Radiotherapy and Oncology 173 (2022) 254–261

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original Article

Plan quality assessment in clinical practice: Results of the 2020 ESTRO survey on plan complexity and robustness

Laura Patricia Kaplan [a,b,1,*], Lorenzo Placidi [c,1,*], Anna Bäck [d,e], Richard Canters [f], Mohammad Hussein [g], Ana Vaniqui [f], Marco Fusella [h], Tomasz Piotrowski [i], Victor Hernandez [j], Nuria Jornet [k], Christian Rønn Hansen [l,m,n], Lamberto Widesott [o]

[a] *Department of Oncology, Aarhus University Hospital;* [b] *Department of Clinical Medicine, Aarhus University, Denmark;* [c] *Fondazione Policlinico Universitario "A. Gemelli" IRCCS, Roma, Italy;* [d] *Department of Therapeutic Radiation Physics, Sahlgrenska University Hospital, Gothenburg;* [e] *Department of Medical Radiation Sciences, Sahlgrenska Academy, University of Gothenburg, Sweden;* [f] *Department of Radiation Oncology (Maastro), GROW School for Oncology, Maastricht University Medical Centre+, the Netherlands;* [g] *Metrology for Med Phys Centre, National Physical Laboratory, Teddington, United Kingdom;* [h] *Department of Med Phys, Veneto Institute of Oncology - IOV IRCCS, Padua, Italy;* [i] *Department of Electroradiology, Poznan University of Medical Sciences and Department of Med Phys, Greater Poland Cancer Centre, Poznan, Poland;* [j] *Department of Med Phys, Hospital Sant Joan de Reus, IISPV, Spain;* [k] *Servei de Radiofísica i Radioprotecció, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain;* [l] *Laboratory of Radiation Physics, Department of Oncology, Odense University Hospital;* [m] *Danish Centre for Particle Therapy, Aarhus University Hospital;* [n] *Department of Clinical Research, University of Southern Denmark, Odense, Denmark;* [o] *Centro di protonterapia, APSS, Trento, Italy*

ARTICLE INFO

Article history:
Received 17 February 2022
Received in revised form 24 May 2022
Accepted 7 June 2022
Available online 14 June 2022

Keywords:
International survey
Plan quality
Plan complexity
Plan robustness

ABSTRACT

Purpose: Plan complexity and robustness are two essential aspects of treatment plan quality but there is a great variability in their management in clinical practice. This study reports the results of the 2020 ESTRO survey on plan complexity and robustness to identify needs and guide future discussions and consensus.

Methods: A survey was distributed online to ESTRO members. Plan complexity was defined as the modulation of machine parameters and increased uncertainty in dose calculation and delivery. Robustness was defined as a dose distribution's sensitivity towards errors stemming from treatment uncertainties, patient setup, or anatomical changes.

Results: A total of 126 radiotherapy centres from 33 countries participated, 95 of them (75%) from Europe and Central Asia. The majority controlled and evaluated plan complexity using monitor units (56 centres) and aperture shapes (38 centres). To control robustness, 98 (97% of question responses) photon and 5 (50%) proton centres used PTV margins for plan optimization while 75 (94%) and 5 (50%), respectively, used margins for plan evaluation. Seventeen (21%) photon and 8 (80%) proton centres used robust optimisation, while 10 (13%) and 8 (80%), respectively, used robust evaluation. Primary uncertainties considered were patient setup (photons and protons) and range calculation uncertainties (protons). Participants expressed the need for improved commercial tools to control and evaluate plan complexity and robustness.

Conclusion: Clinical implementation of methods to control and evaluate plan complexity and robustness is very heterogeneous. Better tools are needed to manage complexity and robustness in treatment planning systems. International guidelines may promote harmonization.

Radiotherapy (RT) treatment plan quality assessment is a broad and complex topic. Plan complexity and robustness are two aspects within plan quality in which there currently is particular interest and controversy within the scientific community.

Radiotherapy dose distributions can be shaped to be highly conformal to a target volume using techniques such as intensity modulated radiotherapy (IMRT), volumetric modulated arc therapy (VMAT), or intensity modulated proton therapy (IMPT). Increased sculpting of the 3D dose distribution often comes hand in hand with increased modulation of many machine parameters such as output rates, gantry speed, or field shapes. This increased modulation, termed plan *complexity*, may lead to increased uncertainty in dose calculation and treatment delivery [1–9]. Inaccurate calculation and delivery of highly complex plans may negatively impact clinical outcomes. Further, they can hinder successful clinical trials

* Corresponding authors at: Department of Oncology and Palliative Units, Zealand University Hospital - Næstved Hospital, Rådmandsengen 5, 4700 Næstved, Denmark. Fondazione Policlinico Universitario "A. Gemelli", IRCCS, UOSD Fisica Medica e Radioprotezione, Ala F, piano 2, Largo Agostino Gemelli 8, 00168, Roma.

E-mail addresses: lakap@regionsjaelland.dk (L.P. Kaplan), lorenzo.placidi@policlinicogemelli.it (L. Placidi).

[1] Contributed equally to this manuscript.

https://doi.org/10.1016/j.radonc.2022.06.005
0167-8140/

and obstruct detection of dose–response effects in big data analyses. These differences between planned and delivered dose can be observed in dosimetry audits [10–12].

