criteria (e.g., 98% of PTV receives at least 95% prescription dose) renormalize to achieve an acceptable value of NTCP or EUD for a given OAR.

VI. SUMMARY OF RECOMMENDATIONS AND PRECAUTIONS FOR CLINICAL USE OF BIOLOGICALLY BASED MODELS

VI.A. General recommendations

(1) Biologically based cost functions for OARs may be preferable to DV constraints because the former typically controls entire portions of the DV domain whereas the latter controls only a single point. For OARs requiring more than one DV constraint for inverse treatment planning, it may be preferable to replace multiple constraints by a single EUD-based cost function with appropriate choice of a volume effect parameter. Because a biological cost function controls greater space in the DV domain, it may be more effective in optimizing a plan toward OAR sparing as compared to the use of DV constraints.

(2) Currently implemented biological cost functions for target volumes control only cold spots. These functions are not essential to obtain good quality plans and may be replaced with minimum dose constraints on target dose even in biologically based optimization.

(3) Biological cost functions for target volumes do not effectively control hot spots. Despite some evidence in favor of less homogeneous target dose distributions, the TG maintains that highly nonuniform dose distributions caused by the optimization technique (as opposed to deliberate and tested nonuniformity as is seen in SRS, simultaneously integrated boost and brachytherapy) should be avoided. To obtain clinically acceptable plans with respect to target dose homogeneity, biological cost functions should be supplemented with maximum dose-type physical cost functions.

(4) At present, the plan evaluation should be performed based on established DV criteria. Therefore, biologically based TPS must present physical parameters (i.e., DVHs, minimum, maximum, mean dose) along with any biological metrics. EUD can be used to rank plans provided the parameter a is calibrated appropriately. Relative estimates of biological indices (i.e., TCP, NTCP) may be used to help select rival plans, provided that the users understand the range of applicability of models and parameter values implemented in their TPS. Use of absolute estimates of TCP/NTCP as main indicators of plan quality is not warranted at this time.

(5) Regardless of advancement of biological tools for treatment planning, the TG recommends that review of 3D dose distributions always remain a part of the treatment plan evaluation process. TPS should allow the dose to be viewed in multiple planes (e.g., axial, coronal, and sagittal, as well as nonstandard planes and an overall volumetric display) since hot spots in nonspecific tissue and heterogeneities in physiology/function within tumor/normal tissues are sometimes more clearly understood in these other planes.

(6) If the parameter a cannot be calibrated in the calculation of EUD, the following generic values may be used as

starting values: $a=1$ for a parallel organ and $a=8$ for a serial organ. An uncertainty analysis should be performed by calculating a type of confidence intervals around the calculated EUD values, for example, by calculating the lower and upper bounds on the EUD using $a=0.5$ and $a=3$ for parallel structures, and $a=4$ and $a=15$ for a serial structures.

VI.B. TPS-specific recommendations

The recommendations below apply to the specific system versions described in Sec. III.

VI.B.1. CMS MONACO

(1) Commonly found values for the cell sensitivity of tumors of $\sim 0.25\,\mathrm{Gy}^{-1}$ frequently do not penalize cold spots as forcefully as demanded by clinical practice, so that higher values of cell sensitivity may have a greater utility, if less biological meaning. However, for the selected test case (Sec. IV A 2), the choice of cell sensitivity parameter had minimal impact on minimum PTV dose.
(2) The Poisson statistics cell kill model should always be used in conjunction with a physical constraint, either the maximum dose or quadratic overdose penalty. Failure to specify the second cost function results in convergence problems and long running times.
(3) Reasonable starting points for EUD constraints can be derived by an EUD computation for a number of acceptable dose distributions in each institution. Typically, without biological optimization the EUD values will be spread out over an interval of values. A good starting point could be the median of this distribution, meaning that half of the patient could get a lower EUD by means of biologically constrained optimization.

VI.B.2. Philips PINNACLE

(1) By adjusting the volume parameter, the Max EUD objective can be used to specify optimization goals for all types of OARs.
(2) In the case of a single PTV, combining the Target EUD objective with the Uniformity constraint yields good results.
(3) For plans with multiple PTVs, using Min dose and Max dose cost functions offers better control over target dose distributions.

VI.B.3. Varian ECLIPSE

(1) Use of Min/Max dose or percentage structure volume at dose provides more reliable control over target dose distributions. Effect of adding EUD or TCP models should be carefully monitored to avoid introducing target dose heterogeneities that would not be accepted in clinical plans.
(2) Adjusting TD50 and n values in the LKB model can be used to shape evolution of OAR DVH in optimizer.
(3) Adjusting volume parameter and target EUD can be used to shape DVH on OARs.

ACKNOWLEDGMENT

The authors wish to thank Dr. Moyed Miften and Dr. Shiva Das for their invaluable comments for improving this report.

