Physics and Imaging in Radiation Oncology 37 (2026) 100900

Contents lists available at ScienceDirect

Physics and Imaging in Radiation Oncology

journal homepage: www.sciencedirect.com/journal/physics-and-imaging-in-radiation-oncology

Guidelines

An ESTRO-EPTN Delphi consensus on robustness evaluation in proton therapy

ARTICLE INFO

Keywords:
Proton therapy
Consensus
Robustness evaluation
European particle therapy network

ABSTRACT

Background and purpose: Robustness evaluation (RE) is vital for proton treatment planning, but lacks international consensus or guidelines, with clinics using varied, self-developed methods focused on selected uncertainties. This ESTRO project surveys expert opinions on clinical RE methods to inform future treatment planning system (TPS) development.

Materials and methods: A study within the European Particle Therapy Network (EPTN) involved 24 European proton therapy centres, with one radiation oncologist and one medical physicist per centre. The goal was to reach a consensus on transitioning from Planning Target Volume (PTV)-based planning to robustly optimized planning, including uncertainties, methods, and reporting of robustness evaluations. An internal committee drafted 39 statements, reviewed by an independent committee. Following a two-round Delphi procedure, consensus was set at a 75% agreement threshold.

Results: Twenty of 24 contacted centers (83.0%) responded to both questionnaire rounds. Consensus was reached on 26 of 39 statements (66.7%), with 5 being high-priority. Strong agreement emerged regarding which uncertainties to include in RE (range, setup, intra-fraction, anatomy changes), methodologies (e.g., for moving targets, combining setup and range), and how to report RE results clinically. Disagreement was found on using the PTV for both planning and dose reporting. The results also offer important implications for TPS vendors and future software development.

Conclusions: The ESTRO Delphi consensus may serve as practical guidance on points where a clear consensus was achieved. For remaining points, the development of guidelines is recommended to standardize methodologies and reporting. Furthermore, TPS vendors are encouraged to align their developments with the community's articulated requirements.

1. Introduction

Proton therapy (PT) has emerged as a promising modality in radiotherapy, offering superior dose conformity compared to conventional X-ray radiotherapy due to the unique physical properties of protons, such as the Bragg peak. However, the clinical implementation of PT is accompanied by significant challenges, particularly in managing uncertainties related to patient setup, anatomical changes, and proton range accuracy. These uncertainties can lead to deviations between planned and delivered dose distributions, potentially compromising target coverage and normal tissue sparing. As a result, the concept of robustness in treatment planning has gained considerable attention in recent years, with a growing emphasis on developing methodologies to evaluate and optimize the robustness of proton therapy plans.

The importance of alternative planning and evaluation approaches, such as robust optimization and robustness evaluation of the treatment plan, that explicitly account for uncertainties in treatment delivery, moving beyond PTV-based methods has been emphasized [1]. This marked a pivotal shift from margin-based strategies to more sophisticated methods that integrate uncertainty scenarios into the optimization process. Both Hernandez et al. [1] and the survey of Kaplan et al. [2] called for a consensus on metrics and methodologies for robustness evaluation, highlighting the need for standardized tools to assess the impact of uncertainties on dose distributions.

Most recently, a systematic review provided a thorough analysis of robustness evaluation methodologies in proton therapy [3]. The review highlighted the widespread adoption of robustness evaluation tools in clinical practice, while also identifying gaps in current approaches, such as the limited consideration of fractionation effects and contouring uncertainties. The review called for the development of more comprehensive and clinically viable robustness evaluation methods, setting the stage for future consensus-building efforts. The concept for the consensus was developed and designed within the European Particle Therapy Network, Work Package 5 (EPTN-WP5), which focuses on treatment planning. Through ESTRO, a broad participation of European particle centres was achieved, aiming to lay a strong foundation for establishing a consensus on robustness evaluation in proton therapy. The initiative also sought to identify key topics in need of standardized guidelines for methodology and reporting. The second aim of this study is guiding future developments in treatment planning systems (TPS) in the robustness evaluation field with solid input from the clinical community.

https://doi.org/10.1016/j.phro.2025.100900
Received 21 October 2025; Received in revised form 16 December 2025; Accepted 28 December 2025
Available online 30 December 2025

2. Materials and methods

2.1. Consensus group description

Twenty-four (24) centres from 14 European countries were contacted to be part of the study. All were centres already treating patients with a wide variety of clinical experience (minimum few months, maximum 29 years). Twenty of them (83 %), from 12 different countries, agreed to contribute to this project. From each centre, one radiation oncologist (RadOnc) and one medical physicist (MP) were requested to participate as reference persons for the consensus with the option to involve the entire local clinical team into the discussion. In total, 42 people actively contributed to this consensus group. Details of the group are reported in Table 1. A subgroup of 9 experts was selected as writing committee to define the consensus statements (4 RadOnc, 5 MP) and another independent group of 10 experts was selected as review group (3 RadOnc, 6 MP and 1 radiation therapist – RTT). The entire Delphi consensus process is summarized in Fig. 1. If responders required clarification on statements or suggested changes to their formulation after the first consensus round, a second round was conducted. In case no change was made on the sentences, only responders requesting the clarification were contacted for the second round.

2.2. Consensus workflow and criteria

Using the systematic review performed by Sterpin et al. [3] and the personal clinical experience of the writing committee members, 39 statements were formulated and divided in two groups: Basic Statements (29) and Advanced Statements (10). The *Basic Statements* group was designed for immediate clinical application, aiming to establish a consensus on standard practices that would be followed if commercial tools were universally available across all clinics. The *Advanced Statements* group was developed based on recent publications in the field of robustness evaluation, where new tools or approaches were proposed but not yet clinically available. This set of statements aimed to assess if an agreement could be reached on how the particle therapy community sees the future (intended as the next five years) in the field of robust treatment planning in order to provide clear, evidence-based guidance for vendor roadmaps and research and development efforts and to help guide future product developments in the TPS. The two statement groups were divided (when possible) into four sections: *From PTV based planning to Robust optimization planning*: This is an introductory section to evaluate if a consensus on the transition methodology from classical planning (PTV-based) to robust optimization planning exists and should be followed. For this section, only basic statements are listed. *Type of uncertainties to consider*: The purpose of this section is to find an agreement on the sources of uncertainties that should be considered during robustness evaluation both in basic and advanced approaches. *Methodologies*: The aim of this section is to identify the best method to run a Robustness Evaluation (i.e. worst-case approach, probabilistic approach, methods to evaluate the magnitude of the uncertainties simulated etc.). *Reporting*: The statements in this section are intended for robustness evaluations performed with a worst-case scenario approach: all the errors are simulated as systematic. For this section, only basic statements are listed.