Several metrics have been proposed to quantify and investigate plan complexity [1–8]. Recent reviews give an overview of these metrics and published studies on their correlation with plan quality assurance results [13–15]. Some complexity metrics provide similar information and can be considered equivalent. However, metrics that focus on different plan parameters yield different results and there is no clear consensus as to which should be used [16].

At the same time, highly conformal treatment plans can come at the cost of inferior robustness towards treatment errors such as those caused by shifts in patient setup, anatomical changes, or range calculation uncertainties (for protons). A lack of robustness towards such errors might in the worst case lead to insufficient target coverage or severe side-effects due to over-dosing of critical healthy tissue.

Until a few years ago, the only universally used method to optimize and evaluate plan robustness was defining margins around the clinical target volume (CTV) and organs at risk (OAR) to obtain the planning target volume (PTV) and planning organ at risk volume (PRV), respectively. Several formulas have been proposed in the literature for the definition of PTV [17,18] and PRV [19,20] margins. The concept of safety margins has inherent limitations, however; the largest being the assumption that the spatial dose distribution does not change shape under the influence of errors (static dose cloud approximation). It has therefore been highlighted that margins are inadequate for use in particle therapy, and possibly also in highly modulated photon treatments [20–23].

Robust optimization (RO) has been suggested as a method to potentially overcome the limitations of safety margins and thus improve reliable CTV coverage and OAR sparing both for photon and proton treatments [21,24–26]. RO directly calculates the dose changes induced by simulated error scenarios, thus taking both the patient-specific anatomy and dose distribution characteristics (field directions, penumbra, dose gradient...) into account. Robust evaluation (RE) follows the same concept as RO: the optimized dose is recalculated and evaluated in simulated error scenarios.

Importantly, the terms 'complexity', 'robust optimization', and 'robust evaluation' cover several different methodologies and metrics, and there is no consensus yet on which to implement or how to use them [23,27–30]. Given the potential of these new tools and their increasing availability in treatment planning systems, the scientific community needs to discuss which methods are most appropriate.

During the third edition of the 'European Society for Radiother Oncol Physics Workshop: Science in Development' in October 2019, a dedicated track was focused on the topic 'Plan quality assessment: dose distribution and robustness metrics'. The workshop group identified treatment plan complexity and robustness as areas of special interest and controversy. The group's first aim was to map current clinical practice concerning plan complexity and robustness assessment both during the plan optimization and evaluation phases. The aim of the 2020 ESTRO online survey on plan quality assessment was to provide this overview.

In this paper, we present and analyse the results of this survey. These results can be used as the first step towards standardizing the clinical optimization and evaluation of plan complexity and robustness.

Methods

Development of survey questions began in October 2019 at the workshop. After the workshop, questions were reviewed by the entire group with a focus on clarity and completeness of multiple-choice answers. The questionnaire was implemented in an online form and trial runs were performed at some of the working group members' centres to verify the questions were understood and the survey was implemented as intended. The final web-based survey (see Supplementary Figure S1 and file "Responses.xlsx") was distributed and promoted via ESTRO mailing lists and national medical physics associations between February and May of 2020.

Respondents were asked to provide only one response per centre. In cases where several responses per centre were recorded, the most complete or (if both were equally complete) the latest response was chosen. Responding centres were included in the analysis if they had answered at least one question. Centres were excluded if any answers clearly were not intended to answer the posed question (e.g., "ABC" as an answer to an open-ended question). Single answers were disregarded if there were obvious incongruencies (e.g., respondent answered "No" to performing proton treatment planning but was still, due to errors in the survey software, shown questions concerning proton planning).

Cross-analyses were performed to investigate whether there was a correlation between how centres regarded plan complexity and robustness, respectively.

Structure of the survey

The survey consisted of four sections (see supplementary Fig. S1). The first included four general questions about the centre's location and the software and hardware used for treatment planning and delivery.

The second covered treatment plan complexity. Complexity was defined as increased modulation of many machine parameters resulting in increased demand on the accuracy of dose calculation and treatment delivery. Centres were asked for which types of treatments they considered plan complexity to be an issue (if any), and how they controlled and evaluated it. No distinction was made between photon and proton treatments in this section.

The third section was focused on plan robustness. Robustness was defined as the dose distribution's sensitivity towards errors stemming from treatment uncertainties, including patient setup, anatomical changes, and proton range uncertainties. The survey aimed at answering two overarching questions about plan robustness: (1) which methods and metrics do centres use to control and evaluate plan robustness? And (2) how do the two main approaches of controlling and evaluating robustness, safety margins and RO/RE, compare in terms of uncertainty types considered and sizes of margins/simulated patient shifts?

Finally, responding centres were asked which other methods and metrics they used to assess plan quality (with emphasis on dosimetric quality), and whether they had any requests for future implementations in commercial treatment planning systems.

Results

General

A total of 126 radiotherapy centres (of 140 recorded responses) were included in the analysis. Of these, 95 (75 %) were from Europe and Central Asia. Details on the characteristics of participating centres are provided in supplementary Tables S1-S5. The number of responses for each separate question varied between 4 and 126. A spreadsheet detailing all questions, responses, and reasons for excluding answers is provided in the supplementary material ("Responses.xlsx"). Percentages stated in the following are in relation to the number of responses given to the respective question. For evaluation of a plan's dosimetric quality (excluding complexity

and robustness), visual inspection and use of DVH curves and metrics were the methods used most widely (66, 69, and 64 of 76 responses; Fig. 1, top).