APPENDIX A: THE FOLLOWING TERMS AND DEFINITIONS ARE USED IN THIS REPORT

A.1. Plan optimization or inverse planning

The process of generating an optimal plan follows the desired objectives. The planner specifies objectives (i.e., optimization criteria) including constraints (limits that should not be violated) and goals for both the target and normal structures. Internally, the planning system represents these objectives in a cost function, which must be maximized or minimized by an optimization algorithm.

A.2. Tumor control or local control

No evidence of tumor recurrence in the region treated with a definitive intent.

A.3. Normal tissue complication

An unfavorable symptom, sign or disease temporally associated with the use of radiation therapy. The Common Terminology Criteria for Adverse Events (CTCAE) version 3.0 (Ref. 62) or other scale[63] are commonly used for grading normal tissue complications.

A.4. Functional subunit (FSU)

Structurally or functionally discrete tissue elements,[64] e.g., nephrons in kidneys or alveoli in lung.

A.5. Volume effect

Modification of normal tissue/organ tolerance with a change in irradiated volume. That is, the tolerance dose increases to a degree that depends on the tissue and complication endpoint as the irradiated volume decreases. The magnitude of the volume effect depends, at least in part, on an underlying anatomic/biological structure of the organ.

A.6. Parallel organ

Normal organ in which FSU functions relatively independently and damage to a sufficiently small region does not render the whole organ dysfunctional. Consequently, a volume threshold or functional reserve may exist. Examples are lung, kidney, and liver.

A.7. Serial organ

Normal organ in which FSU are structured in a series, if one FSU is incapacitated the organ will exhibit complications. Examples are spinal cord, intestines, and optic nerve.

A.8. Serial and parallel response

Normal organ response characterized by small and large volume dependence, respectively. The connection between organ architecture designated as serial or parallel (see above) and serial parallel response is not always exact. Nevertheless, within the framework of this report, an exact correspondence between organ architecture and response is assumed. This means that in parallel organs complications are assumed to occur after a substantial fraction of FSU are damaged, volume effect is large and response is "parallel." In contrast, a serial organ would exhibit complications after a single FSU is incapacitated, volume effect is small and response is "serial".

APPENDIX B: DESCRIPTION OF CURRENTLY AVAILABLE THREE TREATMENT PLANNING SYSTEMS EMPLOYING BIOLOGICAL MODELS

These three systems described are MONACO V1.0 (CMS/Elekta, St. Louis, MO), PINNACLE V8.0 h (Philips Medical Systems, Andover, MA), and ECLIPSE V10.0 (Varian Medical Systems, Palo Alto, CA) systems.

B.1. MONACO

The MONACO system is one of the first commercial IMRT treatment planning systems incorporating biologically based optimization features. MONACO offers three biological cost functions titled Poisson statistics cell kill model, Serial complication model, and Parallel complication model to handle dose prescription for targets and OARs exhibiting serial and parallel behavior. Five physical cost functions are also supplied, i.e., quadratic overdose penalty, quadratic underdose penalty, overdose DVH constraint, underdose DVH constraint, and maximum dose constraint. Despite the presence of conventional DV-based cost functions, the system has been specifically designed to utilize biological models, and produces better plans when the biological optimization features are used to their full potential. The Poisson cell kill model has been made a mandatory cost function for targets. If there is a sole PTV, this function must be assigned to the PTV; additional physical cost functions may also be specified. In case of multiple PTVs, the Poisson cell kill model must be used to create optimization criteria for at least one of the PTVs.

The biological cost functions implemented in MONACO are based on a formalism developed at the University of Tübingen.[65,66] For each of the three functions, a 3D dose distribution in a structure is reduced to a single index that reflects a putative biological response of the structure to radiation. This index is referred to as *isoeffect*. For the Poisson cell kill model and Serial complication model, the isoeffect is expressed in units of dose. For the Parallel complication model the isoeffect is a percentage of the organ that is damaged. Dose or percentage levels specified by the user as optimization goals are referred to as *isoconstraints*. Following each iteration, isoeffects are recomputed and compared with isoconstraints to determine whether user-specified criteria have been met.

Isoeffects for targets (i.e., the Poisson cell kill model) are calculated as

$$D_{eff} = -\frac{1}{\alpha'}\ln\left[\frac{1}{\rho' V}\int_V f(D(\vec{x}))dx^3\right], \tag{B1}$$

where α' is the average cell sensitivity, ρ' is the average clonogen density, V is the total volume of the organ (i.e., number of voxels), and $f(D(\vec{x}))$ is a biological response function given by