For each basic statement, the responders had five answer options listed below along with corresponding explanations:

- *Agree with high priority*. This answer means the responders agree with the entire sentence. With this choice, they declare that the content of the sentence is what they are already doing in their clinical practice or what they are going to implement in the very near future since the tools are available and they can be used for the implementation of the method/procedure declared in the sentence.
- *Agree*. This answer means the responders agree with the sentence but they do not have the tools to implement such an approach in their clinical practice.
- *Neither agree nor disagree*. This choice is used to declare that the sentence is not clear and they do not have the possibility to agree or disagree. In this case, it was recommended to compile the "Comment" section of the form in order to express what isn't clear in the sentence in order to adjust it in a second round of answers.
- *Disagree*. This answer means the responders do not agree with the sentence, or in their clinical practice a different method is adopted or the method proposed is considered outdated but is unlikely to be detrimental to the patient's treatment.
- *Strongly disagree*: This answer means the responders do not agree with the sentence or they strongly believe that if the method proposed is adopted in clinical practice it may be detrimental to the patient's treatment.

For advanced statements, a set of four possible answers were defined as detailed below:

- *Agree with high priority*. The responders do agree with the sentence and they believe that the implementation of the proposed method must have a high priority in future TPS versions and in clinical practice.
- *Agree*. The responders do agree with the sentence and they believe that the implementation of this method must have a low priority in future TPS versions and in clinical practice with respect to other, more important, topics.
- *Neither agree nor disagree*. This choice is used to declare that the sentence is not clear and they cannot agree or disagree. In this case, it is recommended to compile the "Comment" section of the form to express what isn't clear in the sentence in order to adjust it in a second round of answers.
- *Disagree*: The responders do not agree with the sentence and, even if available, they will not use such an approach in their clinical practice.

Table 1
General information about consensus group, writing committee and review committee composition.

Consensus Group		n (%)
Gender	Male	23 (54.8)
	Female	19 (45.2)
Profession	Radiation Oncologist	18 (42.9)
	Medical Physicist	23 (54.8)
	RTT	1 (2.4)
Centres per Country	Belgium	1
	Denmark	1
	France	2
	Germany	2
	Italy	3
	Norway	2
	Poland	1
	Spain	1
	Sweden	1
	Switzerland	1
	The Netherlands	3
	United Kingdom	2
In clinical activity	Yes	20 (100.0)
	Not yet	0 (0.0)
Writing Committee		
Gender	Male	7 (77.8)
	Female	2 (22.2)
Profession	Radiation Oncologist	4 (44.4)
	Medical Physicist	5 (55.6)
Review Committee		
Gender	Male	5 (50.0)
	Female	5 (50.0)
Profession	Radiation Oncologist	3 (30.0)
	Medical Physicist	6 (60.0)
	RTT	1 (10.0)

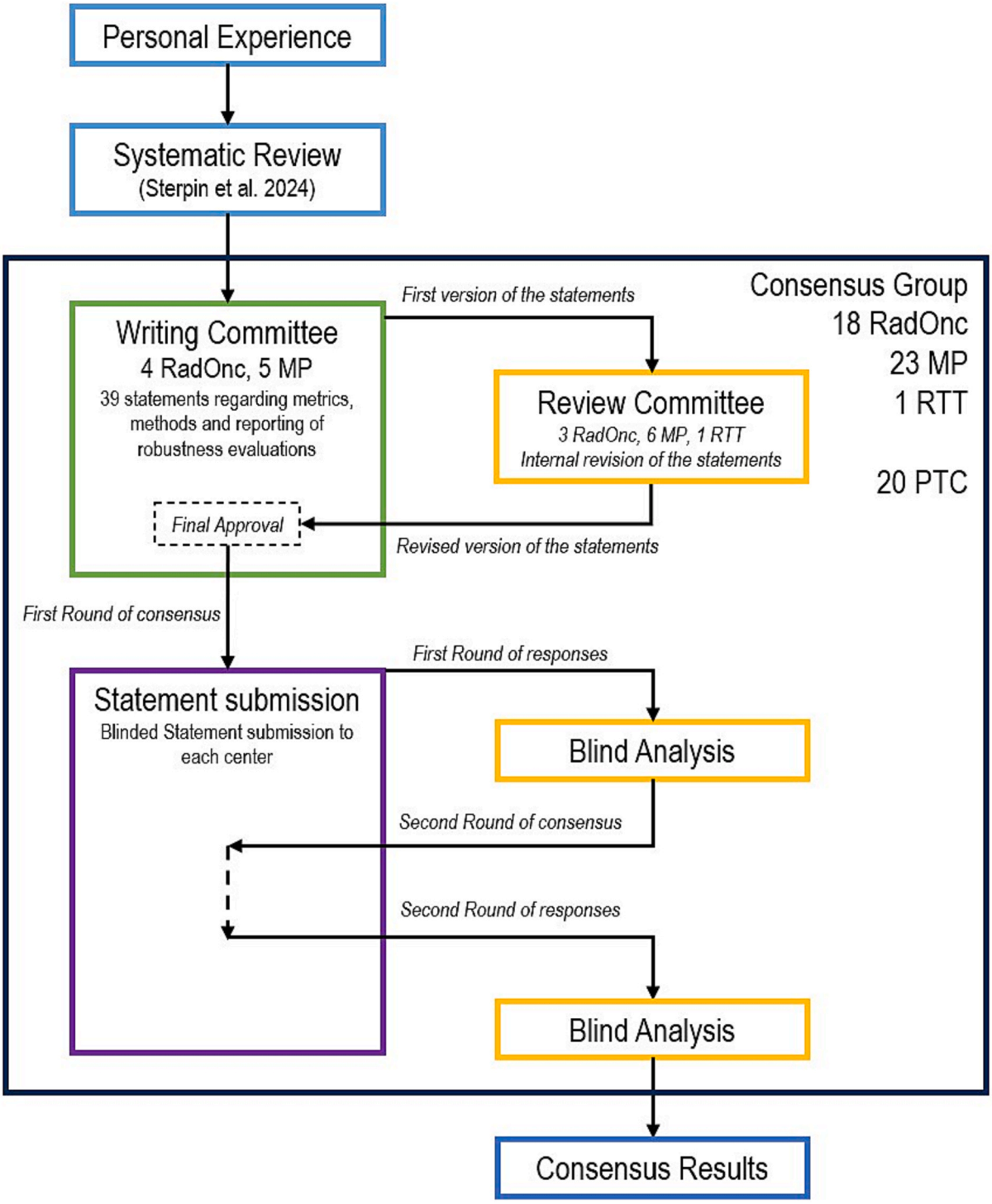

Fig. 1. Schematic representation of the delphi consensus process followed in this study. Abbreviations: RadOnc: Radiation Oncologist, MP: Medical Physicist, RTT: Radiation Therapist, PTC: Proton Therapy Centre. Two rounds of answers were needed to define the consensus.

