Radiotherapy and Oncology 153 (2020) 26–33

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Review Article

What is plan quality in radiotherapy? The importance of evaluating dose metrics, complexity, and robustness of treatment plans

Victor Hernandez [a,1,*], Christian Rønn Hansen [b,c,d,1], Lamberto Widesott [e,1], Anna Bäck [f,g], Richard Canters [h], Marco Fusella [i], Julia Götstedt [j], Diego Jurado-Bruggeman [k], Nobutaka Mukumoto [l], Laura Patricia Kaplan [m], Irena Koniarová [n], Tomasz Piotrowski [o,p], Lorenzo Placidi [q], Ana Vaniqui [r], Nuria Jornet [s]

[a] *Department of Medical Physics, Hospital Sant Joan de Reus, IISPV, Spain;* [b] *Laboratory of Radiation Physics, Odense University Hospital;* [c] *Institute of Clinical Research, University of Southern Denmark;* [d] *Danish Centre for Particle Therapy, Aarhus University Hospital, Denmark;* [e] *Centro di protonterapia, APSS, Trento, Italy;* [f] *Department of Therapeutic Radiation Physics, Medical Physics and Biomedical Engineering, Sahlgrenska University Hospital, Gothenburg;* [g] *Department of Radiation Physics, Institute of Clinical Sciences, Sahlgrenska Academy at the University of Gothenburg, Sweden;* [h] *Department of Radiation Oncology (Maastro), GROW School for Oncology, Maastricht University Medical Centre+, The Netherlands;* [i] *Medical Physics Department, Veneto Institute of Oncology IOV – IRCCS, Padua, Italy;* [j] *Department of Radiation Physics, University of Gothenburg, Göteborg, Sweden;* [k] *Medical Physics and Radiation Protection Department, Institut Català d'Oncologia, Girona, Spain;* [l] *Department of Radiation Oncology and Image-applied Therapy, Graduate, School of Medicine, Kyoto University, Japan;* [m] *Department of Oncology, Aarhus University Hospital, Denmark;* [n] *National Radiation Protection Institute, Prague, Czech Republic;* [o] *Department of Electroradiology, Poznań University of Medical Sciences;* [p] *Department of Medical Physics, Greater Poland Cancer Centre, Poznań, Poland;* [q] *Fondazione Policlinico Universitario "A. Gemelli" IRCCS, UOC Radioterapia Oncologica, Dipartimento di Diagnostica per Immagini, Radioterapia Oncologica ed Ematologia, Roma, Italy;* [r] *Department of Radiation Oncology (Maastro), GROW School for Oncology, Maastricht University Medical Centre+, The Netherlands;* [s] *Servei de Radiofísica i Radioprotecció, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain*

ARTICLE INFO

Article history:
Received 2 July 2020
Received in revised form 17 September 2020
Accepted 17 September 2020
Available online 25 September 2020

Keywords:
Plan quality
Dose metrics
Complexity
Robustness
Robust optimisation
Robust evaluation
Complexity metrics

ABSTRACT

Plan evaluation is a key step in the radiotherapy treatment workflow. Central to this step is the assessment of treatment plan quality. Hence, it is important to agree on what we mean by plan quality and to be fully aware of which parameters it depends on. We understand plan quality in radiotherapy as the clinical suitability of the delivered dose distribution that can be realistically expected from a treatment plan. Plan quality is commonly assessed by evaluating the dose distribution calculated by the treatment planning system (TPS). Evaluating the 3D dose distribution is not easy, however; it is hard to fully evaluate its spatial characteristics and we still lack the knowledge for personalising the prediction of the clinical outcome based on individual patient characteristics. This advocates for standardisation and systematic collection of clinical data and outcomes after radiotherapy. Additionally, the calculated dose distribution is not exactly the dose delivered to the patient due to uncertainties in the dose calculation and the treatment delivery, including variations in the patient set-up and anatomy. Consequently, plan quality also depends on the robustness and complexity of the treatment plan. We believe that future work and consensus on the best metrics for quality indices are required. Better tools are needed in TPSs for the evaluation of dose distributions, for the robust evaluation and optimisation of treatment plans, and for controlling and reporting plan complexity. Implementation of such tools and a better understanding of these concepts will facilitate the handling of these characteristics in clinical practice and be helpful to increase the overall quality of treatment plans in radiotherapy.

Evaluation of treatment plans is a key step in the radiotherapy process that determines the characteristics of the plan selected for treatment and, consequently, how patients undergoing radiotherapy are treated. The goal of plan evaluation is to assess a plan's quality using a number of qualitative and/or quantitative measures. The concept of plan quality can, however, encompass many different characteristics of the treatment plan and there is no global consensus on how exactly to define, measure, and report plan quality. For this reason, the 3rd Physics ESTRO Workshop held in Budapest in October 2019 included a track on 'Plan quality assessment'. The purpose of this work is to summarise the contents discussed during this Workshop and to provide the participants' vision on this topic.

This work comprises three sections focused on different aspects that affect the quality of treatment plans: dose metrics, plan robustness and plan complexity. Finally, we present the working

* Corresponding author at: Department of Medical Physics, Hospital Sant Joan de Reus, IISPV, Spain.
E-mail address: vhernandezmasgrau@gmail.com (V. Hernandez).
[1] The first three authors contributed equally to this work.

https://doi.org/10.1016/j.radonc.2020.09.038
0167-8140/

group's overall conclusions on the subject of treatment plan quality assessment.

Dose metrics

The evaluation of the calculated dose distribution is a fundamental aspect of plan evaluation. It is critical to ensure high plan quality during the treatment planning process and upon treatment approval. However, this evaluation is not easy since there are multiple aspects that commonly play a role in the assessment of plan quality. These include protocols, local requirements, historical customs and personal preferences, and they can affect any final decision.

Prescriptions and treatment planning protocols should incorporate a prioritisation of the dose objectives and constraints on targets and organs at risk (OARs). This prioritisation should be based on expert evaluation of the scientific literature, as done by the DAHANCA group [1,2]. Such a prioritisation would help to standardise the decision-making process and reduce subjectivity when clinical compromises are needed. It is important to be critical towards historical customs and traditions, however. These must always be supported by scientific evidence or at least be regularly re-evaluated based on new experience. Inter-planner variability may be another factor that affects the final treatment plan and the resulting dose distribution. This can be reduced by using automatic planning strategies [3,4]. In any case, personal preferences based on subjective beliefs, such as attempting to make the dose distribution 'look' nice (for instance prioritising conformity over dose to OARs and normal tissues), should be avoided.

To that purpose, the use of class solutions in treatment planning is highly recommended and local planning protocols should be adapted to each institution's specific requirements, be evaluated periodically and incorporate dynamic development [5]. Additionally, it is highly recommended to standardise the collection of patient data such as co-morbidities, systemic treatments, toxicities, local control and survival in a database, regardless of their inclusion in clinical trials. This will help build clinical knowledge based on earlier experiences. The consistency of such data is crucial and any potential sources of bias, such as how the dose distribution is normalised [6–8] and what calculation algorithm and dose quantity is used [9] should be carefully considered and reported.

