

# Modelling the risk of radiation induced alopecia in brain tumor patients treated with scanned proton beams

Giuseppe Palma<sup>1</sup>, PhD, Alberto Taffelli<sup>2</sup>, MSc, Francesco Fellin<sup>3</sup>, Vittoria D'Avino<sup>1</sup>, PhD, Daniele Scartoni<sup>3</sup>, MD, Francesco Tommasino<sup>2,4</sup>, PhD, Emanuele Scifoni<sup>2,5</sup>, PhD, Marco Durante<sup>5,6</sup>, PhD, Maurizio Amichetti<sup>3</sup>, MD, Marco Schwarz<sup>2,3</sup>, PhD, Dante Amelio<sup>3,\*</sup>, MD, Laura Cella<sup>1,\*</sup>, PhD

<sup>1</sup>National Research Council, Institute of Biostructures and Bioimaging, Napoli, Italy

<sup>2</sup>Istituto Nazionale di Fisica Nucleare, Trento Institute for Fundamental Physics and Applications, Trento, Italy

<sup>3</sup>Trento Proton Therapy Center, Azienda Provinciale per I Servizi Sanitari (APSS), Trento, Italy,

<sup>4</sup>University of Trento, Physics Department, Trento, Italy

<sup>5</sup>GSI Helmholtzzentrum für Schwerionenforschung, Biophysics Department, Planckstraße 1, 64291 Darmstadt, Germany

<sup>6</sup>Technische Universität Darmstadt, Institut für Festkörperphysik, Hochschulstraße 6, 64289 Darmstadt, Germany

**Corresponding author:** Laura Cella, National Research Council, Institute of Biostructures and Bioimaging, Via T. De Amicis, 95; 80145 – Napoli, Italy

E-mail address: [laura.cella@cnr.it](mailto:laura.cella@cnr.it)

\* These authors share the senior authorship.

**Short Title:** NTCP for radiation-induced alopecia after proton therapy

**Keywords:** Radiation-induced alopecia, Proton therapy, Brain tumors, NTCP

**Source of Financial Support/Funding Statement:** The research was funded in part by the Italian National Institute for Nuclear Physics (INFN) CSN5 Call “MoVe-IT”.

**Conflict of Interest Notification:** The authors declare no conflict of interest.

## Abstract

### Purpose

To develop normal tissue complication probability (NTCP) models for radiation-induced alopecia (RIA) in brain tumor patients treated with proton therapy (PT).

### Methods and materials

We analyzed 116 brain tumor adult patients undergoing scanning beam PT (median dose 54 GyRBE; range 36-72) for CTCAE v.4 grade 2 (G2) acute ( $\leq 90$  days), late ( $>90$  days) and permanent ( $>12$  months) RIA.

The relative dose-surface histogram (DSH) of the scalp was extracted and used for Lyman-Kutcher-Burman (LKB) modelling. Moreover, DSH metrics ( $S_x$ : the surface receiving  $\geq X$  Gy,  $D_{2\%}$ : near maximum dose,  $D_{\text{mean}}$ : mean dose) and non-dosimetric variables were included in a multivariable logistic regression NTCP model. Model performances were evaluated by the cross-validated area under the receiver operator curve (ROC-AUC).

### Results

Acute, late and permanent G2-RIA was observed in 52%, 35% and 19% of the patients, respectively. The LKB models showed a weak dose-surface effect ( $0.09 \leq n \leq 0.19$ ) with relative steepness  $0.29 \leq m \leq 0.56$ , and increasing tolerance dose values when moving from acute and late (22 and 24 GyRBE) to permanent RIA (44 GyRBE).

Multivariable modelling selected  $S_{21\text{Gy}}$  for acute and  $S_{25\text{Gy}}$  for late G2-RIA as the most predictive DSH factors. Younger age was selected as risk factor for acute G2-RIA while surgery as risk factor for late G2-RIA.  $D_{2\%}$  was the only variable selected for permanent G2-RIA. Both LKB and logistic models exhibited high predictive performances (ROC-AUCs range 0.86-0.90).

### Conclusion

We derived NTCP models to predict G2-RIA after PT, providing a comprehensive modelling framework for acute, late and permanent occurrences that, once externally validated, could be exploited for individualized scalp sparing treatment planning strategies in brain tumor patients.