Complexity

Plans in the head-and-neck region were considered most challenging with respect to plan complexity by most responding centres (72 of 95). Machine output (monitor units per Gray) and

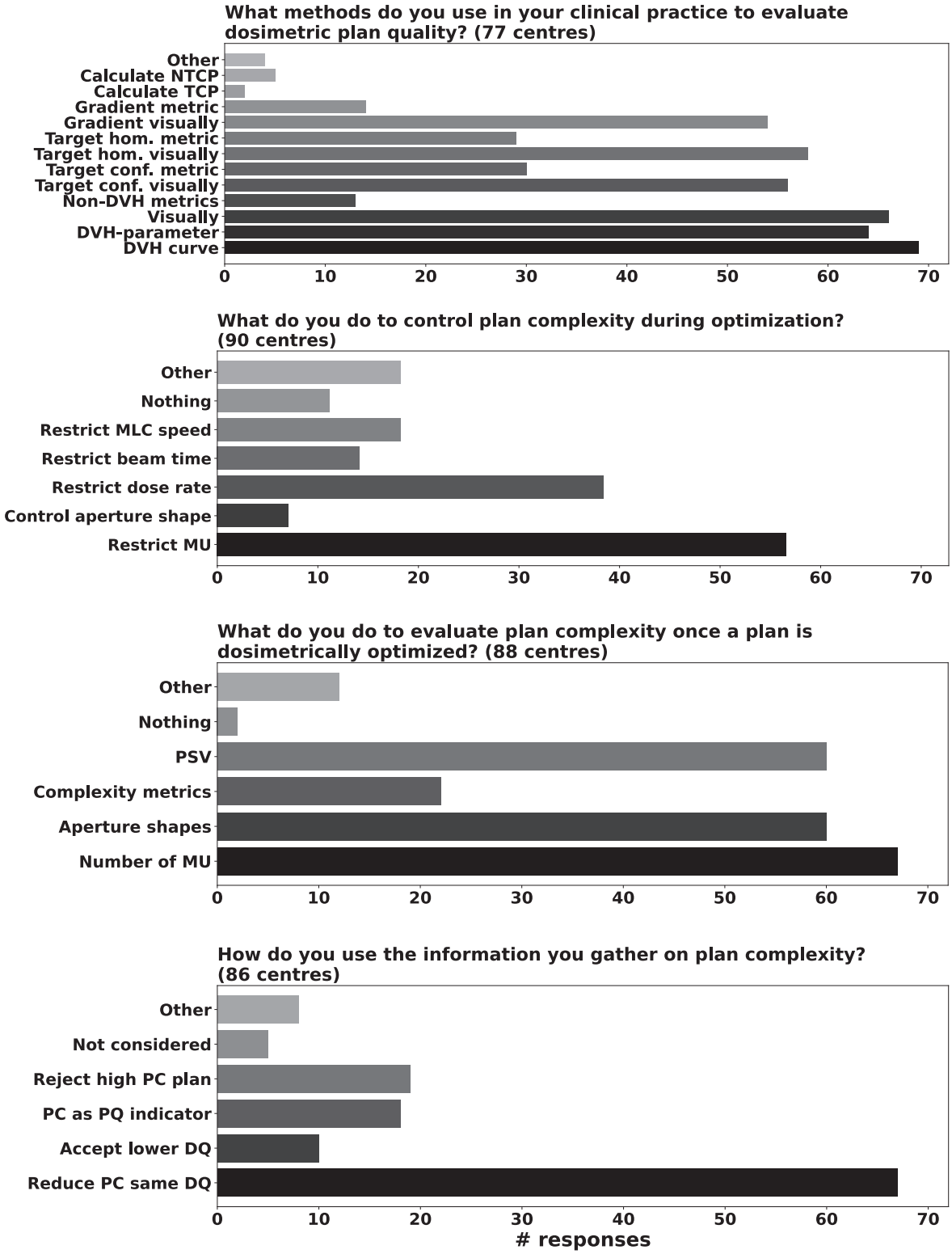

Fig. 1. Plan complexity. Fig. 1: Responses to the questions "Which methods (other than those covered in previous questions) do you in your clinical practice use to evaluate a plan's dosimetric quality?" (a), "How is complexity controlled during optimization?" (b), "How is complexity evaluated after optimization?" (c), and "What do you do with the information gained on complexity?" (d). DQ: Dosimetric quality, PC: Plan complexity, PSV: patient specific verification, PQ: Plan quality, Conf: conformity, Hom.: Homogeneity.

aperture shapes were the methods mentioned most often to control and evaluate complexity (Fig. 1). Additionally, patient-specific verification measurements were reported equally often as an evaluation method.

A large proportion of centres reported that they tried reducing the complexity of highly complex plans if this was possible while keeping dosimetric plan quality constant (67 of 86). Ten centres would even accept lower dosimetric quality in exchange for lower complexity.

Twenty-five centres reported that they did not take complexity into account as part of their clinical evaluation routine at all, due to:

- lack of clear guidance on how to evaluate complexity (14 answers)
- because they had no tools to do so (11 answers)
- because there was no consensus on which metrics to use (8 answers)
- because it was considered too time-consuming (5 answers).

Robustness

Table 1 shows an overview of reported methods used to control and evaluate plan robustness in photon and proton planning.