The evaluation of the calculated dose distribution is often based on dose volume histograms (DVHs), which collapse the 3D dose information in 2D metrics (dose and volume), losing the information on its spatial distribution. Due to this limitation, a slice-by-slice inspection of the dose distribution is recommended to identify potential aspects for further improvement [7]. The target coverage is commonly evaluated by comparing DVH metrics as well as hot and cold regions to protocolised goal values [6–8]. Regarding OARs, studies on clinical tolerances based on DVH metrics have been collected by Emami et. al. [10] and updated in the QUANTEC series of publications by Bentzen et al. [11]. These tolerances rely on a single DVH endpoint and were based on an arbitrary level of acceptable toxicity. This can be problematic since dose metrics are often highly correlated and a single dose metric might not encapsulate the full causal dose response relationship. Existing mathematical methods such as principal component analysis, which can disentangle the collinearity effects of dose metrics [12], can help overcome these problems. Another issue is the fact that, when QUANTEC data were collected, the use of 3D TPSs was not fully implemented and data analysis capacities of computers were limited.

Instead of evaluating the target coverage and the risk of toxicity through a single dosimetric endpoint, the tumour control probability (TCP) and normal tissue complication probability (NTCP) can be calculated from biological models [13,14], which are usually based on the DVH information. Most of these tolerances and NTCP models were generated using data from one or few centres and should, therefore, be used with caution. If possible, a local validation of the NTCP model should be conducted but, as a minimum, an understanding of the underlying model cohort and the limitations of the model is required [15]. It is also important, when possible, to minimise OAR doses below model tolerances to facilitate potential reirradiations in case of tumour relapse [16].

Derived DVH dose metrics can also be useful to objectively quantify the quality of a dose distribution regarding target irradiation [17] and have been adopted in radiotherapy guidelines. These metrics include homogeneity, conformity and gradient indices and some of them indirectly ensure low doses to the normal tissues. They are used in some protocols that require specific dose distributions (e.g., stereotactic treatments and brachytherapy) and different definitions are used depending on the clinical purpose and the considered protocol, each one with its own strengths and drawbacks. Homogeneity indices are usually based on ratios of two or three points in the target DVH, such as different definitions of maximum, minimum and prescribed dose [7,8]. More complex indices can take into account all the points in the DVH curve and quantify the dose dispersion around the average dose [18,19]. Conformity indices (CI) involve ratios of volumes treated at a given isodose to the target volume [20]. The RTOG CI [21] is criticised for not taking into account the location and shape of the prescription isodose with respect to the target volume. To overcome this limitation, other indices have been proposed, such as the Paddick CI, which combines conformity and target coverage [22]. A steep dose fall-off is important in order to decrease toxicity in tissues surrounding the target, especially in treatments involving high doses per fraction. Dose gradient indices quantify this property [23].

It can be practical to collapse all this information into an overall plan score or single plan quality index (PQI), as recently pointed out by Giglioli et al. [24]. PQIs are obtained by weighing DVH values and DVH-derived metrics based on patient group specific protocols and they facilitate comparisons while minimising subjectivity. Large efforts have been made by several authors to define a composite score or figure of merit for quantifying overall plan quality [25–29] and there is a commercial tool available [30,31]. These quality indices are practical as decision-support systems when comparing plans for the same patient [32] and have also been used for evaluating plan quality across different patients and platforms [24] and in treatment planning competitions [33–35]. However, which parameters to consider and their weights is not a straightforward decision and involves a certain degree of subjectivity. Hence, careful evaluation is required in comparisons across different patients and technologies. PQIs are the basis of the cost functions used in plan optimisers and in automatic planning strategies; therefore, these aspects should be carefully considered when plans are evaluated and when automatic planning models are trained [36].

Texture analysis was recently proposed to characterise the spatial dose distribution. This is known as dosiomics [37–42] and is a powerful technique for characterising spatial and statistical distributions of pixel/voxel intensities in an image through the identification of patterns and voxel correlations. Dosiomics is a promising method for parameterising regions of interest and for producing intensity, textural and shape-based dose features that might be able to describe the dose distribution better than DVH-based metrics, as well as to potentially improve the predictive performance of TCP and NTCP models. Moreover, dosiomics can provide novel quality metrics from the extracted features, facilitating a more accurate evaluation and further standardisation of treatment plan quality. Feature extraction methodologies are comparable with

those used by radiomics but they involve 3D dose matrices instead of diagnostic images (CT, MRI, etc.). Among the radiomic features defined in the Image Biomark Standardization Initiative [43], the most promising features for dosiomic studies belong to the first order and textural families (local intensity, intensity-based statistics, grey level co-occurrence matrix, grey level size zone matrix, etc..). A standardisation and stability analysis of the selected dosiomic features is mandatory [44], as well as a feature redundancy investigation to select the dosiomic features that are independent of DVH-based metrics and which can thus provide additional value. Not only anatomy-based ROIs (e.g. delineated target volumes and OAR) should be considered for the dosiomic feature extraction; instead, dedicated dosiomic regions of interest from which to extract the features (isodose volumes, intersections of anatomy-based ROIs, regions of low bath dose, etc.) may offer a better evaluation of the 3D dose distribution. There is also a need for dosiomic features to have a clear clinical interpretation because 'black box' features would be difficult or impossible to prioritise against other metrics.

Ideally, the quality of a dose distribution should be linked to evidence-proven clinical outcomes and should be personalised to account for patient-specific data such as tumour characteristics, complementary treatments, and risk factors. However, this is a truly complex problem involving a large amount of data and it is not well known yet how to carry out this personalisation. This provides a compelling challenge not only for TPS manufacturers but for the whole scientific community. In general, we believe that more effort is warranted by:

- TPS manufacturers, to implement additional tools for evaluating dose distributions, such as radiobiological models, as well as to facilitate the future implementation of new spatially correlated tools, such as dosiomics, and by
- the scientific community, to collect structured data on dose indices, patient and tumour characteristics and follow-up data on tumour control, toxicities and survival. Future analysis of the data with appropriate tools such as Big Data Analytics and artificial intelligence (AI) could help personalise radiotherapy treatments and further link the quality of a dose distribution to the patient's clinical outcome.

Robustness

The traditional method to achieve target coverage and OARs sparing has been the definition of adequate margins around the clinical target volume (CTV) and OARs to obtain the planning target volume (PTV) and organ at risk planning volume (PRV). Several formulas have been proposed for the definition of these margins for the PTV [45,46] and PRVs [47,48].

However, there are several limitations that affect the PTV definition: it relies on the so-called static dose cloud approximation and does not guarantee optimal management for PTVs extending into air. Moreover, whether or not the CTV receives the prescribed dose depends also on the specific dose distribution rather than only on geometric margins. Dose distributions are neither perfectly conformal to the PTV nor equally conformal on all sides of the CTV and non-conformity results in an inherent dosimetric margin [49]. In those regions where the prescription isodose line extends beyond the CTV less or no margin needs to be added to account for treatment uncertainties. In addition to conformity, the required margin also depends on the steepness of the dose fall-off near the target because a naturally gradual fall-off may require a smaller margin than a sharp fall-off. As Stroom et. al. underlined, the PRV concept has even more limitations and it seems necessary to develop alternative ways to include geometric uncertainties of OARs in treatment planning [48]. All these concerns about the use of PTV and PRV are even more important in proton therapy.