## Introduction

Alopecia is one of the most common and emotionally troublesome effects of cancer therapy [1]. Radiation induced alopecia (RIA) occurs less frequently than chemotherapy induced alopecia. However, the former is more likely to result in permanent hair loss [2]. Temporary or permanent RIA are one among the side effects resulting from brain tumor radiation therapy (RT); and, while its importance in affecting long-term quality of life (QoL) has been recognized, it has been accepted as an unavoidable consequence [3]. Only a few studies report on dose response analysis for RIA following photon or proton cranial irradiation [3-6]. Lawenda *et al.* [4] first reported a dose-response relationship for moderate to severe permanent alopecia after photon cranial irradiation in a limited cohort of 26 patients, and found a 50% probability to induce toxicity (TD50) value of 43 Gy.

Proton therapy (PT) is considered a promising RT treatment modality for brain tumors [7]. Indeed, the physical properties of protons render PT potentially ideal for brain tumors requiring prescription doses that exceed tolerance dose levels in surrounding organs at risk. To mitigate radiation-induced adverse effects, protons may also be preferred for brain tumors associated with long-term survival [7, 8]. However, the higher entry dose of the spread-out Bragg peak may represent a disadvantage; thus causing concern over a possible increase in skin adverse effects, such as hair loss. This effect is particularly important in passive scattering PT, but may have a relevant role also in active scanning technique [6, 9] despite this technique offers larger treatment planning optimization flexibility.

A study on a small cohort of 12 pediatric medulloblastoma patients treated with passive scattered PT [5] suggested that permanent RIA was correlated with a dose threshold of about 21-30 Gy. Another very recent study on adult brain tumor patients treated by PT [6] developed and externally validated a Normal Tissue Complication Probability (NTCP) model only for acute RIA. The authors reported that the best predictive variables were skin near maximum dose ( $D_{2\%}$ ) and skin volume receiving at least 25 Gy ( $V_{25Gy}$ ).

Against this background, NTCP models represent the key to maximizing the benefits of technological advances in RT such as PT [10]. Models able to robustly predict RIA, in particular the permanent one, may play an essential role in identifying of an individualized optimal plan to minimize the risk of RIA, thus improving brain tumor patients' QoL.

Aim of the present study was to develop NTCP models for acute, late and permanent RIA in patients treated with scanning beam PT for brain tumors. To the best of our knowledge, this is the first study to model permanent RIA in quite a large cohort of brain tumor patients treated with PT. The modelling procedure was based on the introduction of a novel fully automated procedure for scalp definition as a critical organ for RIA. Two different NTCP modelling strategies were adopted, one based on the pure dosimetric Lyman-Kutcher-Burman (LKB) analysis and the other based on multivariable logistic regression.

## **Methods and materials**

### *Patients and treatments*

Clinical and dosimetric records of a cohort of consecutive adult patients treated for brain tumors (local Ethics Committee approval protocol A489) were retrospectively reviewed. All patients have been evaluated for alopecia before RT (baseline) and monitored weekly for RIA injury during RT, at the end of RT, and every 3 months after RT, as part of clinical routine. For each patient, RIA was graded in consensus by two radiation oncologists (DA and DS) according to the National Cancer Institute's Common Terminology Criteria for Adverse Events version 4 (CTCAE v4) for acute ( $\leq 90$  days after RT completion), late ( $> 90$  days after RT completion) and permanent (persisting for  $> 12$  months after RT completion [4]) injury. For each RIA outcome, the maximum CTCAE v4 severity score was recorded during the corresponding period of time. Time to RIA was computed from the end of RT to

the first RIA event. If an event occurred during the RT treatment, time to RIA was set to 0. Patients were followed up for a median time of 235 days (range 0-923 days).

In order to limit intra- and inter-observer variability, the decision on scoring was accomplished consensually by radiation oncologists with the analysis of pictures (Supplementary figure 1) acquired at each time point. All participants gave their written informed consent. Patient data were analyzed after anonymization. Patient and treatment characteristics are described in Table 1.