Of 105 centres responding to at least one question in the robustness section, 102 performed photon treatment planning and 10 performed proton treatment planning. Seven clinics performed both photon and proton planning.

Photon centres

Ninety-eight photon centres reported using PTV margins for plan optimization, while only 64 used PRV. Substantially more centres used safety margins for plan optimization and evaluation than RO and RE (Table 1). Of those photon centres using RO/RE, all but one reported using both RO/RE and PTV/PRV.

PTV margin sizes were mostly determined based on historical choices (32 of 85, Fig. 2) or using a formula such as the one by van Herk *et al.* [18] (26 of 85). PRV margin sizes were more often determined based on a historical choice than by a formula (35 vs 4, 60 respondents). RO scenarios were calculated most often by assigning fraction-specific uncertainty scenarios and summing these (7 of 17 responses).

The majority of centres reported including possible setup errors in their considerations for margins or robustness shifts (Fig. 3). The distribution of other possible error sources included in margin/shift considerations varied between centres using margins vs those using RO/RE (Fig. 3). For example, 48 of 85 (56 %) and 27 of 60 (45 %) respondents reported including couch positioning or rotation uncertainties in their PTV and PRV margins, while 4 of 17 (24 %) and 1 of 10 (10 %) reported including these in their RO and RE, respectively.

Reported PRV margin sizes were overall smaller than PTV (Supplementary Figure S2). Two centres reported using different shift sizes for RE than for RO. Reported RO shift sizes tended to be smaller than PTV sizes in thoracic and pelvic treatment sites (Supplementary Figure S3).

Table 1
Overview of participants who report using margins or robust optimization/evaluation to control and evaluate treatment plan robustness. Figures in parentheses show the total number of centres responding to each question. RO: robust optimization, RE: robust evaluation.

	Photons	Protons
Optimization		
Use PTV margin	98 (101)	5 (10)
Use RO for target	17 (82)	8 (10)
Of those using RO, also use PTV	16 (17)	3 (8)
Use PRV margin	64 (88)	5 (10)
Use RO for OAR	10 (17)	8 (8)
Of those using RO for OAR, also use PRV	7 (10)	3 (8)
Evaluation		
Use PTV/PRV	75 (80)	1 (7)
Use RE	10 (78)	8 (10)
Answered "No" but commented "Occasionally" or similar	4 (78)	
Of those using RE, also use PTV/PRV	10 (10)	1 (8)

Proton centres

Of the 10 responding proton centres, more reported using RO and RE than PTV or PRV (Table 1). Four centres reported using both PTV and PRV margins for plan optimization, while one centre answered "Yes" only to using PTV and another only to using PRV margins. Four centres reported using both RO/RE and either PTV or PRV (two using both PTV and PRV, one using only PTV, and one using only PRV).

Most centres reported including setup uncertainties in their margin or shift considerations (Figs. 2 and 3). Range uncertainties were included by all centres using RO and seven of eight using RE, while only one centre using PTV and another using PRV considered range (both centres also reported using RO and RE).

All centres except two reported using the same shift sizes for RE as for RO (one used smaller shifts for RE and one did not use RO).

Cross-analysis

Of photon centres that considered complexity, a larger fraction performed RO for targets (19 %), RO for OARs (13 %) and RE (13 %) compared to those that did not consider complexity.

A higher percentage of proton centres included complexity (91 %) in their clinical planning routines compared to photon centres. Of those proton centres who took plan complexity into account, most used RO for targets (70 %), RO for OARs (80 %) and RE (70 %). (Supplementary Figures S4 and S5).

Suggestions for improvements of commercial complexity and robustness tools

When asked whether they needed better commercial tools for plan complexity management, 61 of 109 respondents requested improved tools for controlling complexity during plan optimization, while 70 required tools to evaluate complexity after optimization. Twenty-four centres did not see the need for improvement in currently available tools.

Open-answer suggestions to improve RO and RE tools mainly aimed at making robustness scenarios less simplistic and/or conservative. Suggestions included:

a) Rather than using conservative "worst-case" scenarios, RO/RE scenarios should be based on error distributions (5 comments).
b) Various possible error sources such as couch pitch, roll, or rotation and anatomical variations should be included in tools (2 comments).
c) Better tools for visualization and reporting (1 comment).
d) Tools for RO/RE should be faster (2 comments).

Discussion

The results of this international survey paint a picture of how RT plan complexity and robustness are considered in clinics (as of 2020).