Each round of consensus answers was analysed by a single person and the entire process was blind for every responder. No answer was influenced by other responders or by their comments. The writing committee defined a threshold of 75 % of agreement or disagreement to declare whether a statement met a consensus or not. If an agreement with high priority was met with more than 75 % a high agreement consensus was defined. For statements that reached the consensus threshold in the advanced statements group, the percentage of high-priority agreement was used to rank the new features recommended for future clinical implementation.

3. Results

3.1. Consensus process

Two rounds of answers were needed to finalize the consensus process. All the responding centres gave their answers at every round confirming the initial responding rate of 84.0 %. In total, for 26 out of 39 statements a consensus was found; 18 of them within the basic statements group (out of 29) and 8 in the advanced group (out of 10). For 5 statements in the basic group a high priority consensus was found (see Table 2).

3.2. Basic statements

In the first section of the Delphi process, regarding the transition from the PTV-based planning to a robust optimization approach on the Clinical Target Volume (CTV), the consensus highlights the limitations of the PTV concept for proton therapy. It proposed two methods to establish coverage equivalence between photon- and robustly optimized proton treatments, focusing on CTV coverage while accounting for both the voxel-wise minimum dose distribution and robustness error scenarios.

Specifically, there was a strong consensus obtained for the statement that underlined the need to go beyond the concept of PTV for the evaluation of proton plans. The other agreement of this section was on the equivalence between PTV-CTV coverage in photon treatments and CTV coverage in robustly optimized proton treatments that can be established by defining, for the CTV, an acceptance threshold of the passing rate of error scenarios of the coverage dose index. For the complete report, see the first statement's group in Table 3.

A high priority agreement was reached on three of the five statements concerning the types of uncertainties to be incorporated into robustness evaluation. A 100 % consensus was achieved on the necessity of considering geometric and range uncertainties, and over 90 % agreement was found for the statement that volumetric imaging should be evaluated during the treatment course for tumour sites susceptible to anatomical changes to assess plan robustness. Furthermore, there was agreement that evaluating the dosimetric impact of motion on the nominal plan is crucial for treating moving targets. For more details, refer to the second statement's group in Table 3.

A high priority agreement was reached for one of the seven statements concerning robustness evaluation methods (third statement's group in Table 3), specifying that for static targets, both setup and range systematic errors should be considered simultaneously using a worst-case approach. In five statements an agreement was reached. The statement that received over 95 % consensus indicated that once an in-house evaluation has been conducted to determine the range uncertainty value, this value should be incorporated into the clinically implemented robustness evaluation method.

The section on dose reporting revealed the only one disagreement in this study. The expert panel did not agree that, for proton plans, PTV dose index coverage from the nominal dose distribution should be used to report target volume coverage. On the remaining ten basic statements on reporting, a consensus agreement was reached five times (fourth statement's group Table 3). It was agreed that target volume coverage should be reported using the CTV dose index from the nominal dose distribution, and that the robustness of this coverage should be expressed as the percentage of scenarios meeting the dose prescription (passing rate) for the CTV coverage dose index. Furthermore, it was concurred that normal tissue doses, for both serial and parallel structures, should be reported using the nominal dose distribution and relevant indices, explicitly rejecting the use of planning risk volumes (PRV) for this purpose. For serial organs at risk (OAR), the consensus was to report doses based on the worst-case scenario dose distribution. Finally, the group agreed that dose constraints previously reported on PRVs should now be applied to the worst-case scenario in robustly optimized plans, ensuring that dose limits to normal tissues are maintained even under uncertainty.

3.3. Advanced statements

In the advanced group, a consensus was reached for 8 statements and none of them reached a high priority consensus. No consensus on disagreement with a statement was registered (see Table 4).

Within the robustness evaluation methodologies section, a clear consensus (90.0 %, statement 22) indicates that vendors should clinically implement probabilistic evaluation methods. These tools should account for both systematic and random errors and, optionally, consider delivery errors in multifractional treatment courses (80.0 %, statement 23).

Another important outcome of this section was given by the need for a clinical tool to assess the interplay effect of active scanning delivery time and patient motion for treating moving targets (90.0 % agreement, statement 12). This tool is considered essential for commissioning PBS for moving targets. This view was supported by three other statements (24, 25 and 26), which indicated that the Dynamic 4D Dose (D4DD) method should be used for dose estimation and the evaluation should be treatment site-specific (90.0 %, statement 26) or patient specific (80.0 %, statement 25). There is also strong agreement (85.0 %, statement 14) within the community to incorporate the evaluation of variable Relative Biological Effectiveness (RBE) model impacts on nominal dose distributions in clinical TPS, moving beyond the fixed 1.1 scaling factor. However, the consensus group emphasized the necessity for better validation of these models before clinical implementation, with suggestions including in-vivo validation and further investigation of physical quantities like Linear Energy Transfer (LET) or micro or nanodosimetry metrics.

A consensus was reached also for the need to have an a priori evaluation method of the impact of anatomy changes on nominal dose distributions where these variations are expected for example by generating synthetic CTs with foreseen anatomical changes.

No consensus was found for the possibility to evaluate the impact of anatomical variations together with setup and range uncertainties.

4. Discussion

This ESTRO Delphi Consensus study involved 42 experts from 20 european proton therapy centers with the aim of standardizing clinical methodologies for robustness evaluation and guiding future Treatment Planning System developments in particle therapy. The key findings include a strong consensus against using the PTV concept, advocating for CTV-based robust optimization, and emphasizing the necessity of consistently including both geometric and range uncertainties. High priority was placed on implementing probabilistic robustness evaluation methods and advanced imaging technologies to minimize range uncertainty.

Initially, a systematic review [3] was conducted to examine current clinical robustness assessment methods and emerging research technologies with potential for future clinical application. Subsequently, the focus shifted to achieving consensus among centres to establish a good practice baseline. This baseline is informed by both the systematic literature review and the clinical expertise of 18 RadOnc, 23 MP, and one RTT from 20 european particle therapy centres. These centres range from those with decades of clinical experience to those with professionals at the beginning of their clinical activity. The results of this ESTRO Delphi consensus highlight areas of strong agreement and divergence, providing valuable insights for clinical practice, guidelines and future technology development.

A significant finding was the strong consensus against using the PTV

Table 2
Final classification of statements for each group.

	Number of Statements per section High Priority Agreement	Agreement	Disagreement	Strong Disagreement	No Consensus
Basic Statements (29)	5	12	1	0	11
Advanced Statements (10)	0	8	0	–	2

Table 3
Basic statements group. Agreement, high priority agreement (in brackets), disagreement and neither agree nor disagree response rate are reported.