Robust optimisation follows a different approach, addressing uncertainties explicitly, without the need for margins. Thus, instead of optimising a single scenario (i.e., patient in a fixed nominal position and considering PTV and PRV margins) dose distributions are optimised for n scenarios. Each scenario represents a possible treatment course (i.e. all the treatment fractions) and should include the specification of all the errors that are needed to calculate a final dose distribution. In the context of random setup errors in a fractionated treatment, this would also include the setup errors for all individual fractions. The plan can then be optimised considering all of the n scenarios at once (often in the worst-case scenario, minimax approach) or a combination of different scenarios, each one assigned a certain probability (probabilistic approach).

Regardless of whether the plan has been optimised in the 'classical' way (i.e., PTV-based) or with robust optimisation, it is always possible to make an a posteriori assessment of the plan's robustness. With robustness evaluation one can evaluate how the dose distribution changes compared to the nominal dose by recalculating the plan obtained in different error scenarios. This offers the possibility of quantifying the uncertainties in DVHs and other dose metrics due to e.g. variations in patient set-up and anatomy [50,51].

It has been shown that robust optimisation can potentially solve the PTV/PRV limitations and improve CTV coverage and OAR sparing [52–54] both for photon and proton treatments. The problem of PTV expansion into air is also solved since variations concerning the CTV are explicitly accounted for in the calculation. Robust optimisation and evaluation are generally posed as n-dimensional problems; for instance, n = 1 can take into account the set-up error, n = 2 the range uncertainty for protons, n = 3 the breathing phases of a 4DCT and n = 4 possible anatomical changes (such as cavity filling and tumour shrinkage) to reduce the need for re-planning. In general, robust optimisation reduces doses to normal tissues that would be unnecessarily irradiated with the PTV-margin concept and leads to delivered dose distributions with a higher probability of acceptable target coverage and OAR sparing [52–54].

Only a few years since it was first implemented in a commercial TPS in 2014, robust optimisation has become a standard tool in proton therapy and is also gaining interest in photon therapy. A treatment plan that is both 'good' (i.e., with a clinically acceptable nominal dose distribution), and 'robust' will yield a dose distribution that is suitable for all or the majority of the error scenarios considered by the optimiser. There are different paradigms to translate this notion into mathematical terms. Broadly, these approaches can be categorised as (i) the *probabilistic (stochastic) approach*, which optimises the expected delivered dose distribution, and (ii) the *minimax approach*, which optimises the dose distribution for the worst error considered [55]. Probabilistic optimisation requires associating an occurrence probability to each considered uncertainty scenario. Gaussian probability densities are commonly assumed as a first approximation, which yields treatment plans that are also robust against other probability density functions [56]. Worst-case methods can focus on (i) composite worst-case scenario (*composite* worst-case), (ii) worst-case scenarios for each objective considered independently (*objectivewise* worst-case) and (iii) worst-case scenarios for each voxel considered independently (*voxelwise* worst case) and no particular method has yet proven to be superior to the others in all circumstances [57].

We must bear in mind that the terms 'robust optimisation' and 'robust analysis' encompass several different methodologies and metrics, and there is no consensus on which ones should be imple-

mented or exactly how they should be used [55,58–61]. Robustness evaluation tools are in their infancy and many metrics can be used that provide *global* (DVHs, variations in dosimetric indices, etc.) or *local* (2D dose distribution) information.

A message that clearly emerges from the literature is that there can be significant differences, both in terms of target coverage and OAR sparing, between the nominal dose distribution approved by the clinician and the dose distribution in the different scenarios evaluated in the robustness analysis. This is true both for photon and proton plans, which is why robustness analysis should not be restricted to the proton therapy world alone.

In order to make robustness analysis clinically useful, it is extremely important to define a benchmark for the robust analysis tools and agree on what constitutes a safe degree of variability. Is it acceptable, for instance, if in the worst-case scenario the D1% to the spinal cord increases by 20 Gy compared to the nominal dose, even though it remains within tolerance? Calibration of robust evaluation methods against existing PTV-based methods could be a solution for this [58]. Another important consideration must be made when robustness is applied to OAR dose constraints in addition to target coverage. The OAR constraints reported in the QUANTEC papers [11] are all based on nominal DVHs and no robust analysis was considered in the setting of those constraints. Therefore, the request that these OAR constraints are fulfilled in many, if not all, the considered scenarios may be too restrictive and compromise target coverage.

Robustness methods address variations in the patient's anatomy throughout the course of treatment, but the uncertainty in volume delineation is typically not considered. Automatic contouring algorithms that improve inter-observer consistency are currently available for clinical use, especially for organs at risk [62]. Volume delineation still constitutes an important potential source of uncertainty in radiotherapy, however. Interestingly, robustness methods could also be applied to delineation uncertainties, but convincing applications of robust optimisation in this context are limited [55].

Given the potential of these new tools and their current availability in treatment planning systems, it is important to reach consensus on which methods are the most appropriate for both robust optimisation and evaluation. Retrospectively analysing plan robustness for patients already treated with protons (with follow up data on toxicity and tumour control) could be an interesting solution. These data could be useful for establishing appropriate references for robustness parameters, as proposed by Malyapa et. al. [61]. We believe the same should be done for typical IMRT and VMAT plans in different treatment sites in order to enable a transition to margin-less robust optimisation and evaluation in the photon world.

In this context we consider it essential to:

- Implement more and faster tools in TPSs to facilitate both robust optimisation and robust evaluation of treatment plans. These tools should preferably include not only the worst-case scenario, but also the probabilistic approach and be available for both proton and photon planning.
- Further perform robust optimisation and evaluation of clinical plans, also for photon beam therapy.
- Establish a common benchmark for plan robustness evaluation that is used for both particle and photon beam therapy and allows a fair comparison of the robustness of proton and photon plans.
- Reach consensus in the scientific community on which methods and metrics are the most appropriate for robust optimisation and analysis, e.g. to quantify the degree of robustness of a given plan.
- Generate databases of clinical data and robustness parameters and share these data among different centres. This will enable tumour control and OAR toxicity to be correlated with the results of robustness evaluations.

Plan complexity

Modern modulated radiotherapy techniques involve modulation of many machine parameters, placing high demands on treatment machines and TPSs. This increase in modulation of machine parameters is commonly referred to as an increase in treatment plan complexity. More complex plans have larger uncertainties in dose calculation and treatment delivery [63–65] compared to non-modulated plans.

It is challenging to agree on a unique definition of complexity, even among scientists [66]. We understand the complexity of a radiotherapy treatment plan as an estimation of the degree of dose uncertainty as a result of its calculation and delivery, which depends on all the machine parameters that make up the treatment plan.

Several complexity metrics have been proposed in the literature for MLC-based treatments. Initially, they were based on fluence maps, but these maps are not always available. Furthermore, the same fluence maps can be produced through different variations of machine parameters depending on the optimisation and sequencing algorithms used, which cannot be taken into account by fluence-based metrics [67]. Therefore, fluence-based metrics have been gradually replaced by metrics that directly depend on the treatment plan parameters. Different complexity metrics focus on different aspects of plan complexity [68], such as aperture modulation [67,69], the size and irregularity of beam apertures [67,70–73] and modulation of specific machine parameters, e.g. the distance travelled by the leaves [74] and the variations in dose rate and gantry speed [75,76]. A detailed explanation of complexity metrics can be found in two recent reviews [77,78].