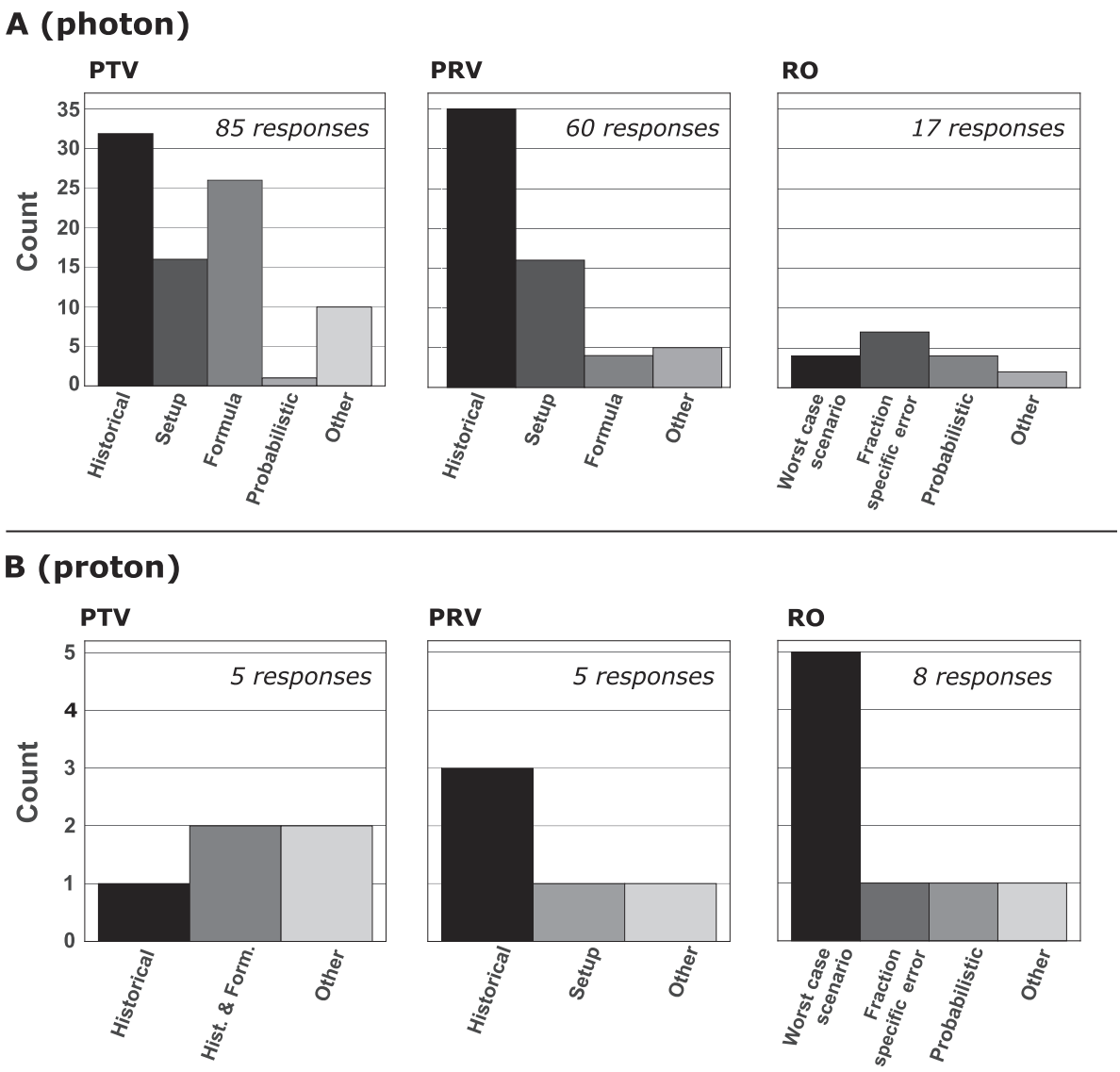

Fig. 2. How are safety margins and RO calculated?. Fig. 2: Methods used by photon (A) and proton (B) centres to calculate safety margins and robust optimization shifts. Answer options for PTV/PRV were "Based on a historical approach", "Translate setup errors directly into margin", "According to the formula by van Herk *et al.* or others", "We use a probabilistic approach", and "Other" (with the option to specify). "Other" responses included e.g. "A combination of several approaches", "Depends on the diagnosis", "Depends on the specific machine uncertainty", and "Depends on the physician".

We saw that different methods, often more than one, were employed in clinical practice to evaluate a plan's dosimetric quality. Some of the evaluations were performed by inspecting the spatial 3D dose distribution visually. Evaluating the DVH curve and specific DVH metrics remained the dominant methods to assess dosimetric quality.

Additional methods have recently been proposed which could aid in better quantifying and objectively assessing a plan's dosimetric quality. Tools such as a total plan quality index (PQI), which collapses several indices into an overall plan score [31–35], the use of machine and deep learning [36–38], and new kinds of dose distribution metrics [39–43] can provide quantification of parameters which are currently mainly evaluated subjectively. We believe using such tools will facilitate consensus and standardization on how to define a treatment plan's quality.

An interesting result of this survey was that 26 of the responding centres (21 %) did not consider plan complexity at all in their daily practice. Only five of these centres stated that the reason for this was lack of time and resources. Instead, the reason most often given for why plan complexity was not considered was because there was a lack of clear guidance and consensus on how to evaluate complexity. The different approaches used by participants to control and evaluate complexity also highlight the absence of both international consensus and of any known tolerance levels for the different metrics. A method to translate complexity into plan quality and deliverability could facilitate the use of plan complexity in clinical practice.