Statements Group Name	*ID*	*Statement definition*	*Agreement (High Priority Agreement)*	*Neither Agree nor Disagree*	*Disagreement*
From PTV to robust CTV optimization methods	1	PTV margin concept assumes the dose distribution invariance in presence of geometrical uncertainties. Dose distribution invariance mostly holds for photon therapy plans, as opposed to proton therapy plans.	**95.0 %** (**85.0 %**)	5.0 %	0.0 %
	2	Transitioning to robust optimization requires some form of cross-calibration against PTV based photon or proton planning techniques in terms of coverage requirements.	65.0 % (35.0 %)	15.0 %	20.0 %
	3	Equivalence between target coverage in PHOTON treatments (see premise above) and CTV coverage in robustly optimized proton treatments can be established by defining, in the voxel-wise minimum dose distribution, an acceptance threshold of the CTV coverage dose index.	73.7 % (26.3 %)	5.3 %	21.1 %
	4	Equivalence between PTV-CTV coverage in PHOTON treatments and CTV coverage in robustly optimized proton treatments can be established by defining, for the CTV, an acceptance threshold of the passing rate of error scenarios of the coverage dose index.	**84.2 %** (47.4 %)	10.5 %	5.3 %
	5	Equivalence between PTV-CTV coverage in PROTON treatments and CTV coverage in robustly optimized proton treatments can be established by defining, in the voxel-wise minimum dose distribution, an acceptance threshold of the CTV coverage dose index.	31.6 % (15.8 %)	5.3 %	63.2 %
	6	Equivalence between PTV-CTV coverage in PROTON treatments and CTV coverage in robustly optimized proton treatments can be established by defining, for the CTV, an acceptance threshold of the passing rate of error scenarios of the coverage dose index.	31.6 % (21.1 %)	15.8 %	52.6 %
Types of uncertainties	7	Geometrical uncertainty should be considered in robustness evaluation	**100.0 %** (**100.0 %**)	0.0 %	0.0 %
	8	Range uncertainty should be considered in robustness evaluation	**100.0 %** (**100.0 %**)	0.0 %	0.0 %
	9	Intra-fraction uncertainties: A per patient dosimetric evaluation of the impact of the motion, above a site-specific motion threshold, is a necessary condition to treat moving targets	**90.0 %** (70.0 %)	5.0 %	5.0 %
	10	For tumour sites prone to anatomical changes, volumetric imaging (repeated CTs, synthetic CTs from CBCT, CT-On-Rails etc.) should be evaluated during the treatment course to assess the plan robustness.	**100.0 %** (**95.0 %**)	0.0 %	0.0 %
	11	The dose averaged LET distribution should be considered in robustness evaluation.	60.0 % (0.0 %)	15.0 %	25.0 %
Methods	15	Range Uncertainty: once an in-house evaluation to determine the value of range uncertainty has been performed, it should be simulated within the robustness evaluation method clinically implemented/used.	**95.0 %** (65.0 %)	5.0 %	0.0 %
	16	Setup uncertainty: values for setup errors are derived as the CTV to PTV margins and are used accordingly within the robustness evaluation method clinically implemented/used.	**80.0 %** (70.0 %)	15.0 %	5.0 %
	17	Static lesions: at least set-up and range systematic errors should be considered simultaneously with a worst-case approach	**95.0 %** (**85.0 %**)	0.0 %	5.0 %
	18	Moving target: at least a recalculation of the entire plan on *extreme* phases of 4DCT should be evaluated to establish the robustness of the plan with respect to motion uncertainties.	**75.0 %** (65.0 %)	10.0 %	15.0 %
	19	Moving target: at least a recalculation of the entire plan on *selected* phases of 4DCT should be evaluated to establish the robustness of the plan with respect to motion uncertainties. If you agree or strongly agree, please state which phase you would select in the comments section.	60.0 % (45.0 %)	15.0 %	25.0 %
	20	Moving target: periodic 4DCT should be acquired during the treatment to assess if the plan is still robust with respect to the motion during the fraction.	**85.0 %** (40.0 %)	10.0 %	5.0 %
	21	Inter-fraction uncertainties: an 'a posteriori analysis' on repeated imaging during the treatment, a restricted cohort of patients should be used to assess the best planning approach for tumour sites with a high probability of anatomical changes during treatment.	**85.0 %** (60.0 %)	0.0 %	15.0 %
Reporting	29	PTV dose index coverage from nominal dose distribution should be used to report target volume coverage.	10.0 % (10.0 %)	0.0 %	**90.0 %**
	30	CTV dose index coverage from nominal dose distribution should be used to report target volume coverage.	**85.0 %** (65.0 %)	0.0 %	15.0 %
	31	Voxel-wise minimum dose distribution should be used to report CTV coverage dose index.	45.0 % (10.0 %)	10.0 %	45.0 %
	32	Worst case scenario dose distribution should be used to report CTV coverage dose index	70.0 % (30.0 %)	5.0 %	25.0 %
	33	Percentage of scenarios respecting the dose prescription (passing rate) should be used to report CTV coverage dose index	**80.0 %** (35.0 %)	0.0 %	20.0 %
	34	Nominal dose distribution and relevant dose indexes should be used to report organs (both serials and parallels) at risk dose (no PRV should be used to report nominal dose distributions).	**80.0 %** (60.0 %)	5.0 %	15.0 %
	35	When the near maximum dose for organs at risk is clinically relevant (serial OARs), the robustness evaluation should be performed on Voxel-wise max dose distribution.	40.0 % (20.0 %)	15.0 %	45.0 %

(*continued on next page*)

Table 3 (*continued*)

Statements Group Name	*ID*	*Statement definition*	*Agreement (High Priority Agreement)*	*Neither Agree nor Disagree*	*Disagreement*
	36	Worst case scenario dose distribution should be used to report serial organs at risk doses.	**80.0 %** (40.0 %)	0.0 %	20.0 %
	37	Worst case scenario dose distribution should be used to report parallel organs at risk doses.	35.0 % (15.0 %)	5.0 %	60.0 %
	38	Percentage of scenarios respecting the clinical goal (passing rate) should be used to report organs at risk (both parallels and serials) doses.	60.0 % (30.0 %)	5.0 %	35.0 %
	39	The dose to normal tissues that was reported on corresponding PRVs, should be reported on the worst-case scenario for robustly optimized plans	**80.0 %** (40.0 %)	0.0 %	20.0 %

Table 4
Advanced statements group. Agreement, high priority agreement (in brackets), disagreement and neither agree nor disagree response rate are reported. Results above 75% have been highlighted in bold.