All patients underwent PT by pencil beam scanning at Trento Proton Therapy Center with standard fractions (1.8 or 2 GyRBE) between October 2014 and November 2017. The study eligibility criteria include evaluation of RIA status and availability of 3-dimensional dose maps. The resulting cohort consisted of 116 patients. For all the patients in the resulting cohort, the planning-CT matrix size was 512x512 in plane, with a CT slice thickness of 1.5 mm. Treatment plans were designed using a single field optimization technique based on pencil beam scanning. According to the moment in time when the patient was treated, the plan was designed either in XiO (Elekta AB, Stockholm, Sweden) or in RayStation (Ray-Search AB, Stockholm, Sweden). The dose grid size ranged between 1.0x1.0x1.0 mm<sup>3</sup> and 1.5x1.5x1.5 mm<sup>3</sup>.

The relative biological effectiveness (RBE) was considered to be constant and set to 1.1, *i.e.* D (GyRBE) = 1.1 D (Gy) [11]. ICRU is currently revising the dose units in proton therapy to avoid the use of different units [12]. In this paper we will refer only to the Gy unit for proton dose values although they include the RBE corrections. Physical doses are 10% lower than those reported in the paper.

Target definition was based on CT and on morphological magnetic resonance (MR) imaging with contrast enhancement medium. When required, information coming from positron-emission tomography (PET) were also employed. Macroscopic residual tumor did represent the gross tumor volume (resection cavity was included for patients with surgery prior RT). Depending on the tumor

type, variable margins were added to take into account microscopic tumor infiltration (clinical target volume - CTV). Planning target volume was generated by adding a 3-4 mm margin to CTV.

### *Dosimetric analysis*

For each patient, individual DICOM RT plans (computed tomography (CT) scans, doses and contoured organ structures) were converted into Matlab-readable format (MathWorks, Natick, MA, USA) using the CERR (Computational Environment for Radiotherapy Research) software [13].

Healthy hair roots are found in the fat tissue (connective tissue) of the second scalp layer, on average 4 mm under the skin surface [14, 15]. In order to accurately evaluate the dose to scalp hair follicles and to take into account the scalp anatomy, the scalp structure was automatically extracted as follows. First, a gross estimate of the connective tissue in the planning CT field-of-view was derived as:

$$\text{RawConnectiveTissue} = [(\text{Body} \ominus B[2 \text{ mm}]) \setminus (\text{Body} \ominus B[6 \text{ mm}])]$$

where  $B[r]$  is a spherical structuring element of radius  $r$ ,  $\setminus$  represents the set difference, and  $\oplus$  and  $\ominus$  stand for morphological dilation and erosion [16], respectively. In order to select just the portion that covers the cranial vault, we then defined a gross scalp structure as:

$$\text{RawScalp} = \text{RawConnectiveTissue} \cap (\text{Brain} \oplus B[25 \text{ mm}])$$

Finally, the fine estimate of the Scalp structure was obtained by excluding from RawScalp possible regions of bone tissue (radiometrically identified by a lower threshold set to 100 Hounsfield Unit (HU) on planning CT scan):

$$\text{Scalp} = \text{RawScalp} \cap \overline{[(\text{HU} > 100) \cdot B[1 \text{ mm}]}$$

where  $\cdot$  represents the morphological closing [16].

Using an in-house developed library for Matlab [17, 18], we then computed the dose-surface histograms (DSHs) of the scalp thus obtained. The following DSH metrics were extracted: the relative scalp surface receiving  $\geq X$  Gy ( $S_x$ ) in step of 1 Gy, the minimum dose given to the hottest  $x\%$  scalp surface in step of 5% ( $D_x$ ), the near maximum dose ( $D_{2\%}$ ) and the mean dose ( $D_{\text{mean}}$ ) to the scalp surface.

### *Statistical analysis*

The analysis was performed considering 3 different endpoints: acute, late and permanent G2 RIA status (a G2 RIA event was defined as the occurrence of a RIA of grade 2). All the extracted scalp dose parameters along with patient and treatment-related factors (age at RT, prescription dose, number of fractions, CTV size, scalp surface, gender, surgery, chemotherapy, re-irradiation, diagnosis) were analyzed for the above described groupings by univariate statistical methods: categorical variables were tested by Pearson's  $\chi^2$ -test or Fisher's exact test when appropriate; continuous variables were tested by Mann-Whitney  $U$ -test. Bonferroni-corrected  $\alpha$ -level of 0.05 was used as significance level. Receiver operator characteristic (ROC) curve analysis was performed and cut-off values on the ROC curve were determined by Youden's J statistic [19].