$$f(D(\vec{x})) = \rho(\vec{x})\exp[-\alpha(\vec{x})D(\vec{x})], \tag{B2}$$

where $\rho(\vec{x})$ is the local density of clonogenic tumor cells, $\alpha(\vec{x})$ is the cell sensitivity in a particular voxel, and $D(\vec{x})$ is the absorbed dose in this voxel. Equations (B1) and (B2) are ready to accommodate information about spatially heterogeneous clonogen density and/or clonogen radiosensitivity that will become available in the future, pending advances in biological imaging techniques. However, at present, spatial variations in either clonogen density or cell sensitivity are not taken into account, i.e., $\rho(\vec{x}) \equiv \rho'$ and $\alpha(\vec{x}) \equiv \alpha'$. Parameter ρ' has been hard-coded to 10^6 clonogens per voxel and presently its value has no impact on isoeffect calculations because $\rho(\vec{x})$ in Eq. (18) and ρ' in Eq. (B1) cancel out. Parameter α' takes on user-specified values in the range $0.1–1.0\,\mathrm{Gy}^{-1}$ with a default value of $0.25\,\mathrm{Gy}^{-1}$. Equation (B2) represents a simplified (only linear component of cell killing is taken into account) expression for the number of surviving clonogens in a voxel based on the standard Poisson statistics-based TCP model. Equation (B1) is an inverted form of Eq. (B2) and conceptually represents the EUD.

Isoeffects for serial-type OARs are specified in terms of an effective dose given by

$$D_{\mathrm{eff}} = \left[\frac{1}{V}\int_V g(D(\vec{x}))dx^3\right]^{1/k}, \tag{B3}$$

where k is the volume effect parameter, V is the total number of voxels, and a response function applied to each voxel takes the form

$$g(D(\vec{x})) = [D(\vec{x})]^k. \tag{B4}$$

Combination of Eqs. (B3) and (B4) is mathematically equivalent to the gEUD formula[14] with $k \equiv a$. D_{eff} approaches maximum dose as k increases.

Although the Serial complication model with $k = 1$ may be used to handle situations when mean organ dose needs to be controlled, i.e., for parallel-type structures, MONACO provides an additional cost function for this purpose. Isoeffects for the Parallel complication model are computed in terms of the mean organ damage

$$v_{\mathrm{eff}} = 100\% \times \frac{1}{V}\int_V h(D(\vec{x}))dx^3, \tag{B5}$$

where V is the total number of voxels and a voxel response function h is calculated as

$$h(D(\vec{x})) = \left[1 + \left(\frac{d_0}{D(\vec{x})}\right)^k\right]^{-1}, \tag{B6}$$

where d_0 is referred to as the reference dose, i.e., a dose that results in 50% complication rate, k is the power law exponent, which determines the steepness of the sigmoid curve described by Eq. (B6). As a rule of thumb, one may choose $k = 0.15\,\mathrm{Gy}^{-1} \times D_{50}$.[67] When expressed as a fraction rather than a percentage, the mean organ damage is mathematically equivalent to the f_{dam} concept of the parallel complication model. Properties of biological cost functions employed in MONACO are summarized in Table I. Limits imposed on parameter or isoconstraint values are shown in parentheses. Because the Poisson cell kill model does not include a mechanism to control hot spots in target volumes, a physical cost function, either the quadratic overdose penalty or maximum dose, must be added to create optimization goals for target volumes. In our practice, it has always been possible to design good quality treatment plans using the three biological functions listed in Table I and the Quadratic overdose penalty cost function.[42]

MONACO supports the concept of constrained optimization. That is, the two biological cost functions used for OARs and all physical cost functions are treated as hard constraints. All optimization criteria specified using these cost functions will be met by the TPS. The Poisson cell kill model is only an objective, meaning that the system finds the optimal cell kill subject to satisfying the hard constraints. As a result, the treatment planner does not have to specify any weights, i.e., effectively all cost functions except the Poisson cell kill model are assigned infinitely large weights. Because target dose is only an objective, achieving this objective may often be limited by one or more constraints on dose in nearby OARs or constraints on hot spots in target volumes. A Sensitivity Analysis tool[68] is provided to help the planner to identify the limiting constraints. Desired target coverage could then be obtained by relaxing (increasing) isoconstraint values for the restrictive cost functions.

In addition to primary biological constraints for OARs (i.e., Serial and Parallel complication models), MONACO allows specification of secondary optimization objectives with these functions. This is referred to as the multicriterial option. This option could be used to attempt to further reduce OAR doses when adequate target coverage had already been achieved or in special cases when additional OAR sparing is more important than adequate target coverage, such as retreatments for recurrent tumors.

B.2. Philips PINNACLE

B.2.1. Plan optimization tools

PINNACLE system offers biological optimization features incorporated into its P^{3}IMRT inverse treatment planning module. The biological objective functions have been developed by RaySearch Laboratories AB (Stockholm, Sweden).[69] As opposed to MONACO, PINNACLE is not a designated biologically based optimization system, but rather uses biological cost functions to enhance the traditional, DV-based optimization approach. In addition to a number of DV cost functions (Min dose, Max dose, Uniform dose, Min DVH, Max DVH, and Uniformity), PINNACLE has at its disposal three biological cost functions denoted Min EUD, Target EUD, and Max EUD. These cost functions are defined as[69]

$$F(\mathrm{EUD}) = \theta(\mathrm{EUD}, \mathrm{EUD}_0)\left(\frac{\mathrm{EUD} - \mathrm{EUD}_0}{\mathrm{EUD}_0}\right)^2, \tag{B7}$$

where EUD_0 is the desired dose level specified by the user. Actually attained dose, EUD, is computed according to the gEUD formalism. Function θ is defined as

$$\theta(\mathrm{EUD}, \mathrm{EUD}_0) = \begin{cases} H(\mathrm{EUD} - \mathrm{EUD}_0) & \text{for Max EUD} \\ 1 & \text{for Target EUD} \\ H(\mathrm{EUD}_0 - \mathrm{EUD}) & \text{for Min EUD} \end{cases}, \tag{B8}$$

where H is the Heaviside step function.