Statements Group Name	*ID*	*Statements definition*	*Agreement (High Priority Agreement)*	*Neither agree nor disagree*	*Disagreement*
Type of uncertainties	12	Interplay effect: A dosimetric evaluation by means of the combination of active scanning delivery time structure and patient motion is a necessary prerequisite to treat moving targets above a certain site-specific motion threshold.	**90.0 %** (50.0 %)	5.0 %	5.0 %
	13	It is necessary to evaluate the impact of treatment delivery uncertainties (spot positioning, beam intensity etc.) in robustness evaluation by using machine log file or by using periodic machine QA's data.	65.0 % (30.0 %)	0.0 %	35.0 %
	14	It is important to have, in the near future, the possibility of evaluating the biological dose by considering variable RBE models in robustness evaluation.	**85.0 %** (40.0 %)	5.0 %	10.0 %
Methods	22	A probabilistic approach to evaluate the plan robustness should be implemented instead of a worst-case scenario approach.	**90.0 %** (60.0 %)	10.0 %	0.0 %
	23	It is necessary to simulate the impact of random and systematic set-up errors, geometrical/mechanical and systematic range error and, optionally, delivery errors (e.g. spot positioning, spot symmetry variations, output dose fluctuations etc.) in a multiple fraction treatment to obtain more realistic estimates of robustness.	**80.0 %** (35.0 %)	15.0 %	5.0 %
	24	Interplay effect: the Dynamic 4D Dose (D4DD) method should be used to evaluate the interplay effect by considering the motion (breathing, heart motion) pattern of the patient and the average delivery time parameters (average energy switching time, mean delivery time per MU, mean scanning speed etc.).	**90.0 %** (35.0 %)	5.0 %	5.0 %
	25	Moving target: the choice of 4D dosimetric evaluation method should be patient specific (i.e. by using the real breathing pattern of the patient together with planning 4DCT)	**80.0 %** (45.0 %)	15.0 %	5.0 %
	26	Moving target: the choice of 4D dosimetric evaluation method should be indication/tumour specific based on defined classification criteria	**90.0 %** (30.0 %)	5.0 %	5.0 %
	27	For clinical cases where anatomical changes are expected, these kinds of inter-fraction uncertainties should be evaluated as part of pre-treatment plan evaluation (i.e. by generating synthetic CTs with foreseen anatomical changes).	**75.0 %** (40.0 %)	10.0 %	15.0 %
	28	Anatomical changes: the impact of setup, range and anatomical changes uncertainty should be assessed together. (select disagree if you believe that anatomical changes should be evaluated separately with respect to setup and range uncertainties)	55.0 % (30.0 %)	10.0 %	35.0 %

concept for evaluating proton plans, acknowledging its limitations in accounting for the dose distribution changes due to geometrical, particle range, and anatomical uncertainties. This shift towards CTV-based robust optimization and evaluation, as well as considering error scenario passing rates, marks a necessary evolution in proton therapy [4,5]. On the other hand, voxel-wise minimum dose, in this consensus, has not been widely recognized as an instrument to commission robustly optimized plans against PTV based photon or protons plans.

The consensus formally established the necessity to include geometric and range uncertainties in robustness evaluations, and these errors should be considered simultaneously using a worst-case approach for static targets. This consensus agreed that the values of setup uncertainties that should be used, as a starting point, can be derived from the CTV to PTV margin used in the PTV based optimization (statement 16) [6–9]. It should be noted that those margins were defined in order to implicitly include part of the anatomy changes and contouring uncertainties. A concerted effort is needed to quantify those uncertainties, which remain a significant challenge in the radiotherapy community in order to explicitly include them in the optimization process or, at least, in the robustness evaluation. For moving targets, the importance of evaluating motion impact on the nominal dose was acknowledged and was considered as a necessary condition to start the treatment of moving lesions to ensure safe and effective treatments. The action levels to be used in these cases can vary depending on tumour site, modality, and institutional practice. With a general agreement recalculating plans on extreme or selected 4DCT phases was identified as a minimum requirement for these evaluations. For sites prone to anatomical changes, the importance of volumetric imaging during the treatment to assess plan robustness was highly agreed upon. The need for in-house determination of range uncertainty values for simulation was also strongly emphasized.

Regarding reporting, the clear rejection of PTV and PRV dose indices for reporting target volume and normal tissue doses further underscores the shift away from traditional margin-based therapy concepts. The consensus supports reporting CTV dose index from the nominal dose distribution, along with the percentage of calculated scenarios (for lack of any better method available) meeting the minimal required dose coverage dose prescription. We have to make explicit here that the percentage of scenarios meeting minimal dose coverage objective cannot be confused with a confidence level related to the absolute probability of meeting this minimal dose coverage, as would be the case in typical PTV margin recipes (for instance, 95 % minimal dose coverage of the GTV/CTV for 90 % of the patient population). This is because the scenarios are a limited set of all possible scenarios. For normal tissues, the nominal dose distribution and relevant indices were preferred, with worst-case scenario distributions suggested for serial OARs and dose

constraints. It should be noted that the voxel-wise minimum dose, despite a Dutch consensus exists [10], is not the recommendation of this consensus group for reporting target-related robustness evaluation results. Statement 31 of the consensus group revealed a disagreement on this approach, with answers split into two groups. This indicates that the field is still actively working to define the most appropriate standards for robust treatment planning evaluation and the suitability of any chosen metric for robustness assessments will only be validated by local control and normal tissue complication data.

A key issue that has found consensus is to consider in the analysis of the robustness of serial organs, the dose limits historically accepted for their PRV. If we consider, for example, the Danish guidelines for head and neck photon treatment [11] for the chiasm and optic nerves, the dose limit is $D_{max} < 54$ Gy for the organ and 60 Gy for their PRV. Following this example, to the consensus group it seemed reasonable to require for proton plans a surrogate dose index for maximum dose (i.e. $D_{0.03cc}$ as stated in [12]) at the organ of 54 Gy for the nominal plan and 60 Gy for all robustness evaluation scenarios to be respected. It is important to underline here how the dose limits to the normal tissues reported in QUANTEC [13] and subsequent papers [14], are always relative to the nominal dose indices. It could be too cautious for the normal tissue, with detriment to target coverage, to require that for all robustness evaluation scenarios of a proton plan, the maximum dose at the chiasm be below the nominal value reported in QUANTEC [15], considering that photon plans often have high dose gradients near serial organs, so that the normal tissue dose indices would likely have been much worse than the nominal had robustness evaluation been conducted. In this perspective, since robustness evaluation is used to make clinical decisions, it will be important to validate with prospective trials the correlation between robustness evaluation metrics and side effects.

In the advanced statements, a clear ranking of the upcoming innovations that should be implemented in clinical treatment planning software is provided. Among the proposed innovations, the implementation of probabilistic evaluation methods received the highest priority score for clinical integration into commercial software, as it enables a more realistic and reliable assessment of the impact of random and systematic uncertainties on dose distribution quality throughout the patient's treatment course. This topic has been widely studied using various clinical approaches [16–20], and the consensus clearly encourages vendors to consolidate these findings in order to standardize such methods in clinical practice. In particular, the possibility to compute confidence levels about the minimal dose coverage objective and organ at risk sparing is critical. As mentioned before, current methods of quantifying the number of scenarios passing a given objective/constraint do not give an accurate estimate of the confidence we have that this objective/constraint will be actually respected; because those scenarios are not necessarily equally likely, and because the scenario set only represents a limited percentage of all possible scenarios (some scenarios are a priori rejected because considered too unlikely; yet they should be counted in the estimation of the confidence). For a given dose objective/constraint, the associated global probability would be a welcome metric. In conclusion, the expert panel unanimously agrees that current robustness evaluation in particle therapy is insufficient, even if a crucial first step. Significant improvements are needed, highlighting a clear call to action for vendors and clinical centres to develop and implement advanced robustness evaluation tools, as the field has not yet reached its full potential.