Even though we saw different approaches among centres on how to deal with plan complexity, there was nevertheless a strong and shared demand for access to commercial tools able to control and evaluate complexity during and after optimization. Several other authors [3,44] suggest in particular incorporating complexity metrics into optimisation algorithms. Quantification of plan complexity using dedicated metrics can prevent unacceptably high dosimetric uncertainties - which could lead to substantial differences between planned and delivered dose if undetected. For this reason, the RATING guidelines for planning studies [45] explicitly recommend quantitative reporting on the complexity of treatment

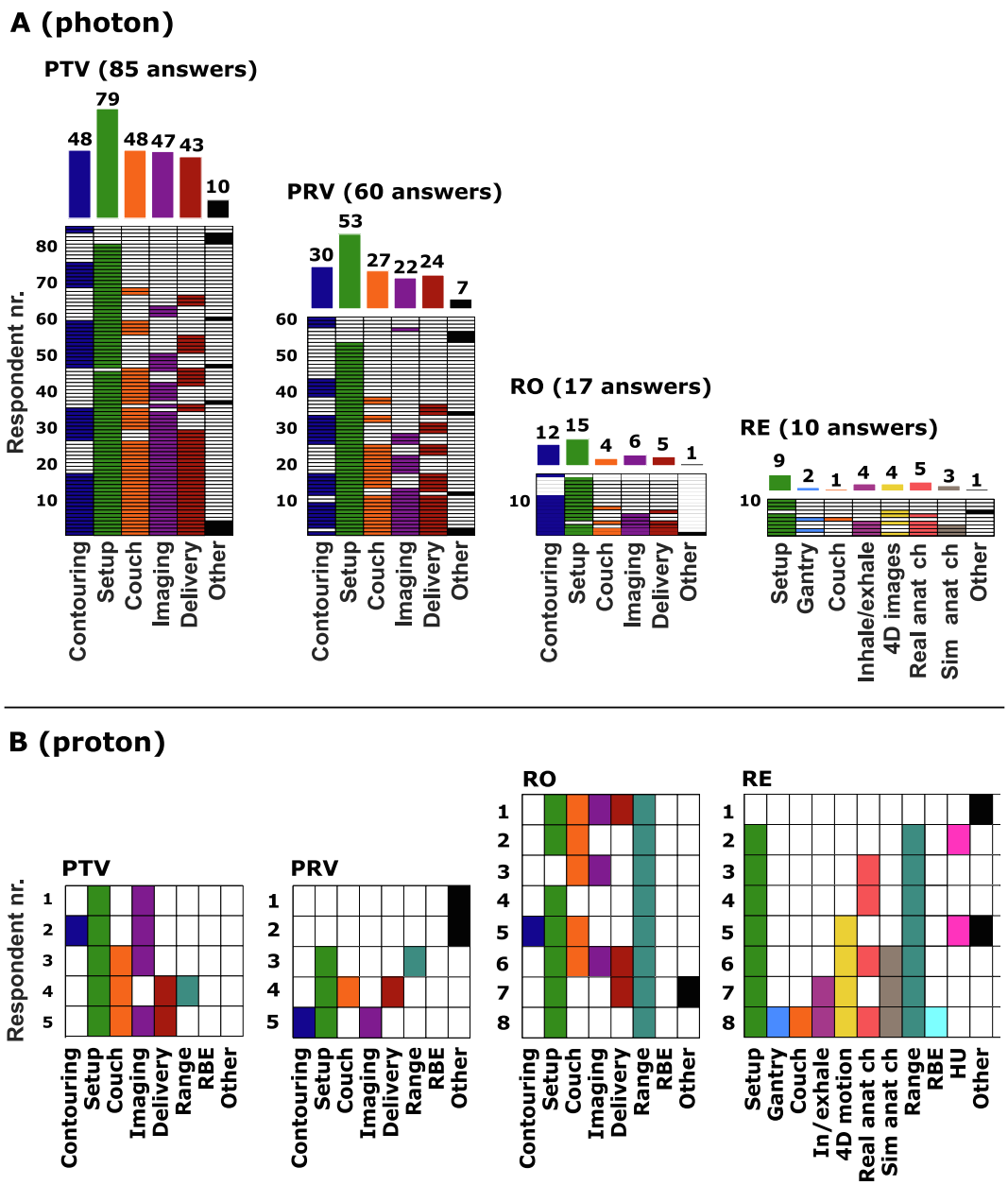

Fig. 3. Which uncertainties are included in margin/shift considerations?. Fig. 3: Uncertainties included in the margin or robust optimization/evaluation considerations by participating photon (A) and proton (B) centres. Each row corresponds to one responding centre. Numbers do not correspond to the same centre in each of the four plots. Abbreviations: anat.: anatomical, ch.: changes, sim.: simulated.

plans. Better commercial tools for controlling complexity during optimisation, in particular for automated planning, are highly desirable. For evaluation, vendors are encouraged to enable calculation and reporting of complexity metrics.

In the questions concerning plan complexity, no distinction was made between photon and proton centres. Traditionally, plan complexity has been described mainly in the context of photon treatments with conventional linear accelerators, but some complexity metrics may be valid for other techniques such as robotic radiosurgery [46]. We believe similar concepts that reduce complexity and increase efficiency could also be applied to proton therapy. Two of three responding centres who performed only proton planning commented that they evaluated complexity in terms of spot weighting and spot positions, while one answered that they did not control or evaluate complexity for proton plans since there was no data on the subject and they considered all modulated spot-scanning proton plans equally complex.

Concerning plan robustness, we saw large variations between responding centres in how safety margin and RO/RE shift sizes were chosen, which approach was used for RO, and which types of uncertainties were considered in margin or RO/RE shift sizes. This variability highlights the need for broad international consensus and clear guidelines on how plan robustness should be optimized, evaluated, and reported.

Some respondents admitted to not knowing how margin sizes were specified, which uncertainties were considered, or the actual margin sizes, as these decisions were made solely by the treating physicians. This is troubling, since detailed knowledge of the underlying physics is essential to understand the effect of uncertainties and errors on the delivered dose. As is recommended in

the International Commission on Radiation Units and Measurements Report 62 [47], medical physicists should always be involved in decisions of this nature alongside radiation oncologists.