A certain degree of complexity in treatment plans is necessary because it is often required to achieve an acceptable dose distribution. However, many investigators have reported that a high degree of plan complexity may affect the accuracy of dose calculation and treatment delivery [69,71,72,74,75,79,80]. This is because more complex plans typically involve smaller and more irregular beam apertures, larger tongue-and-groove effects and larger modulation of machine parameters. Such complexities affect the uncertainties in dose calculations due to limitations in the calculation algorithm or in the beam model, e.g. in the MLC configuration [81]. They also influence the sensitivity of the delivered dose to small deviations in machine parameters during treatment delivery (even within their tolerances) and to variations in patient geometry, e.g. respiratory motion [82]. Plan complexity can be interpreted as an assessment of the robustness of treatment plans to all these uncertainties. Additionally, highly complex plans require longer beam-on times, which can increase the risk of intrafraction movements in some treatment sites [83]. A high degree of plan complexity may therefore compromise the overall accuracy and quality of the treatment. Consequently, high degrees of plan complexity should be avoided to maximise accuracy and robustness of radiation treatments. AAPM pointed out the need to quantify plan modulation [84] and to adapt tolerance limits to the degree of plan complexity [65].

It is known that 'dosimetric' plan quality is often not correlated to plan complexity [85–87] and that similar dose distributions are achievable with more or less complex treatment plans because inverse optimisation can introduce unnecessary complexity [70,85,88]. For these reasons, many investigators have recommended incorporating complexity metrics into the cost function used by optimisation algorithms [69,70,79,86,89]. This is particu-

larly relevant in automated planning, where there is an added risk of inadvertently increasing plan complexity [90–92]. There is no consensus on which complexity metric should be used. Many of them are correlated and multiple metrics can be used to account for the different uncertainties and sources of plan complexity [68]. A very comprehensive table was recently provided by Antoine et al. [78], summarising the main proposed complexity metrics that could be used. Several different metrics may have a similar impact in reducing plan complexity and further studies are needed. In our opinion, the more relevant aspects in plan complexity are the degree of aperture modulation and the irregularity of the beam apertures. Hence, a combination of metrics addressing these two aspects could be used to control plan complexity during the optimisation process. Additionally, it would be desirable that TPSs clearly describe the metrics that they use and report the obtained values to facilitate the management of plan complexity in clinical practice.

Information on the complexity of treatment plans can be helpful to evaluate trade-offs between dosimetric performance and plan complexity. Additionally, a plan complexity metric can serve as a plan verification tool aimed at reducing the QA workload, e.g. because plans with a low degree of complexity may not require as much verification as highly complex plans [77,78]. This can be important for the QA of online adaptive radiotherapy, where measurement-based verifications are not feasible and complexity metrics can facilitate a fast verification of the adapted plans [93].

For the reasons discussed above, plan complexity is a relevant aspect when evaluating treatment plan quality. The quality of a treatment plan depends not only on the calculated dose distribution (dosimetric performance), but also on the accuracy of the dose calculation and beam delivery. Therefore, plan quality must incorporate an assessment of plan complexity. Plan complexity has been discussed in the literature mainly for photon radiotherapy based on conventional linear accelerators, but we believe that some of these complexity metrics would be valid for other techniques such as robotic radiosurgery (CyberKnife) and similar concepts that reduce complexity and increase efficiency could also be applied to proton therapy.

Many commercial TPSs offer the possibility to control the number of MUs or the treatment time of the plans they produce. Some TPSs incorporate methods to restrict complexity during optimization, e.g. the aperture shape controller (ASC) tool in the Eclipse TPS [94] and the modulation factor (MF) in the TomoTherapy Hi-Art treatment system [95–97], but most TPSs still do not directly incorporate advanced complexity metrics [77]. Thus, unnecessary plan complexity is sometimes a consequence of inverse planning and better tools are needed to handle plan complexity [98]. To that aim, we encourage TPS manufacturers to:

- Minimise and better control plan complexity by incorporating complexity metrics into the optimisation algorithms. In our opinion, a combination of metrics focused on aperture modulation and aperture irregularity could be used, but further studies are necessary.
- Include tools for scoring and reporting on different aspects of the complexity of treatment plans to facilitate the handling of plan complexity during the treatment planning process.

In our opinion, implementation of such tools in TPSs will be useful for reducing and rationalising the complexity of clinical plans and their associated uncertainties, as well as for facilitating a better evaluation of the overall quality of radiotherapy treatment plans.

Plan quality

The quality of a radiotherapy treatment plan should indicate its clinical suitability, which can be assessed from the delivered dose distribution that can realistically be expected. To that purpose, the existing uncertainties in the calculation and delivery of the plan and their impact on the delivered dose must be accounted for. Plan quality is typically assessed by evaluating to what degree a 'nominal' calculated dose distribution fulfils the desired dose objectives specified in the radiation oncologist's prescription. It is important to bear in mind, however, that the dose 'on the screen' is not the dose actually delivered to the patient. Differences may arise due to limitations in the models and algorithms implemented in treatment planning systems [99], which can be even more relevant for navigated plans in Pareto fronts using multi-criteria optimisation (MCO)[100,101]. Differences can also be due to uncertainties in the delivery of the treatment plan [102,103], including those associated with the treatment unit as well as with patient set-up, intrafraction motion, and variations in the patient's anatomy [104].

As discussed in the previous sections, the impact of these uncertainties on the dose delivered to the patient depends on the robustness and complexity of treatment plans. Despite that, plan quality is commonly quantified solely in terms of the calculated dose distribution in a single scenario [25,27–30], with few exceptions where aspects related to plan robustness, complexity, and treatment efficiency were considered [59,105–108]. In our opinion, when only the calculated dose distribution in the nominal situation is examined this should be clearly acknowledged and reported; in these cases, terms such as 'dosimetric plan quality' or 'dosimetric performance' could be used to differentiate them from overall plan quality.

In conclusion, the dose distribution delivered to the patient depends not only on the calculated dose distribution but also on the robustness and complexity of the treatment plan. Therefore, all these characteristics need to be taken into account when plan quality is evaluated. Plan quality should be linked to clinical outcome and we believe that, to that aim, collecting structured data on dose metrics, robustness and complexity as well as clinical information from patients is essential. A better understanding of these concepts and further implementation of appropriate tools in commercial TPSs will facilitate the handling of these characteristics in clinical practice and be helpful for increasing the overall quality of treatment plans in radiotherapy.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Jensen K, Friborg J, Hansen CR, Samsøe E, Johansen J, Andersen M, et al. The Danish Head and Neck Cancer Group (DAHANCA) 2020 radiotherapy guidelines. Radiother Oncol 2020;151:149–51. https://doi.org/10.1016/j.radonc.2020.07.037. In press.
[2] Hansen CR, Johansen J, Samsøe E, Andersen E, Petersen JBB, Jensen K, et al. Consequences of introducing geometric GTV to CTV margin expansion in DAHANCA contouring guidelines for head and neck radiotherapy. Radiother Oncol 2018;126:43–7. https://doi.org/10.1016/j.radonc.2017.09.019.
[3] Tol JP, Delaney AR, Dahele M, Slotman BJ, Verbakel WFAR. Evaluation of a knowledge-based planning solution for head and neck cancer. Int J Radiat Oncol Biol Phys 2015;91:612–20. https://doi.org/10.1016/j.ijrobp.2014.11.014.
[4] Hussein M, Heijmen BJM, Verellen D, Nisbet A. Automation in intensity modulated radiotherapy treatment planning-a review of recent innovations. Br J Radiol 2018;1092:20180270. https://doi.org/10.1259/bjr.20180270.