The second high-scored priority topic was the need for tools to assess interplay effects in moving targets indicating a desire for more realistic robustness assessments against motion [21–25]. The group's agreement, in the basic statements, on the need to recalculate the entire plan on selected 4DCT phases to assess the effect of motion on the nominal dose distribution is not in contradiction with the need for a reliable interplay evaluation tool in commercial treatment planning systems. Such tools are essential for calculating D4DD [25–28], particularly when commissioning new techniques or introducing novel fractionation schemes, such as hypo fractionated lung treatments.

This consensus group agrees on the necessity of variable RBE models, indicating a desire to move beyond the limitations of a fixed RBE. The consensus process highlighted the importance of validating these models. While one centre reported using such models for plan evaluation, they are not yet used for clinical decision-making. This area presents a significant challenge for the future of proton therapy. Initial progress includes a European survey [29] revealing that European proton centres acknowledge RBE variability but adhere to current guidelines prescribing a constant RBE. To manage the uncertainty and potential for increased side effects due to RBE, centres employ measures like robust optimization and LET distribution evaluations during treatment planning, with LET being considered by a recent European survey as the recommended quantity beyond dose to be reported for proton plans [30]. A further advancement is the establishment of a national standard for evaluating and reporting LET and RBE-weighted dose in proton therapy for intracranial lesions [31].

While the primary focus of this consensus has been proton therapy, the principles discussed herein may also apply to other forms of particle therapies and serve as a useful guide. In this consensus we didn't include any statement regarding the effect of contouring uncertainty on plan robustness because this is an intrinsic source of uncertainty related to radiotherapy that, at the time being, cannot be simulated or integrated in a commercial TPS. We focused our consensus on physical and technical aspects of uncertainty quantification and robustness evaluation related to proton therapy. The contouring uncertainty should be considered in future works, potentially in a dedicated consensus exercise that focuses on clinical and workflow-related uncertainties considering the impact of instruments like AI-based contouring, multi-modality imaging, double-checking contours, and international guidelines to try to mitigate contour variability.

Ultimately, given the recognized limitations of the PTV and PRV concepts even in photon radiotherapy, especially with regard to superficial lesions (such as head and neck, or thoracic lesions), it is important to share these new tools throughout the radiotherapy community in order to achieve shared criteria for evaluating and comparing treatment plans. This work, focused on proton therapy, should therefore be seen as a first step towards achieving a broader consensus that also includes photon techniques.

This ESTRO Delphi consensus study offers crucial insights into robustness evaluation in proton therapy. A key conclusion is a strong consensus against using the PTV concept, instead advocating for CTV-based robust optimization and evaluation. It recommends consistently including geometric and range uncertainties, using a worst-case approach for static targets and 4DCT-based robustness evaluation for moving ones and a systematic evaluation of the impact of anatomy changes during the treatment.

There is a high priority consensus for probabilistic robustness evaluation methodologies and advanced imaging technologies aimed at reducing range uncertainty. Finally, the study highlights the need to move beyond a fixed RBE of 1.1 and emphasizes the importance of standardizing methodologies and aligning software developments with clinical needs. This lays the groundwork for more comprehensive, uncertainty-aware proton therapy and encourages further standardization and the guideline development in areas where consensus was not reached.

5. Disclaimer

ESTRO guidelines present scientific and medical opinions for educational and informational purposes only. Commercial use of any content in the guidelines without the prior written consent of ESTRO is strictly prohibited. The guidelines are not, and nor are they intended or implied to be, a substitute for professional medical advice or medical care. The advice of a medical professional should always be sought prior to commencing any form of medical treatment. Under no circumstances will ESTRO be held liable for any decision or action taken as a result of reliance on the content of the guidelines. ESTRO and all of its staff,

agents and members disclaim any and all warranties and representations (implied or otherwise) with regards to the information contained on the guidelines. Each guideline is based on information available at the time the corresponding writing panel conducted its research and discussions on the topic. Guidelines should neither be deemed inclusive of all proper methods of care or factors influencing the decision making, nor exclusive of other methods reasonably directed to obtaining the same results. There may be new developments since publication that are not reflected in any one guideline and that may, over time, form the basis for ESTRO to revisit and update the guideline.

CRediT authorship contribution statement

Francesco Fracchiolla: Conceptualization, Data curation, Formal analysis, Methodology, Project administration, Supervision, Visualization, Writing – original draft, Writing – review & editing. **Arturs Meijers:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Ester Orlandi:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Erik Korevaar:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Gillian Whitfield:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Kenneth Jensen:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Mischa Hoogeman:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Robin Wijsman:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Emmanuel Jouglar:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Eva Van Weerd:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Juan M. Pérez:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Ilaria Rinaldi:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Magdalena Garbacz-Stryszewska:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Marco Cianchetti:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Marta Montero Feijoo:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Silvia Molinelli:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Ulrik Vindelev Elstrøm:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Alessia Pica:** Writing – original draft, Writing – review & editing. **Andrew Gosling:** Writing – original draft, Writing – review & editing. **Beata Koczur:** Writing – original draft, Writing – review & editing. **Christina Vallhagen Dahlgren:** Writing – original draft, Writing – review & editing. **Daniela Alterio:** Writing – original draft, Writing – review & editing. **Einar Dale:** Writing – original draft, Writing – review & editing. **Goudjil Farid:** Writing – original draft, Writing – review & editing. **Inge Compter:** Writing – original draft, Writing – review & editing. **Jérôme Doyen:** Writing – original draft, Writing – review & editing. **Jon Espen Dale:** Writing – original draft, Writing – review & editing. **Johanna Färlin:** Writing – original draft, Writing – review & editing. **Konrad Urbanek:** Writing – original draft, Writing – review & editing. **Lars Fredrik Fjæra:** Writing – original draft, Writing – review & editing. **Maarten Lambrecht:** Writing – original draft, Writing – review & editing. **Maren Ugland:** Writing – original draft, Writing – review & editing. **Maria Tschiche:** Writing – original draft, Writing – review & editing. **Marie Vidal:** Writing – original draft, Writing – review & editing. **Matthew Lowe:** Writing – original draft, Writing – review & editing. **Rebecca Bütof:** Writing – original draft, Writing – review & editing. **Sarah Gulliford:** Writing – original draft, Writing – review & editing. **Stefania Comi:** Writing – original draft, Writing – review & editing. **Steven Habraken:** Writing – original draft, Writing – review & editing. **Thomas Tessonnier:** Writing – original draft, Writing – review & editing. **Edmond Sterpin:** Conceptualization, Methodology, Supervision, Writing – original draft, Writing – review & editing. **Lamberto Widesott:** Conceptualization, Data curation, Formal analysis, Methodology, Supervision, Visualization, Writing – original draft, Writing – review & editing.