A substantial proportion (21 %/13 %) of photon centres reported using RO or RE, some routinely and some for select cases only. Most of these also used margins for optimization and evaluation of plans, though. This suggests that the photon community is starting to embrace the new paradigm of RO and RE but has not yet adopted it to the same degree as the proton community has.

Most centres using RO and RE stated that they used the same shift sizes for both. It has previously been suggested that using the same error distributions to optimize and evaluate treatment plans may introduce an estimation bias [48]. Evaluation with different error distributions than those used for optimization is therefore recommended.

Moreover, many centres reported the same shift sizes for simulating systematic and random errors. We believe this to be due to a misunderstanding of the question.

Caution is recommended when directly comparing RO/RE shift sizes to margin sizes. It is difficult to translate the PTV concept, a margin providing 95 % of a prescribed dose to the CTV in 90 % of the treated population, to the RO framework. Most currently available commercial RO solutions provide only uncertainty scenarios based on setup and range errors (anatomical variations have only been implemented by one of the major vendors to date). PTV margins, on the other hand, directly incorporate the known distributions of errors from several other sources (e.g. contouring or machine uncertainties). Attempts at a PTV-to-RO translation have been made by e.g. Korevaar *et al.* [27] or Perkó *et al.* [49]. The ability to translate the underlying goal of the PTV to RO will be important in future studies to connect "old" PTV-based outcomes to those of RO-based plans.

In the transition from PRV to RO/RE for organs at risk, two crucial points must be taken into consideration: 1) To date, in commercial treatment planning systems the target and OARs are robustly optimized using the same parameters. This is in contrast to the tradition of some centres to use smaller margins for organs at risk than the target [50], which is also reflected in our survey results. 2) Toxicity data and NTCP models are based on the nominal DVH of organs at risk. Requiring the same (or even lower) OAR dose limits than those published by QUANTEC in all uncertainty scenarios may be too restrictive. Research into this issue is needed to translate OAR dose constraints to the RO/RE framework [26].

Respondents pinpointed the need to define more realistic and versatile uncertainty scenarios for RO and RE including, for instance, errors caused by anatomical variations or mechanical and contouring uncertainties. Further, the need for faster tools was highlighted to make the requested precision compatible with clinical workflows. We strongly believe improvement in these areas will lead to wider adoption of the RO/RE framework.

The survey showed a tendency for photon centres that considered plan complexity important to more often use RO and/or RE. However, due to the small number of responses, it is difficult to draw statistically rigorous conclusions on this question. In the context of another large survey, it was hypothesized that centres with more staffing resources were more likely to implement "non-standard" techniques (such as complexity assessment and RO/RE) [51]. We did not include a question concerning centre size and resources in this survey, and were unable to gather this information subsequently.

Based on the results of this study, we recommend that centres moving from using margins to robust optimization and robust evaluation carefully consider how to include the various types of uncertainties in the new framework to maximize plan robustness towards relevant uncertainties and improve transparency in reporting results. Commercial tools must facilitate the inclusion of various types and sources of errors.

A limitation of this study lies in the geographical distribution of the participants: most respondents were from European countries, possibly giving rise to bias. Moreover, although a fair number of centres responded to the survey, this represents a small proportion of operating radiotherapy clinics and may not fully capture patterns of clinical practice at a global level. In the future, it would be desirable to include more centres from a larger number of countries and continents in such a survey. This includes especially countries with a higher density of particle centres, since these centres have a long tradition of using robust optimization and evaluation methods.

This working group aims to support work towards broad consensus regarding which metrics and methods are used to control and assess plan complexity and robustness. The results obtained in this survey highlight trends in how complexity and robustness are currently controlled and evaluated. Further, they show that there is a need for both practical education and broader discussion concerning these topics, as well as for development of commercial tools to facilitate their assessment in clinical practice.

In the future, it will be important to evaluate the impact of such new tools through carefully designed studies [44]. Shared consensus on metrics will be essential for the performance and interpretability of such studies. International guidelines on how to manage plan complexity and robustness in the different steps of the radiotherapy workflow are strongly desired and, to that aim, further efforts of the community are warranted.

Funding declarations

LPK was supported by DCCC Radiotherapy - The Danish National Research Centre for Radiotherapy, Danish Cancer Society (Grant No. R191-A11526) and Danish Comprehensive Cancer Centre, and the Danish Cancer Society (Grant No R167-A11003).

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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.radonc.2022.06.005.

References

[1] McNiven AL, Sharpe MB, Purdie TG. A new metric for assessing IMRT modulation complexity and plan deliverability. Med Phys 2010;37:505–15.
[2] Nauta M, Villarreal-Barajas JE, Tambasco M. Fractal analysis for assessing the level of modulation of IMRT fields. Med Phys 2011;38:5385–93.
[3] 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.
[4] 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.
[5] 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.
[6] 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.
[7] 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.
[8] 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.