[5] Tol JP, Doornaert P, Witte BI, Dahele M, Slotman BJ, Verbakel WFAR. A longitudinal evaluation of improvements in radiotherapy treatment plan quality for head and neck cancer patients. Radiother Oncol 2016;119:337–43. https://doi.org/10.1016/j.radonc.2016.04.011.
[6] Xiao Y, Papiez L, Paulus R, Timmerman R, Straube WL, Bosch WR, et al. Dosimetric evaluation of heterogeneity corrections for RTOG 0236: stereotactic body radiotherapy of inoperable stage I-II non-small-cell lung cancer. Int J Radiat Oncol Biol Phys 2009;73:1235–42. https://doi.org/10.1016/j.ijrobp.2008.11.019.
[7] ICRU. ICRU Report 83. International Commission on Radiation Units and Measurements. Prescribing, recording, and reporting photon-beam intensity-modulated radiation therapy (IMRT). ICRU Report 83. J ICRU 2010;10:1–106.
[8] ICRU. ICRU Report 91. Prescribing, recording and reporting of stereotactic treatments with small photon beams. J ICRU 2014;14:1–160.
[9] Andreo P. Dose to water-like media or dose to tissue in MV photons radiotherapy treatment planning: still a matter of debate. Phys Med Biol 2014;60:309–37. https://doi.org/10.1088/0031-9155/60/1/309.
[10] Emami B, Lyman J, Brown A, Cola L, Goitein M, Munzenrider JE, et al. Tolerance of normal tissue to therapeutic irradiation. Int J Radiat Oncol Biol Phys 1991;21:109–22. https://doi.org/10.1016/0360-3016(91)90171-y.
[11] Bentzen SM, Constine LS, Deasy JO, Eisbruch A, Jackson A, Marks LB, et al. Quantitative analyses of normal tissue effects in the clinic (QUANTEC): an introduction to the scientific issues. Int J Radiat Oncol Biol Phys 2010;76:S3–9. https://doi.org/10.1016/j.ijrobp.2009.09.040.
[12] Hansen CR, Bertelsen A, Zukauskaite R, Johnsen L, Bernchou U, Thwaites DI, et al. Prediction of radiation-induced mucositis of H&N cancer patients based on a large patient cohort. Radiother Oncol 2020;147:15–21. https://doi.org/10.1016/j.radonc.2020.03.013.
[13] Grégoire V. Tumor control probability (TCP) and normal tissue complication probability (NTCP) in head and neck cancer. Rays 2005;30:105–8.
[14] Sharabiani M, Clementel E, Andratschke N, Hurkmans C. Generalizability assessment of head and neck cancer NTCP models based on the TRIPOD criteria. Radiother Oncol 2020;146:143–50. https://doi.org/10.1016/j.radonc.2020.02.013.
[15] Hansen CR, Friborg J, Jensen K, Samsøe E, Johnsen L, Zukauskaite R, et al. NTCP model validation method for DAHANCA patient selection of protons versus photons in head and neck cancer radiotherapy. Acta Oncol (Madr) 2019;58:1410–5. https://doi.org/10.1080/0284186X.2019.1654129.
[16] Dionisi F, Fiorica F, D'Angelo E, Maddalo M, Giacomelli I, Tornari E, et al. Organs at risk's tolerance and dose limits for head and neck cancer re-irradiation: A literature review. Oral Oncol 2019;98:35–47. https://doi.org/10.1016/j.oraloncology.2019.08.017.
[17] Yaparpalvi R, Garg MK, Shen J, Bodner WR, Mynampati DK, Gafar A, et al. Evaluating which plan quality metrics are appropriate for use in lung SBRT. Br J Radiol 2018;91:20170393. https://doi.org/10.1259/bjr.20170393.
[18] Yoon M, Park SY, Shin D, Lee SB, Pyo HR, Kim DY, et al. A new homogeneity index based on statistical analysis of the dose–volume histogram. J Appl Clin Med Phys 2007;8:9–17. https://doi.org/10.1120/jacmp.v8i2.2390.
[19] Yan L, Xu Y, Chen X, Xie X, Liang B, Dai J. A new homogeneity index definition for evaluation of radiotherapy plans. J Appl Clin Med Phys 2019;20:50–6. https://doi.org/10.1002/acm2.12739.
[20] Feuvret L, Noël G, Mazeron J-J, Bey P. Conformity index: a review. Int J Radiat Oncol Biol Phys 2006;64:333–42. https://doi.org/10.1016/j.ijrobp.2005.09.028.
[21] Shaw E, Kline R, Gillin M, Souhami L, Hirschfeld A, Dinapoli R, et al. Radiation Therapy Oncology Group: radiosurgery quality assurance guidelines. Int J Radiat Oncol Biol Phys 1993;27:1231–9. https://doi.org/10.1016/0360-3016(93)90548-a.
[22] Paddick I, Lippitz B. A simple dose gradient measurement tool to complement the conformity index. J Neurosurg 2006;105:194–201. https://doi.org/10.3171/sup.2006.105.7.194.
[23] Reynolds TA, Jensen AR, Bellairs EE, Ozer M. Dose gradient index for stereotactic radiosurgery/radiation therapy. Int J Radiat Oncol Biol Phys 2020;106:604–11. https://doi.org/10.1016/j.ijrobp.2019.11.408.
[24] Giglioli FR, Garibaldi C, Blanck O, Villaggi E, Russo S, Esposito M, et al. Dosimetric multicenter planning comparison studies for stereotactic body radiation therapy: methodology and future perspectives. Int J Radiat Oncol Biol Phys 2020;106:403–12. https://doi.org/10.1016/j.ijrobp.2019.10.041.
[25] Akpati H, Kim C, Kim B, Park T, Meek A. Unified dosimetry index (UDI): a figure of merit for ranking treatment plans. J Appl Clin Med Phys 2008;9:99–108. https://doi.org/10.1120/jacmp.v9i3.2803.
[26] Ruan D, Shao W, Demarco J, Tenn S, King C, Low D, et al. Evolving treatment plan quality criteria from institution-specific experience. Med Phys 2012;39:2708–12. https://doi.org/10.1118/1.4704497.
[27] Jornet N, Carrasco P, Beltrán M, Calvo JF, Escudé L, Hernández V, et al. Multicentre validation of IMRT pre-treatment verification: comparison of in-house and external audit. Radiother Oncol 2014;112:381–8. https://doi.org/10.1016/j.radonc.2014.06.016.
[28] Ventura T, do Carmo Lopes M, Ferreira BC, Khouri L. SPIDERplan: A tool to support decision-making in radiation therapy treatment plan assessment. Reports Pract Oncol Radiother 2016;21:508–16. https://doi.org/10.1016/j.rpor.2016.07.002.