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.

Acknowledgement

We acknowledge the comprehensive review of this recommendation by Marco Schwarz, Mirko Unipan and Ingrid Fagerström Kristensen whose advice and constructive comments were highly appreciated and helped to improve the quality of the manuscript.

References

[1] Hernandez V, Hansen C, 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. https://doi.org/10.1016/j.radonc.2020.09.038.

[2] Kaplan LP, Placidi L, Bäck A, Canters R, Hussein M, Vaniqui A, et al. Plan quality assessment in clinical practice: results of the 2020 ESTRO survey on plan complexity and robustness. Radiother Oncol 2022;173:254–61. https://doi.org/10.1016/j.radonc.2022.06.005.

[3] Sterpin E, Widesott L, Poels K, Hoogeman M, Korevaar E, Lowe M, et al. Robustness evaluation of pencil beam scanning proton therapy treatment planning: a systematic review. Radiother Oncol 2024;197:110365. https://doi.org/10.1016/j.radonc.2024.110365.

[4] 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.

[5] Fredriksson A. A characterization of robust radiation therapy treatment planning methods-from expected value to worst case optimization. Med Phys 2012;39:5169–81. https://doi.org/10.1118/1.4737113.

[6] Stroom JC, de Boer HC, 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.

[7] 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 2000;47:1121–35. https://doi.org/10.1016/S0360-3016(00)00518-6.

[8] 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.

[9] 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.

[10] Rojo-Santiago J, Habraken S, Lathouwers D, Méndez Romero A, Perkó Z, Hoogeman M. Accurate assessment of a dutch practical robustness evaluation protocol in clinical PT with pencil beam scanning for neurological tumors. Radiother Oncol 2021;163:121–7. https://doi.org/10.1016/j.radonc.2021.07.028.

[11] 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.

[12] Lambrecht M, Eekers DBP, Alapetite C, Burnet NG, Calugaru V, Coremans IEM, et al. Radiation dose constraints for organs at risk in neuro-oncology; the European particle therapy network consensus. Radiother Oncol 2018;128:26–36. https://doi.org/10.1016/j.radonc.2018.05.001.

[13] Bentzen S, Constine L, Deasy J, Eisbruch A, Jackson A, Marks L, 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(3 Suppl):S3–9. https://doi.org/10.1016/j.ijrobp.2009.09.040.

[14] Bisello S, Cilla S, Benini A, Cardano R, Nguyen NP, Deodato F, et al. Dose-volume Constraints fOr oRganS At risk In Radiotherapy (CORSAIR): an "all-in-one" multicenter-multidisciplinary practical summary. Curr Oncol Tor Ont 2022;29:7021–50. https://doi.org/10.3390/curroncol29100552.

[15] Mayo C, Martel MK, Marks LB, Flickinger J, Nam J, Kirkpatrick J. Radiation dose–volume effects of optic nerves and chiasm. Int J Radiat Oncol Biol Phys 2010;76:S28–35. https://doi.org/10.1016/j.ijrobp.2009.07.1753.

[16] 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. https://doi.org/10.1088/0031-9155/61/12/4646.

[17] Yang Z, Li H, Li Y, Li Y, Chang Y, Li Q, et al. Statistical evaluation of worst-case robust optimization intensity-modulated proton therapy plans using an exhaustive sampling approach. Radiat Oncol Lond Engl 2019;14:129. https://doi.org/10.1186/s13014-019-1335-8.

[18] Wahl N, Hennig P, Wieser HP, Bangert M. Efficiency of analytical and sampling-based uncertainty propagation in intensity-modulated proton therapy. Phys Med Biol 2017;62:5790–807. https://doi.org/10.1088/1361-6560/aa6ec5.
[19] Vazquez I, Gronberg MP, Zhang X, Court LE, Zhu XR, Frank SJ, et al. A deep learning-based approach for statistical robustness evaluation in proton therapy treatment planning: a feasibility study. Phys Med Biol 2023;68:095014. https://doi.org/10.1088/1361-6560/accc08.
[20] Fracchiolla F, Widesott L, Righetto R, Algranati C, Amelio D, Trianni A, et al. Development and clinical application of a probabilistic robustness evaluation tool for pencil beam scanning proton therapy treatments. Phys Med 2025;131:104938. https://doi.org/10.1016/j.ejmp.2025.104938.
[21] Taasti VT, Hattu D, Vaassen F, Canters R, Velders M, Mannens J, et al. Treatment planning and 4D robust evaluation strategy for proton therapy of lung tumors with large motion amplitude. Med Phys 2021;48:4425–37. https://doi.org/10.1002/mp.15067.
[22] Shan J, Yang Y, Schild SE, Daniels TB, Wong WW, Fatyga M, et al. Intensity-modulated proton therapy (IMPT) interplay effect evaluation of asymmetric breathing with simultaneous uncertainty considerations in patients with non-small cell lung cancer. Med Phys 2020;47:5428–40. https://doi.org/10.1002/mp.14491.
[23] Siregar H, Bäumer C, Blanck O, Chan M, Engwall E, Plaude S, et al. Mitigation of motion effects in pencil-beam scanning - impact of repainting on 4D robustly optimized proton treatment plans for hepatocellular carcinoma. Z Med Phys 2022;32:63–73. https://doi.org/10.1016/j.zemedi.2020.08.001.
[24] Batista V, Richter D, Chaudhri N, Naumann P, Herfarth K, Jäkel O. Significance of intra-fractional motion for pancreatic patients treated with charged particles. Radiat Oncol Lond Engl 2018;13:120. https://doi.org/10.1186/s13014-018-1060-8.
[25] Fan X, Wang S, Li W, Wang R, Yin Y, Dai T. Quantifying interplay effects in lung cancer IMPT: a comprehensive analysis of treatment planning parameter sensitivity. BMC Cancer 2025. https://doi.org/10.1186/s12885-025-15416-x.
[26] Meijers A, Knopf A-C, Crijns APG, Ubbels JF, Niezink AGH, Langendijk JA, et al. Evaluation of interplay and organ motion effects by means of 4D dose reconstruction and accumulation. Radiother Oncol 2020;150:268–74. https://doi.org/10.1016/j.radonc.2020.07.055.
[27] Inoue T, Widder J, van Dijk LV, Takegawa H, Koizumi M, Takashina M, et al. Limited impact of setup and range uncertainties, breathing motion, and interplay effects in robustly optimized intensity modulated proton therapy for stage III non-small cell lung cancer. Int J Radiat Oncol Biol Phys 2016;96:661–9. https://doi.org/10.1016/j.ijrobp.2016.06.2454.
[28] Rana S, Rosenfeld AB. Investigating volumetric repainting to mitigate interplay effect on 4D robustly optimized lung cancer plans in pencil beam scanning proton therapy. J Appl Clin Med Phys 2021;22:107–18. https://doi.org/10.1002/acm2.13183.
[29] Heuchel L, Hahn C, Pawelke J, Sørensen BS, Dosanjh M, Lühr A. Clinical use and future requirements of relative biological effectiveness: Survey among all European proton therapy centres. Radiother Oncol 2022;172:134–9. https://doi.org/10.1016/j.radonc.2022.05.015.
[30] Lühr A, Wagenaar D, Eekers DBP, Glimelius L, Habraken SJM, Harrabi S, et al. Recommendations for reporting and evaluating proton therapy beyond dose and constant relative biological effectiveness. Phys Imaging Radiat Oncol 2025;33:100692. https://doi.org/10.1016/j.phro.2024.100692.
[31] Wagenaar D, Habraken SJM, Rinaldi I, Eekers DBP, Kramer M, Jaspers JPM, et al. Evaluating and reporting LET and RBE-weighted dose in proton therapy for glioma - the dutch approach. Radiother Oncol 2025;202:110653. https://doi.org/10.1016/j.radonc.2024.110653.