[9] Li JS, Lin T, Chen L, Price RA, Ma C-M. Uncertainties in IMRT dosimetry. Med Phys 2010;37:2491–500.
[10] Kouloulias VE, Poortmans PM, Bernier J, Horiot J-C, Johansson K-A, Davis B, et al. The Quality Assurance programme of the Radiotherapy Group of the European Organization for Research and Treatment of Cancer (EORTC): a critical appraisal of 20 years of continuous efforts. Eur J Cancer 2003;39:430–7.
[11] Ibbott GS, Followill DS, Molineu HA, Lowenstein JR, Alvarez PE, Roll JE. Challenges in credentialing institutions and participants in advanced technology multi-institutional clinical trials. Int J Radiat Oncol Biol Phys 2008;71:S71–5.
[12] Budiharto T, Musat E, Poortmans P, Hurkmans C, Monti A, Bar-Deroma R, et al. Profile of European radiotherapy departments contributing to the EORTC Radiation Oncology Group (ROG) in the 21st century. Radiother Oncol 2008;88:403–10.
[13] 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.
[14] 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.
[15] 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 Med 2019;64:98–108.
[16] 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.
[17] 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.
[18] 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.
[19] McKenzie A, van Herk M, Mijnheer B. Margins for geometric uncertainty around organs at risk in radiotherapy. Radiother Oncol 2002;62:299–307.
[20] Stroom JC, Heijmen BJM. Limitations of the planning organ at risk volume (PRV) concept. Int J Radiat Oncol Biol Phys 2006;66:279–86.
[21] 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.
[22] Wagenaar D, Kierkels RG, Free J, Langendijk JA, Both S, Korevaar EW. Composite minimax robust optimization of VMAT improves target coverage and reduces non-target dose in head and neck cancer patients. Radiother Oncol 2019;136:71–7.
[23] Unkelbach J, Alber M, Bangert M, Bokrantz R, Chan TC, Deasy JO, et al. Robust radiotherapy planning. Phys Med Biol 2018;63. 22TR02.
[24] 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.
[25] Zhang X, Rong Yi, 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.
[26] Noufal MP, Widesott L, Sharma SD, Righetto R, Cianchetti M, Schwarz M. The role of plan robustness evaluation in comparing protons and photons plans - an application on IMPT and IMRT plans in skull base chordomas. J Med Phys 2020;45:206–14.
[27] 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.
[28] 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.
[29] 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.
[30] 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.
[31] 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.
[32] Ventura T, Lopes MdC, Ferreira BC, Khouri L. SPIDERplan: a tool to support decision-making in radiation therapy treatment plan assessment. Rep Pract Oncol Radiother 2016;21:508–16.
[33] Akpati H, Kim ChangSeon, 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.
[34] 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.
[35] 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.
[36] Kalet AM, Luk SM, Phillips MH. Radiation therapy quality assurance tasks and tools: the many roles of machine learning. Med Phys 2020;47:e168–77.
[37] Fan J, Xing L, Ma M, Hu W, Yang Y. Verification of the machine delivery parameters of a treatment plan via deep learning. Phys Med Biol 2020;65. 195007.
[38] Pillai M, Adapa K, Das SK, Mazur L, Dooley J, Marks LB, et al. Using artificial intelligence to improve the quality and safety of radiation therapy. J Am Coll Radiol 2019;16:1267–72.
[39] 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.
[40] 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 Med 2020;77:30–5.
[41] Vaniqui A, Canters R, Vaassen F, Hazelaar C, Lubken I, Kremer K, et al. Treatment plan quality assessment for radiotherapy of rectal cancer patients using prediction of organ-at-risk dose metrics. Phys Imaging Radiat Oncol 2020;16:74–80.
[42] Kaplan LP, Korreman SS. A systematically compiled set of quantitative metrics to describe spatial characteristics of radiotherapy dose distributions and aid in treatment planning. Physica Medica 2021;90:164–75.
[43] 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.
[44] Hernandez V, Hansen CR, Widesott L, Bäck A, Canters R, Fusella M, et al. What is plan quality in radiotherapy? The importance of evaluating dose metrics, complexity, and robustness of treatment plans. Radiother Oncol 2020;153:26–33.
[45] Hansen CR, Crijns W, Hussein M, Rossi L, Gallego P, Verbakel W, et al. Radiotherapy Treatment plannINg study Guidelines (RATING): A framework for setting up and reporting on scientific treatment planning studies. Radiother Oncol 2020;153:67–78.
[46] Masi L, Hernandez V, Saez J, Doro R, Livi L. Robotic MLC-based plans: A study of plan complexity. Med Phys 2021;48:942–52.
[47] Landberg T, Chavaudra J, Dobbs J, Gerard J-P, Hanks G, Horiot J-C, et al. ICRU Report 62. J ICRU 1999;os32:NP–P.
[48] Bohoslavsky R, Witte MG, Janssen TM, van Herk M. Probabilistic objective functions for margin-less IMRT planning. Phys Med Biol 2013;58:3563–80.
[49] Perkó Z, van der Voort SR, van de Water S, Hartman CMH, Hoogeman M, Lathouwers D. Fast and accurate sensitivity analysis of IMPT treatment plans using Polynomial Chaos Expansion. Phys Med Biol 2016;61:4646–64.
[50] Tommasino F, Widesott L, Fracchiolla F, Lorentini S, Righetto R, Algranati C, et al. Clinical implementation in proton therapy of multi-field optimization by a hybrid method combining conventional PTV with robust optimization. Phys Med Biol 2020;65. 045002.
[51] Healy BJ, Budanec M, Ourdane B, Peace T, Petrovic B, Sanz DE, et al. An IAEA survey of radiotherapy practice including quality assurance extent and depth. Acta Oncol 2020;59:503–10.