Properties of biological cost functions implemented in PINNACLE for the purpose of plan optimization are summarized in Table II. Each function requires specification of a single volume parameter, a, which has the same interpretation as described by Niemierko.[15] For negative a values, cold spots influence EUD to a greater extent, and for positive a values, EUD is most influenced by hot spots. Generally, negative a values are an appropriate choice for targets, positive a values should be used for serial structures, and $a = 1$ should be used for parallel structures. Biophysically meaningful ranges of the volume parameter for a corresponding cost function are shown in parentheses in Table II. However, in contrast to MONACO, PINNACLE does not impose any limits on values of the volume parameter or EUD.

PINNACLE employs the traditional unconstrained optimization approach. Target and OAR cost functions contribute to the overall cost function in proportion to user-specified weights. Also for any cost function (with the exception of the uniformity, which can only be used as a constraint and the uniform dose, which can only be used as an objective), a treatment planner has an option to use it as either an objective or constraint. The latter effectively sets a very high penalty for violating an optimization goal specified using this cost function.

Philips PINNACLE allows a gradual transition to biologically based inverse planning through combining conventional DV-based and novel EUD-based cost functions. Especially for targets, it is advised to supplement EUD-based objectives with DV-based ones to better control target dose distributions.[69]

B.2.2. Plan evaluation tools

PINNACLE provides two separate tools, titled NTCP/TCP Editor and Biological Response panel, for plan evaluation with the help of biological models. The NTCP/TCP Editor is used to obtain NTCP and TCP estimates. NTCP is calculated according to the LKB model. A database of model parameters originating from Burman *et al.*[70] is available, and a user is given the option to customize parameter values. TCP is calculated using an empirical sigmoid curve corresponding to the cumulative distribution function (CDF) of the normal

distribution. Users are responsible for specifying their own estimates of two model parameters describing dose to control 50% of tumors, D_{50}, and a measure of a slope of the sigmoid curve, m.

Users licensed for Biological Evaluation may take advantage of an enhanced plan evaluation tool that includes a database of endpoint- and tumor-stage-specific parameter values (accompanied by literature references) for calculation of NTCP and TCP, capability to compare alternate treatment plans side-by-side, graphical representation of NTCP/TCP for individual structures, and composite estimates of NTCP, TCP, and probability of complication-free tumor control for the entire plan. Models and parameter estimates implemented in the Biological Response panel are based on the expertise collected at the Karolinska Institute and Stockholm University.[71,72] The Källman *s*-model,[73] also known as the relative seriality model, is used to calculate NTCP. The Poisson model with LQ cell survival is used to describe response of the entire organ to uniform irradiation. TCP is calculated with the Poisson model. The majority of default parameter values provided for NTCP calculations come from a PhD Thesis, which were likely obtained by refitting the relative seriality model to the Emami *et al.*[41] data. Default values of D_{50} and γ provided for TCP calculations are taken from old literature dating back to the 1960 s, with the most recent report being from 1993. Both TCP and NTCP parameter databases are customizable, but the choice of models is fixed. The models used for biological plan evaluation in Philips PINNACLE are summarized in Table III.

B.3. Varian ECLIPSE

B.3.1. Plan optimization tools

The ECLIPSE system provides biological optimization through the use of a "plug-in" to an application by RaySearch Laboratories (Sveavägen 25 111 34 Stockholm Sweden). Selecting biological optimization transfers all patient and plan information to the application for fluence optimization with a separate rapid calculation algorithm. The optimized fluences are then returned to the ECLIPSE dose engine for the final dose calculation.

The optimizer differentiates between biological and physical functions used in the optimization. The biological models used in the optimization include TCP Poisson-LQ, NTCP Poisson-LQ, and NTCP Lyman, and are tabulated in Table IV. The TCP Poisson-LQ and NTCP Poisson-LQ models are identical, respectively, to the TCP and NTCP models implemented in the Biological Response panel in PINNACLE. The NTCP Lyman model is the LKB model calculated based on an LQ-corrected DVH. This model is thus somewhat different from the "Lyman-Kutcher" model in the NTCP/TCP Editor in PINNACLE, which does not take an extra α/β parameter and is calculated based on noncorrected DVHs. The models allow specification of repair time for NTCP models and repair/repopulation times for TCP. The biological functions allow the user to specify the weight (constraint bound percentage) used in calculation of the cost function. Physical functions do not allow assignment of a weight but are regarded as constraints that cannot be compromised. Physical functions include: maximum dose, maximum dose for percentage structure volume, maximum gEUD, and uniformity. For structures defined as targets one may additionally define minimum values for dose, dose for percentage of structure volume, and gEUD. Each structure may also have a conformity constraint which specifies the dose gradient near to the structure.