Francesco Fracchiolla[a,*], Arturs Meijers[b], Ester Orlandi[c,d], Erik Korevaar[e], Gillian Whitfield[f], Kenneth Jensen[g], Mischa Hoogeman[h,i], Robin Wijsman[e], Emmanuel Jouglar[j], Eva Van Weerd[i], Juan M. Pérez[k], Ilaria Rinaldi[l], Magdalena Garbacz-Stryszewska[m], Marco Cianchetti[n], Marta Montero Feijoo[k], Silvia Molinelli[o], Ulrik Vindelev Elstrøm[g], Alessia Pica[b], Andrew Gosling[p], Beata Koczur[q], Christina Vallhagen Dahlgren[r], Daniela Alterio[s], Einar Dale[t], Goudjil Farid[j], Inge Compter[l], Jérôme Doyen[u], Jon Espen Dale[v], Johanna Färlin[w], Konrad Urbanek[x], Lars Fredrik Fjæra[y], Maarten Lambrecht[z], Maren Ugland[v], Maria Tschiche[aa], Marie Vidal[ab], Matthew Lowe[ac], Rebecca Bütof[aa,ad], Sarah Gulliford[p,ae], Stefania Comi[af], Steven Habraken[i,ag], Thomas Tessonnier[q], Edmond Sterpin[ah,ai,aj], Lamberto Widesott[a]

[a] *UO Fisica Sanitaria, APSS, Trento, Italy*
[b] *Center for Proton Therapy, Paul Scherrer Institut, Forschungsstrasse 111, Villigen 5232, Switzerland*
[c] *Department of Clinical, Surgical, Diagnostic and Pediatric Sciences, University of Pavia, Italy*
[d] *Radiation Oncology Unit, Clinical Department, CNAO National Center for Oncological Hadrontherapy, Italy*
[e] *Department of Radiation Oncology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands*
[f] *Department of Radiotherapy, The Christie NHS Foundation Trust, Manchester, UK*
[g] *Danish Centre for Particle Therapy, Aarhus University Hospital, Aarhus, Denmark*
[h] *Erasmus MC Cancer Institute, University Medical Center Rotterdam, Department of Radiotherapy, the Netherlands*
[i] *Holland Proton Therapy Center, Delft, the Netherlands*
[j] *Institut Curie, PSL Research University, Department of Radiation Oncology, Orsay Protontherapy Center, F-91400 Orsay, France*
[k] *Centro de Protonterapia Quironsalud, Madrid, Spain*
[l] *Department of Radiation Oncology (Maastro), GROW School for Oncology, Maastricht University Medical Centre+, Maastricht, the Netherlands*
[m] *Cyclotron Centre Bronowice, Institute of Nuclear Physics Polish Academy of Sciences, ul. Radzikowskiego 152, 31-342 Krakow, Poland*
[n] *UO Protonterapia, APSS, Trento, Italy*
[o] *Medical Physics Unit, Clinical Department, CNAO National Center for Oncological Hadrontherapy, Italy*
[p] *Department of Radiotherapy Physics, University College London Hospitals NHS Foundation Trust, London, UK*
[q] *Heidelberg Ion-Beam Therapy Center (HIT), Department of Radiation Oncology, Heidelberg University Hospital, Im Neuenheimer Feld 450, 69120 Heidelberg, Germany*
[r] *The Skandion Clinic, Uppsala, Sweden*
[s] *Division of Radiotherapy, IEO European Institute of Oncology, IRCCS, Milan, Italy*
[t] *Department of Oncology, Oslo University Hospital, Oslo, Norway*
[u] *Department of Radiation Therapy, Antoine Lacassagne Cancer Center, University of Côte d'Azur, Nice, France*
[v] *Cancer Clinic, Haukeland University Hospital Bergen, Norway*
[w] *Department of Diagnostics and Intervention, Oncology, Umeå University, Umeå, Sweden*
[x] *Department of Radiotherapy, Maria Skłodowska-Curie National Research Institute of Oncology, Kraków Branch, Kraków, Poland*
[y] *Department of Medical Physics, Oslo University Hospital, Oslo, Norway*
[z] *Department of Radiotherapy-Oncology, Leuvens Kanker Instituut, Universitair Ziekenhuis Leuven, Leuven, Belgium*
[aa] *Department of Radiotherapy and Radiation Oncology, Faculty of Medicine and University Hospital Carl Gustav Carus, TUD Dresden University of Technology, Dresden, Germany*
[ab] *Institut Méditerranéen de Protonthérapie - Centre Antoine Lacassagne, Nice, France*
[ac] *Christie Medical Physics and Engineering, The Christie NHS Foundation Trust, Manchester, UK*
[ad] *OncoRay – National Center for Radiation Research in Oncology, Faculty of Medicine, Dresden, Germany*
[ae] *Department of Medical Physics and Biomedical Engineering, University College London, UK*
[af] *Unit of Medical Physics, European Institute of Oncology (IEO) IRCCS, Milano, Italy*
[ag] *Leiden University Medical Center, Department of Radiotherapy, Leiden, the Netherlands*
[ah] *KU Leuven, Department of Oncology, Experimental Radiotherapy Lab, Leuven, Belgium*
[ai] *UCLouvain, Institution de Recherche Expérimentale et Clinique, Center of Molecular Imaging Radiotherapy and Oncology (MIRO), Brussels, Belgium*
[aj] *Particle Therapy Interuniversity Center Leuven – PARTICLE, Leuven, Belgium*

[*] Corresponding author at: Centro di Protonterapia, APSS, Via Al Desert 14, 38122, Trento, Italy.
E-mail address: francesco.fracchiolla@apss.tn.it (F. Fracchiolla).