[29] Mambretti M, Romanò C, Marvaso G, Comi S, Cambria R, Ciardo D, et al. A global Unified Dosimetry Index (gUDI) to evaluate simultaneous integrated boost radiotherapy plans in prostate cancer. Radiother Oncol 2018;128:315–20. https://doi.org/10.1016/j.radonc.2018.06.002.
[30] Nelms BE, Robinson G, Markham J, Velasco K, Boyd S, Narayan S, et al. Variation in external beam treatment plan quality: an inter-institutional study of planners and planning systems. Pract Radiat Oncol 2012;2:296–305. https://doi.org/10.1016/j.prro.2011.11.012.
[31] Ahmed S, Nelms B, Gintz D, Caudell J, Zhang G, Moros EG, et al. A method for a priori estimation of best feasible DVH for organs-at-risk: Validation for head and neck VMAT planning. Med Phys 2017;44:5486–97. https://doi.org/10.1002/mp.12500.
[32] Alfonso JCL, Herrero MA, Nunez L. A dose-volume histogram based decision-support system for dosimetric comparison of radiotherapy treatment plans. Radiat Oncol 2015;10:263. https://doi.org/10.1186/s13014-015-0569-3.
[33] Moustakis C, Chan MKH, Kim J, Nilsson J, Bergman A, Bichay TJ, et al. Treatment planning for spinal radiosurgery. Strahlentherapie Und Onkol 2018;194:843–54. https://doi.org/10.1007/s00066-018-1314-2.
[34] Hardcastle N, Bignell F, Nelms B, Siva S, Kneebone A, Lao L, et al. The challenge of planning vertebral body SBRT: Optimizing target volume coverage. Med Dosim 2020;45:302–7. https://doi.org/10.1016/j.meddos.2020.02.005.
[35] Radiation Knowledge n.d. https://radiationknowledge.org/.
[36] Fusella M, Scaggion A, Pivato N, Rossato MA, Zorz A, Paiusco M. Efficiently train and validate a RapidPlan model through APQM scoring. Med Phys 2018;45:2611–9. https://doi.org/10.1002/mp.12896.
[37] Monti S, Palma G, D'Avino V, Gerardi M, Marvaso G, Ciardo D, et al. Voxel-based analysis unveils regional dose differences associated with radiation-induced morbidity in head and neck cancer patients. Sci Rep 2017;7:1–8. https://doi.org/10.1038/s41598-017-07586-x.
[38] Faught AM, Yamamoto T, Castillo R, Castillo E, Zhang J, Miften M, et al. Evaluating which dose-function metrics are most critical for functional-guided radiation therapy. Int J Radiat Oncol Biol Phys 2017;99:202–9. https://doi.org/10.1016/j.ijrobp.2017.03.051.
[39] Liang B, Yan H, Tian Y, Chen X, Yan L, Zhou Z, et al. Dosiomics: extracting 3D spatial features from dose distribution to predict incidence of radiation pneumonitis. Front Oncol 2019;9:269. https://doi.org/10.3389/fonc.2019.00269.
[40] Lakshminarayanan P, Jiang W, Robertson SP, Cheng Z, Han P, Bowers M, et al. Radio-morphology: Parametric shape-based features in radiotherapy. Med Phys 2019;46:704–13. https://doi.org/10.1002/mp.13323.
[41] Rossi L, Bijman R, Schillemans W, Aluwini S, Cavedon C, Witte M, et al. Texture analysis of 3D dose distributions for predictive modelling of toxicity rates in radiotherapy. Radiother Oncol 2018;129:548–53. https://doi.org/10.1016/j.radonc.2018.07.027.
[42] Lee SH, Han P, Hales R, Voong KR, Noro K, Sugiyama S, et al. Multi-view radiomics and dosiomics analysis with machine learning for predicting acute-phase weight loss in lung cancer patients treated with radiotherapy. Phys Med Biol 2020. https://doi.org/10.1088/1361-6560/ab8531. In press.
[43] Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology 2020;295:328–38. https://doi.org/10.1148/radiol.2020191145.
[44] Placidi L, Lenkowicz J, Cusumano D, Boldrini L, Dinapoli N, Valentini V. Stability of dosomics features extraction on grid resolution and algorithm for radiotherapy dose calculation. Phys Medica 2020;77:30–5. https://doi.org/10.1016/j.ejmp.2020.07.022.
[45] Stroom JC, de Boer HCJ, Huizenga H, Visser AG. Inclusion of geometrical uncertainties in radiotherapy treatment planning by means of coverage probability. Int J Radiat Oncol Biol Phys 1999;43:905–19. https://doi.org/10.1016/s0360-3016(98)00468-4.
[46] Van Herk M, Remeijer P, Rasch C, Lebesque JV. The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy. Int J Radiat Oncol Biol Phys 2000;47:1121–35. https://doi.org/10.1016/s0360-3016(00)00518-6.
[47] McKenzie A, van Herk M, Mijnheer B. Margins for geometric uncertainty around organs at risk in radiotherapy. Radiother Oncol 2002;62:299–307. https://doi.org/10.1016/s0167-8140(02)00015-4.
[48] Stroom JC, Heijmen BJM. Limitations of the planning organ at risk volume (PRV) concept. Int J Radiat Oncol Biol Phys 2006;66:279–86. https://doi.org/10.1016/j.ijrobp.2006.05.009.
[49] Gordon JJ, Siebers JV. Evaluation of dosimetric margins in prostate IMRT treatment plans. Med Phys 2008;35:569–75. https://doi.org/10.1118/1.2826558.
[50] Lomax AJ. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 1: The potential effects of calculational uncertainties. Phys Med Biol 2008;53:1027–42. https://doi.org/10.1088/0031-9155/53/4/014.
[51] Lomax AJ. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 2: The potential effects of inter-fraction and inter-field motions. Phys Med Biol 2008;53:1043–56. https://doi.org/10.1088/0031-9155/53/4/015.
[52] Unkelbach J, Chan TCY, Bortfeld T. Accounting for range uncertainties in the optimization of intensity modulated proton therapy. Phys Med Biol 2007;52:2755–73. https://doi.org/10.1088/0031-9155/52/10/009.
[53] Liu W, Frank SJ, Li X, Li Y, Zhu RX, Mohan R. PTV-based IMPT optimization incorporating planning risk volumes vs robust optimization. Med Phys 2013;40:21709. https://doi.org/10.1118/1.4774363.
[54] Zhang X, Rong Y, Morrill S, Fang J, Narayanasamy G, Galhardo E, et al. Robust optimization in lung treatment plans accounting for geometric uncertainty. J Appl Clin Med Phys 2018;19:19–26. https://doi.org/10.1002/acm2.12291.