A library of tissue-specific parameter values for the TCP and NTCP models is provided, enabling selection of standardized values within the biological optimizer application. The library may be edited and modified to include user-specified parameter sets and tissues.

The optimizer screen displays the evolution of items used to monitor and modify parameters used in optimization: optimization functions and parameter values, color wash of coronal and sagittal views of the evolving dose distribution, beam fluence patterns, charts and tables of evolving cost function values for each constraint, DVH, and a graph plotting sensitivity of the biological response for a change in dose per fraction for constant total number of fractions. Dose distributions from plans calculated in ECLIPSE may be specified to be used as a base dose from which the optimization proceeds. During optimization the user iterates through addition and modification of optimization functions to achieve the desired DVH characteristics or until the optimization thresholds are reached.

B.3.2. Plan evaluation tools

After a treatment plan has been generated in ECLIPSE, it may be evaluated using the biological evaluation module. The same biological functions used in the optimizer may be used to calculate NTCP and TCP values for structures. Evaluation of EUD is not supported in this module. Models used for plan evaluation are included in Table V.

The tool may be used to evaluate effect on NTCP and TCP values of changing fractionation schedules (e.g., twice vs once per day), changing number of fractions or scaling total dose. In addition to the conventional dose-volume histogram, two other graphs are available for plan evaluation. The LQ-Scaled DVH utilizes the α/β ratio specified in the NTCP or TCP model to scale the DVH to equivalent values for 2 Gy fractions as a standard benchmark for plan comparison. The Radiobiological Response graph shows the potential effect of altering total plan dose by plotting the value of NTCP or TCP vs a scale factor (0.7–3.0) for the total dose.

Since the RaySearch libraries are used in PINNACLE also, the general comments in Appendix Sec. B 2 on parameters are applicable for the ECLIPSE evaluation tools.

[a]The full Task Group report is available at www.aapm.org.
[b]Author to whom correspondence should be addressed. Electronic mail: ali@mcw.edu

[1]C. C. Ling and X. A. Li, "Over the next decade the success of radiation treatment planning will be judged by the immediate biological response of tumor cells rather than by surrogate measures such as dose maximization and uniformity," Med. Phys. **32**, 2189–2192 (2005).