[55] Unkelbach J, Alber M, Bangert M, Bokrantz R, Chan TCY, Deasy JO, et al. Robust radiotherapy planning. Phys Med Biol 2018;63:22TR02.. https://doi.org/10.1088/1361-6560/aae659.

[56] Wieser H-P, Karger CP, Wahl N, Bangert M. Impact of Gaussian uncertainty assumptions on probabilistic optimization in particle therapy. Phys Med Biol 2020;65:. https://doi.org/10.1088/1361-6560/ab8d77145007.

[57] Fredriksson A, Bokrantz R. A critical evaluation of worst case optimization methods for robust intensity-modulated proton therapy planning. Med Phys 2014;41:. https://doi.org/10.1118/1.4883837081701.

[58] Korevaar EW, Habraken SJM, Scandurra D, Kierkels RGJ, Unipan M, Eenink MGC, et al. Practical robustness evaluation in radiotherapy–A photon and proton-proof alternative to PTV-based plan evaluation. Radiother Oncol 2019;141:267–74. https://doi.org/10.1016/j.radonc.2019.08.005.

[59] Yock AD, Mohan R, Flampouri S, Bosch W, Taylor PA, Gladstone D, et al. Robustness analysis for external beam radiation therapy treatment plans: describing uncertainty scenarios and reporting their dosimetric consequences. Pract Radiat Oncol 2019;9:200–7. https://doi.org/10.1016/j.prro.2018.12.002.

[60] McGowan SE, Albertini F, Thomas SJ, Lomax AJ. Defining robustness protocols: a method to include and evaluate robustness in clinical plans. Phys Med Biol 2015;60:2671–84. https://doi.org/10.1088/0031-9155/60/7/2671.

[61] Malyapa R, Lowe M, Bolsi A, Lomax AJ, Weber DC, Albertini F. Evaluation of robustness to setup and range uncertainties for head and neck patients treated with pencil beam scanning proton therapy. Int J Radiat Oncol Biol Phys 2016;95:154–62. https://doi.org/10.1016/j.ijrobp.2016.02.016.

[62] van Dijk LV, Van den Bosch L, Aljabar P, Peressutti D, Both S, Steenbakkers Roel JHM, et al. Improving automatic delineation for head and neck organs at risk by Deep Learning Contouring. Radiother Oncol 2020;142:115–23. https://doi.org/10.1016/j.radonc.2019.09.022.

[63] LoSasso T, Chui C-S, Ling CC. Physical and dosimetric aspects of a multileaf collimation system used in the dynamic mode for implementing intensity modulated radiotherapy. Med Phys 1998;25:1919–27. https://doi.org/10.1118/1.598381.

[64] Das IJ, Ding GX, Ahnesjö A. Small fields: nonequilibrium radiation dosimetry. Med Phys 2008;35:206–15. https://doi.org/10.1118/1.2815356.

[65] Miften M, Olch A, Mihailidis D, Moran J, Pawlicki T, Molineu A, et al. Tolerance limits and methodologies for IMRT measurement-based verification QA: recommendations of AAPM Task Group No. 218. Med Phys 2018;45:e53–83. https://doi.org/10.1002/mp.12810.

[66] Johnson NF. Two's company, three is complexity. Simply Complex. A Clear Guid. to Complex. theory. Oneworld Publications; 2009. p. 3.

[67] Du W, Cho SH, Zhang X, Hoffman KE, Kudchadker RJ. Quantification of beam complexity in intensity-modulated radiation therapy treatment plans. Med Phys 2014;41:. https://doi.org/10.1118/1.4861821021716.

[68] Hernandez V, Saez J, Pasler M, Jurado-Bruggeman D, Jornet N. Comparison of complexity metrics for multi-institutional evaluations of treatment plans in radiotherapy. Phys Imaging Radiat Oncol 2018;5:37–43. https://doi.org/10.1016/j.phro.2018.02.002.

[69] McNiven AL, Sharpe MB, Purdie TG. A new metric for assessing IMRT modulation complexity and plan deliverability. Med Phys 2010;37:505–15. https://doi.org/10.1118/1.3276775.

[70] Younge KC, Matuszak MM, Moran JM, McShan DL, Fraass BA, Roberts DA. Penalization of aperture complexity in inversely planned volumetric modulated arc therapy. Med Phys 2012;39:7160–70. https://doi.org/10.1118/1.4762566.

[71] Götstedt J, Karlsson Hauer A, Bäck A. Development and evaluation of aperture-based complexity metrics using film and EPID measurements of static MLC openings. Med Phys 2015;42:3911–21. https://doi.org/10.1118/1.4921733.

[72] Kairn T, Crowe SB, Kenny J, Knight RT, Trapp JV. Predicting the likelihood of QA failure using treatment plan accuracy metrics. J Phys Conf Ser 2014;489:12051. https://doi.org/10.1088/1742-6596/489/1/012051.

[73] Crowe SB, Kairn T, Kenny J, Knight RT, Hill B, Langton CM, et al. Treatment plan complexity metrics for predicting IMRT pre-treatment quality assurance results. Australas Phys Eng Sci Med 2014;37:475–82. https://doi.org/10.1007/s13246-014-0274-9.

[74] Masi L, Doro R, Favuzza V, Cipressi S, Livi L. Impact of plan parameters on the dosimetric accuracy of volumetric modulated arc therapy. Med Phys 2013;40:. https://doi.org/10.1118/1.4810969071718.

[75] Park JM, Park S-Y, Kim H, Kim JH, Carlson J, Ye S-J. Modulation indices for volumetric modulated arc therapy. Phys Med Biol 2014;59:7315–40. https://doi.org/10.1088/0031-9155/59/23/7315.

[76] Park JM, Wu HG, Kim JH, Carlson JNK, Kim K. The effect of MLC speed and acceleration on the plan delivery accuracy of VMAT. Br J Radiol 2015;88:20140698. https://doi.org/10.1259/bjr.20140698.

[77] Chiavassa S, Bessieres I, Edouard M, Mathot M, Moignier A. Complexity metrics for IMRT and VMAT plans: a review of current literature and applications. Br J Radiol 2019;92:20190270. https://doi.org/10.1259/bjr.20190270.

[78] Antoine M, Ralite F, Soustiel C, Marsac T, Sargos P, Cugny A, et al. Use of metrics to quantify IMRT and VMAT treatment plan complexity: A systematic review and perspectives. Phys Medica 2019;64:98–108. https://doi.org/10.1016/j.ejmp.2019.05.024.

[79] Nauta M, Villarreal-Barajas JE, Tambasco M. Fractal analysis for assessing the level of modulation of IMRT fields. Med Phys 2011;38:5385–93. https://doi.org/10.1118/1.3633912.

[80] Younge KC, Roberts D, Janes LA, Anderson C, Moran JM, Matuszak MM. Predicting deliverability of volumetric-modulated arc therapy (VMAT) plans using aperture complexity analysis. J Appl Clin Med Phys 2016;17:124–31. https://doi.org/10.1120/jacmp.v17i4.6241.

[81] Vieillevigne L, Khamphan C, Saez J, Hernandez V. On the need for tuning the dosimetric leaf gap for stereotactic treatment plans in the Eclipse treatment planning system. J Appl Clin Med Phys 2019;20:68–77. https://doi.org/10.1002/acm2.12656.

[82] Hubley E, Pierce G. The influence of plan modulation on the interplay effect in VMAT liver SBRT treatments. Phys Medica 2017;40:115–21. https://doi.org/10.1016/j.ejmp.2017.07.025.

[83] De Muinck Keizer DM, Kontaxis C, Kerkmeijer LGW, Der Voort Van, Van Zyp JRN, Van Den Berg CAT, et al. Dosimetric impact of soft-tissue based intrafraction motion from 3D cine-MR in prostate SBRT. Phys Med Biol 2020;65. https://doi.org/10.1088/1361-6560/ab6241. 025012.

[84] Ezzell GA, Burmeister JW, Dogan N, LoSasso TJ, Mechalakos JG, Mihailidis D, et al. IMRT commissioning: Multiple institution planning and dosimetry comparisons, a report from AAPM Task Group 119. Med Phys 2009;36:5359–73. https://doi.org/10.1118/1.3238104.

[85] Craft D, Süss P, Bortfeld T. The tradeoff between treatment plan quality and required number of monitor units in intensity-modulated radiotherapy. Int J Radiat Oncol Biol Phys 2007;67:1596–605. https://doi.org/10.1016/j.ijrobp.2006.11.034.

[86] McGarry CK, Agnew CE, Hussein M, Tsang Y, McWilliam A, Hounsell AR, et al. The role of complexity metrics in a multi-institutional dosimetry audit of VMAT. Br J Radiol 2016;89:20150445. https://doi.org/10.1259/bjr.20150445.

[87] Jurado-Bruggeman D, Hernández V, Sáez J, Navarro D, Pino F, Martinez T, et al. Multi-centre audit of VMAT planning and pre-treatment verification. Radiother Oncol 2017;124:302–10. https://doi.org/10.1016/j.radonc.2017.05.019.

[88] Lee MT, Purdie TG, Eccles CL, Sharpe MB, Dawson LA. Comparison of simple and complex liver intensity modulated radiotherapy. Radiat Oncol 2010;5:115. https://doi.org/10.1186/1748-717X-5-115.

[89] Matuszak MM, Larsen EW, Fraass BA. Reduction of IMRT beam complexity through the use of beam modulation penalties in the objective function. Med Phys 2007;34:507–20. https://doi.org/10.1118/1.2409749.

[90] Babier A, Boutilier JJ, McNiven AL, Chan TCY. Knowledge-based automated planning for oropharyngeal cancer. Med Phys 2018;45:2875–83. https://doi.org/10.1002/mp.12930.

[91] Hansen CR, Nielsen M, Bertelsen AS, Hazell I, Holtved E, Zukauskaite R, et al. Automatic treatment planning facilitates fast generation of high-quality treatment plans for esophageal cancer. Acta Oncol (Madr) 2017;56:1495–500. https://doi.org/10.1080/0284186X.2017.1349928.

[92] Hansen CR, Bertelsen A, Hazell I, Zukauskaite R, Gyldenkerne N, Johansen J, et al. Automatic treatment planning improves the clinical quality of head and neck cancer treatment plans. Clin Transl Radiat Oncol 2016;1:2–8. https://doi.org/10.1016/j.ctro.2016.08.001.

[93] Crijns W, Defraene G, Van Herck H, Depuydt T, Haustermans K, Maes F, et al. Online adaptation and verification of VMAT. Med Phys 2015;42:3877–91. https://doi.org/10.1118/1.4921615.

[94] Scaggion A, Fusella M, Agnello G, Bettinelli A, Pivato N, Roggio A, et al. Limiting treatment plan complexity by applying a novel commercial tool. J Appl Clin Med Phys 2020;21:27–34. https://doi.org/10.1002/acm2.12908.

[95] Ryczkowski A, Piotrowski T, et al. Influence of the modulation factor on the treatment plan quality and execution time in Tomotherapy in head and neck cancer: In-phantom study. J Cancer Res Ther 2013;9:618–23.. https://doi.org/10.4103/0973-1482.126458.

[96] Boyd R, Jeong K, Tomé WA. Determining efficient helical IMRT modulation factor from the MLC leaf-open time distribution on precision treatment planning system. J Appl Clin Med Phys 2019;20:64–74. https://doi.org/10.1002/acm2.12581.

[97] Santos T, Ventura T, Mateus J, Capela M, Lopes M do C. On the complexity of helical tomotherapy treatment plans. J Appl Clin Med Phys 2020;21:107–18. https://doi.org/10.1002/acm2.12895.

[98] Kamperis E, Kodona C, Hatziioannou K, Giannouzakos V. Complexity in Radiation Therapy: It's Complicated. Int J Radiat Oncol Biol Phys 2020;106:182–4. https://doi.org/10.1016/j.ijrobp.2019.09.003.

[99] Li JS, Lin T, Chen L, Price Jr RA, Ma C-M. Uncertainties in IMRT dosimetry. Med Phys 2010;37:2491–500. https://doi.org/10.1118/1.3413997.

[100] McGarry CK, Bokrantz R, O'Sullivan JM, Hounsell AR. Advantages and limitations of navigation-based multicriteria optimization (MCO) for localized prostate cancer IMRT planning. Med Dosim 2014;39:205–11. https://doi.org/10.1016/j.meddos.2014.02.002.

[101] Kyroudi A, Petersson K, Ghandour S, Pachoud M, Matzinger O, Ozsahin M, et al. Discrepancies between selected Pareto optimal plans and final deliverable plans in radiotherapy multi-criteria optimization. Radiother Oncol 2016;120:346–8. https://doi.org/10.1016/j.radonc.2016.05.018.

[102] Schwarz M, Cattaneo GM, Marrazzo L. Geometrical and dosimetrical uncertainties in hypofractionated radiotherapy of the lung: A review. Phys Medica 2017;36:126–39. https://doi.org/10.1016/j.ejmp.2017.02.011.

[103] Liebl J, Paganetti H, Zhu M, Winey BA. The influence of patient positioning uncertainties in proton radiotherapy on proton range and dose distributions. Med Phys 2014;41:. https://doi.org/10.1118/1.4892601091711.

[104] Nassef M, Simon A, Cazoulat G, Duménil A, Blay C, Lafond C, et al. Quantification of dose uncertainties in cumulated dose estimation compared to planned dose in prostate IMRT. Radiother Oncol 2016;119:129–36. https://doi.org/10.1016/j.radonc.2016.03.007.
[105] Blanck O, Wang L, Baus W, Grimm J, Lacornerie T, Nilsson J, et al. Inverse treatment planning for spinal robotic radiosurgery: an international multi-institutional benchmark trial. J Appl Clin Med Phys 2016;17:313–30. https://doi.org/10.1120/jacmp.v17i3.6151.
[106] Masi L, Zani M, Doro R, Calusi S, Di Cataldo V, Bonucci I, et al. CyberKnife MLC-based treatment planning for abdominal and pelvic SBRT: Analysis of multiple dosimetric parameters, overall scoring index and clinical scoring. Phys Medica 2018;56:25–33. https://doi.org/10.1016/j.ejmp.2018.11.012.
[107] Villaggi E, Hernandez V, Fusella M, Moretti E, Russo S, Vaccara EML, et al. Plan quality improvement by DVH sharing and planner's experience: Results of a SBRT multicentric planning study on prostate. Phys Medica 2019;62:73–82. https://doi.org/10.1016/j.ejmp.2019.05.003.
[108] Russo S, Esposito M, Hernandez V, Saez J, Rossi F, Paoletti L, et al. Does deep inspiration breath hold reduce plan complexity? Multicentric experience of left breast cancer radiotherapy with volumetric modulated arc therapy. Phys Medica 2019;59:79–85. https://doi.org/10.1016/j.ejmp.2019.02.018.