[2]M. V. Graham, J. A. Purdy, B. Emami, W. Harms, W. Bosch, M. A. Lockett, and C. A. Perez, "Clinical dose-volume histogram analysis for pneumonitis after 3D treatment for non-small lung cancer (NSCLC)," Int. J. Radiat. Oncol., Biol., Phys. **45**, 323–329 (1999).
[3]L. B. Marks, S. M. Bentzen, J. O. Deasy, F. M. Kong, J. D. Bradley, I. S. Vogelius, I. El Naqa, J. L. Hubbs, J. V. Lebesque, R. D. Timmerman, M. K. Martel, and A. Jackson, "Radiation dose-volume effects in the lung," Int. J. Radiat. Oncol., Biol., Phys. **76**(3 Suppl), S70–S76 (2010).
[4]J. O. Deasy, "Multiple local minima in radiotherapy optimization problems with dose-volume constraints," Med. Phys. **24**, 1157–1161 (1997).
[5]Q. Wu and R. Mohan, "Multiple local minima in IMRT optimization based on dose-volume criteria," Med. Phys. **29**, 1514–1527 (2002).
[6]R. D. Stewart and X. A. Li, "BGRT: Biologically guided radiation therapy—The future is fast approaching!," Med. Phys. **34**, 3739–3751 (2007).
[7]A. Brahme, "Optimized radiation therapy based on radiobiological objectives," Semin. Radiat. Oncol. **9**, 35–47 (1999).
[8]C. C. Ling, J. Humm, S. Larson, H. Amols, Z. Fuks, S. Leibel, and J. A. Koutcher, "Towards multidimensional radiotherapy (MD-CRT): Biological imaging and biological conformality," Int. J. Radiat. Oncol., Biol., Phys. **47**, 551–560 (2000).
[9]Y. Yang and L. Xing, "Towards biologically conformal radiation therapy (BCRT): Selective IMRT dose escalation under the guidance of spatial biology distribution," Med. Phys. **32**, 1473–1484 (2005).
[10]S. M. Bentzen, "Theragnostic imaging for radiation oncology: Dose-painting by numbers," Lancet Oncol. **6**, 112–117 (2005).
[11]Y. Kim and W. A. Tomé, "Risk-Adaptive Optimization: Selective Boosting of high-risk tumor subvolumes," Int J. Radiat. Oncol., Biol., Phys. **66**, 1528–42 (2006).
[12]A. Søvik, E. Malinen, H. K. Skogmo, S. M. Bentzen, O. S. Bruland, and D. R. Olsen, "Radiotherapy adapted to spatial and temporal variability in tumor hypoxia," Int. J. Radiat. Oncol., Biol., Phys. **68**, 1496–1504 (2007).
[13]www.AAPM.org
[14]A. Niemierko, "Reporting and analyzing dose distributions: A concept of equivalent uniform dose," Med. Phys. **24**, 103–110 (1997).
[15]A. Niemierko, "A generalized concept of equivalent uniform dose (EUD)," Med. Phys. **26**, 1101 (1999).
[16]B. Choi and J. O. Deasy, "The generalized equivalent uniform dose function as a basis for intensity-modulated treatment planning," Phys. Med. Biol. **47**, 3579–3589 (2002).
[17]H. E. Romeijn, J. F. Dempsey, and J. G. Li, "A unifying framework for multi-criteria fluence map optimization models," Phys. Med. Biol. **49**, 1991–2013 (2004).
[18]M. Alber and R. Reemtsen, "Intensity modulated radiation therapy planning by use of a barrier-penalty multiplier method," *Optimization, Methods and Software (OMS)* (Taylor & Francis, London, 2007), Vol. 22, pp. 391–411.
[19]Q. Wu, D. Djajaputra, Y. Wu, J. Zhou, H. H. Liu, and R. Mohan, "Intensity-modulated radiotherapy optimization with gEUD-guided dose-volume objectives," Phys. Med. Biol. **48**, 279–291 (2003).
[20]D. N. Mihailidis, B. Plants, L. Farinash, M. Harmon, L. Whaley, P. Paja, and P. Tomara, "Superiority of equivalent uniform dose (EUD)-based optimization for breast and chest wall," Med. Dosim. **35**, 67–76 (2010)
[21]P. Källman, B. K. Lind, and A. Brahme, "An algorithm for maximizing the probability of complication-free tumour control in radiation therapy," Phys. Med. Biol. **37**, 871–890 (1992).
[22]Q. Wu, R. Mohan, A. Niemierko, and R. Schmidt-Ullrich, "Optimization of intensity-modulated radiotherapy plans based on the equivalent uniform dose," Int. J. Radiat. Oncol., Biol., Phys. **52**, 224–235 (2002).
[23]W. R. De Gersem, S. Derycke, C. O. Colle, C. De Wagter, and W. J. De Neve, "Inhomogeneous target-dose distributions: A dimension more for optimization?," Int. J. Radiat. Oncol., Biol., Phys. **44**, 461–468 (1999).
[24]QUANTEC, Int. J. Radiat. Oncol. Biol. Phys. **76**(3 Suppl) (2010).
[25]S. H. Benedict, K. M. Yenice, D. Followill, J. M. Galvin, W. Hinson, B. Kavanagh, P. Keall, M. Lovelock, S. Meeks, L. Papiez, T. Purdie, R. Sadagopan, M. C. Schell, B. Salter, D. J. Schlesinger, A. S. Shiu, T. Solberg, D. Y. Song, V. Stieber, R. Timmerman, W. A. Tomé, D. Verellen, L. Wang, and F. F. Yin, "Stereotactic body radiation therapy: The report of AAPM Task Group 101," Med Phys. **37**, 4078–4101 (2010).
[26]I. Kawrakow, "The effect of Monte Carlo statistical uncertainties on the evaluation of dose distributions in radiation treatment planning," Phys. Med. Biol. **49**, 1549–1556 (2004).
[27]R. K. Ten Haken, T. S. Lawrence, and L. A. Dawson, "Prediction of radiation-induced liver disease by Lyman normal-tissue complication probability model in three-dimensional conformal radiation therapy for primary liver carcinoma: In regards to Xu *et al*. (Int J Radiat Oncol Biol Phys 2006;65:189-195)," Int. J. Radiat. Oncol., Biol., Phys. **66**, 1272 (2006).
[28]E. S. Koh, A. Sun, T. H. Tran, R. Tsang, M. Pintilie, D. C. Hodgson, W. Wells, R. Heaton, and M. K. Gospodarowicz, "Clinical dose-volume histogram analysis in predicting radiation pneumonitis in Hodgkin's lymphoma," Int. J. Radiat. Oncol., Biol., Phys. **66**, 223–228 (2006).
[29]J. T. Lyman, "Complication probability as assessed from dose-volume histograms," Radiat. Res. Suppl. **8**, S13–S19 (1985).
[30]G. J. Kutcher and C. Burman, "Calculation of complication probability factors for non-uniform normal tissue irradiation: The effective volume method," Int. J. Radiat. Oncol., Biol., Phys. **16**, 1623–1630 (1989).
[31]L. A. Dawson, D. Normolle, J. M. Balter, C. J. McGinn, T. S. Lawrence, and R. K. Ten Haken, "Analysis of radiation-induced liver disease using the Lyman NTCP model," Int. J. Radiat. Oncol., Biol., Phys. **53**, 810–821 (2002). Erratum in: Int. J. Radiat. Oncol., Biol., Phys. **53**, 1422 (2002).
[32]Y. Seppenwoolde, J. V. Lebesque, K. de Jaeger, J. S. Belderbos, L. J. Boersma, C. Schilstra, G. T. Henning, J. A. Hayman, M. K. Martel, and R. K. Ten Haken, "Comparing different NTCP models that predict the incidence of radiation pneumonitis. Normal tissue complication probability," Int. J. Radiat. Oncol., Biol., Phys. **55**, 724–735 (2003).
[33]Z. Y. Xu, S. X. Liang, J. Zhu, X. D. Zhu, J. D. Zhao, H. J. Lu, Y. L. Yang, L. Chen, A. Y. Wang, X. L. Fu, and G. L. Jiang, "Prediction of radiation-induced liver disease by Lyman normal-tissue complication probability model in three-dimensional conformal radiation therapy for primary liver carcinoma," Int. J. Radiat. Oncol., Biol., Phys. **65**, 189–195 (2006).
[34]M. Zaider and H. I. Amols, "A little to a lot or a lot to a little: Is NTCP always minimized in multiport therapy?," Int. J. Radiat. Oncol., Biol., Phys. **41**, 945–950 (1998).
[35]V. Moiseenko, J. Battista, and J. Van Dyk, "Normal tissue complication probabilities: Dependence on choice of biological model and dose-volume histogram reduction scheme," Int. J. Radiat. Oncol., Biol., Phys. **46**, 983–993 (2000).
[36]L. P. Muren, N. Jebsen, A. Gustafsson, and O. Dahl, "Can dose-response models predict reliable normal tissue complication probabilities in radical radiotherapy of urinary bladder cancer? The impact of alternative radiation tolerance models and parameters," Int. J. Radiat. Oncol., Biol., Phys. **50**, 627–637 (2001).
[37]B. Sanchez-Nieto and A. E. Nahum, "The delta-TCP concept: A clinically useful measure of tumor control probability," Int. J. Radiat. Oncol., Biol., Phys. **44**, 369–380 (1999).
[38]W. A. Tomé and J. F. Fowler, "Selective boosting of tumor subvolumes," Int. J. Radiat. Oncol., Biol., Phys. **48**, 593–599 (2000).
[39]W. A. Tomé and J. F. Fowler, "On cold spots in tumor subvolumes," Med. Phys. **29**, 1590–1598 (2002).
[40]P. Okunieff, D. Morgan, A. Niemierko, and H. D. Suit, "Radiation dose-response of human tumors," Int. J. Radiat. Oncol., Biol., Phys. **32**, 1227–1237 (1995).
[41]B. Emami, J. Lyman, A. Brown, L. Coia, M. Goitein, J. E. Munzenrider, B. Shank, L. J. Solin, and M. Wesson, "Tolerance of normal tissue to therapeutic irradiation," Int. J. Radiat. Oncol., Biol., Phys. **21**, 109–122 (1991).
[42]V. A. Semenenko, B. Reitz, E. Day, X. S. Qi, M. Miften, and X. A. Li, "Evaluation of a commercial biologically based IMRT treatment planning system," Med. Phys. **35**, 5851–5860 (2008).
[43]X. S. Qi, V. A. Semenenko, X. A. Li, "Improved critical structure sparing with biologically-based IMRT optimization," Med. Phys. **36**, 1790–1799 (2009).
[44]G. J. Kutcher, L. Coia, M. Gillin, W. F. Hanson, S. Leibel, R. J. Morton, J. R. Palta, J. A. Purdy, L. E. Reinstein, G. K. Svensson, M. Weller, and L. Wingfield, "Comprehensive QA for radiation oncology: Report of AAPM Radiation Therapy Committee Task Group 40," Med. Phys. **21**, 581–618 (1994).
[45]B. Fraass, K. Doppke, M. Hunt, G. Kutcher, G. Starkschall, R. Stern, and J. Van Dyke, "American Association of Physicists in Medicine Radiation Therapy Committee Task Group 53: Quality assurance for clinical radiotherapy treatment planning," Med. Phys. **25**, 1773–1829 (1998).
[46]G. A. Ezzell, J. W. Burmeister, N. Dogan, T. J. Losasso, J. G. Mechalakos, D. Mihailidis, A. Molineu, J. R. Palta, C. R. Ramsey, B. J. Salter, J. Shi, P. Xia, N. J. Yue, and Y. Xiao, "IMRT commissioning: Multiple institution planning and dosimetry comparisons, a report from AAPM Task Group 119," Med. Phys. **36**, 5359–5373 (2009).

[47]E. E. Klein, J. Hanley, J. Bayouth, F. F. Yin, W. Simon, S. Dresser, C. Serago, F. Aguirre, L. Ma, B. Arjomandy, C. Liu, C. Sandin, and T. Holmes, "Task Group 142 report: Quality assurance of medical accelerators," Med Phys. **36**, 4197–212 (2009).
[48]J. O. Deasy, A. I. Blanco, and V. H. Clark, "CERR: A computational environment for radiotherapy research," Med. Phys. **30**, 979–985 (2003).
[49]http://radium.wustl.edu/CERR/about.php
[50]B. Sanchez-Nieto and A. E. Nahum, "BIOPLAN: Software for the biological evaluation of radiotherapy treatment plans," Med. Dosim. **25**, 71–76 (2000).
[51]J. Uzan and A. E. Nahum, "BioSuite, new software for radiobiological customisation of dose and fraction size in external-beam radiothearapy," 10th Biennial ESTRO meeting (2009).
[52]http://www.aapm.org/org/committees/TG166/TG166prostate.zip.
[53]http://www.aapm.org/org/committees/TG166/TG166headneck.zip.
[54]http://www.aapm.org/org/committees/TG166/TG166brain.zip.
[55]http://www.aapm.org/org/committees/TG166/EUD_Monaco_Pinnacle_Eclipse.xls.
[56]ICRU, "Prescribing, recording and reporting photon beam therapy," ICRU Report No. 50 (International Commission on Radiation Units and Measurements, Washington, DC, 1993).
[57]J. H. Killoran, H. M. Kooy, D. J. Gladstone, F. J. Welte, and C. J. Beard, "A numerical simulation of organ motion and daily setup uncertainties: Implications for radiation therapy," Int. J. Radiat. Oncol., Biol., Phys. **37**, 213–221 (1997).
[58]D. J. Brenner and R. K. Sachs, "A more robust biologically based ranking criterion for treatment plans," Int. J. Radiat. Oncol., Biol., Phys. **43**, 697–698 (1999).
[59]S. Das, "A role for biological optimization within the current treatment planning paradigm," Med Phys. **36**, 4672–4682 (2010).
[60]Y. Lu, D. R. Spelbring, and G. T. Chen, "Functional dose-volume histograms for functionally heterogeneous normal organs," Phys. Med. Biol. **42**, 345–56 (1997).
[61]L. B. Marks, G. W. Sherouse, M. T. Munley, G. C. Bentel, and D. P. Spencer, "Incorporation of functional status into dose-volume analysis," Med. Phys. **26**, 196–9 (1999).
[62]http://ctep.cancer.gov
[63]J. D. Cox, J. Stetz, and T. F. Pajak, "Toxicity criteria of the radiation therapy oncology group (RTOG) and the European organization for research and treatment of cancer (EORTC)," Int. J. Radiat. Oncol., Biol., Phys. **31**, 1341–1346 (1995).
[64]H. R. Withers, J. M. Taylor, and B. Maciejewski, "Treatment volume and tissue tolerance," Int. J. Radiat. Oncol., Biol., Phys. **14**, 751–759 (1988).
[65]M. Alber and F. Nüsslin, "An objective function for radiation treatment optimization based on local biological measures," Phys. Med. Biol. **44**, 479–493 (1999).
[66]M. Alber, "A concept for the optimization of radiotherapy," Ph.D. dissertation, University of Tübingen, Tübingen, Germany, 2000.
[67]M. Alber and C. Belka, "A normal tissue dose response model of dynamic repair processes," Phys. Med. Biol. **51**, 153–172 (2006).
[68]M. Alber, M. Birkner, and F. Nüsslin, "Tools for the analysis of dose optimization: II. Sensitivity analysis," Phys. Med. Biol. **47**, N265–N270 (2002).
[69]B. Hårdemark, A. Liander, H. Rehbinder, J. Löf, and D. Robinson, "P^{3}IMRT Biological optimization and EUD," Pinnacle3 White Paper, available at www.medical.philips.com/main/products/ros/assets/docs/white_papers/EUD_Final.pdf.
[70]C. Burman, G. J. Kutcher, B. Emami, and M. Goitein, "Fitting of normal tissue tolerance data to an analytic function," Int. J. Radiat. Oncol., Biol., Phys. **21**, 123–135 (1991).
[71]B. K. Lind, P. Mavroidis, S. Hyödynmaa, and C. Kappas, "Optimization of the dose level for a given treatment plan to maximize the complication-free tumor cure," Acta Oncol. **38**, 787–798 (1999).
[72]G. Kåver, B. K. Lind, J. Löf, A. Liander, and A. Brahme, "Stochastic optimization of intensity modulated radiotherapy to account for uncertainties in patient sensitivity," Phys. Med. Biol. **44**, 2955–2969 (1999).
[73]P. Källman, A. Ågren, and A. Brahme, "Tumour and normal tissue responses to fractionated non-uniform dose delivery," Int. J. Radiat. Biol. **62**, 249–262 (1992).