### *Normal tissue complication probability modelling*

For each endpoint, two different NTCP models were analyzed: the LKB model built on generalized equivalent uniform dose (gEUD) [20, 21] and recast for DSHs [18] as well as the multivariable logistic model. The LKB model was fitted as previously described in [22]. The Maximum Likelihood method was used to find the best-fit values of the LKB parameters (TD50,  $m$  and  $n$ ) by maximizing the logarithm of the likelihood (LLH). TD50 is the value of the uniform dose given to the entire organ

surface corresponding to the 50% probability to induce toxicity;  $m$  is inversely proportional to the slope of the dose-response curve; and  $n$  accounts, in this specific case, for surface effects. The LLH function was numerically maximized using an in-house developed library for Matlab. 95% confidence intervals for parameters estimates were obtained using the profile likelihood method.

In order to evaluate the impact of dosimetric and non-dosimetric factors, multivariable stepwise logistic regression methods for NTCP modelling were also applied [23-25]. For each analyzed outcome, only the variables most highly correlated with RIA ( $p < 0.1$ ) at the univariate analysis were included in the subsequent multivariable analysis. Highly inter-correlated variables (correlation  $|R_s| > 0.75$ ) were further removed keeping only the variables with the highest correlation with RIA to avoid collinearity problems.

The Leave-One-Out (LOO) method was applied to the whole statistical pipeline to cross validate the model performance. Model performance was evaluated by the area under the ROC curve (AUC) and by the Brier Score (BS). Calibration plots were generated for graphical assessment of the agreement between observed outcomes and LOO predictions.

A hierarchical regression approach [26]) was also applied in order to take into account a potential role of RIA history in the development of late/permanent effects and, thus, the consequential component of RIA [27]. Accordingly, acute RIA status was provided as second level for late/permanent RIA regression and late RIA status as third level for permanent RIA regression.

## Results

The incidence of RIA is reported in Table 1. Data for acute RIA were available for 116 out of 117 initially included patients. Data for late and permanent RIA were available for 103 patients out of 116. Acute G2 RIA was found in 60 of the 116 (52%) patients at a median time of 35 days (range: 0-58 days) from the end of PT. Late G2 RIA was found in 41 of the 103 patients (35%) at a median time of

113 days (range: 91-671 days) from PT. Twenty of the 103 (19.4%) patients developed permanent G2 alopecia. Non-dosimetric variables, significantly associated with G2 RIA status, were reported in Table 1. Strong correlations between temporary and permanent G2 RIA statuses were identified.

For all the analyzed RIA status, the average DVHs of those patients who developed alopecia were separated from the average DVHs of unaffected patients well beyond the associated standard errors on a large range of doses (Figs. 1-3).

The training of the LKB models achieved high prediction performances ( $AUC_{acuteRIA} = 0.89$ ;  $AUC_{lateRIA} = 0.87$ ;  $AUC_{permanentRIA} = 0.88$ ; see Table 2, Figs. 1-3) and highlighted increasing TD50 values (acute G2 RIA TD50 = 22 Gy, late G2 RIA TD50 = 24 Gy, permanent G2 RIA TD50 = 44 Gy) when moving from temporary (acute or late) to permanent G2 RIA.

Multivariable logistic modelling of both temporary G2 RIA outcomes led to the selection of one dosimetric variable ( $S_{21Gy}$  and  $S_{25Gy}$  for acute and late G2 RIA, respectively) and of one non-dosimetric variable (age and surgery for acute and late G2 RIA, respectively) as predictive factors (Table 3). Conversely, the only variable selected for permanent G2 RIA modelling was the near maximum scalp dose  $D_{2\%}$ . ROC analysis of acute G2 RIA identified the value of 5.1% as optimal cut-off for  $S_{21Gy}$ ; for late G2 RIA the cut-off for  $S_{25Gy}$  was 2.0%, and for permanent effect the cut-off for  $D_{2\%}$  was 48 Gy.

As for the LKB models, all logistic models provided high performances measured in terms of AUC ( $AUC_{acute} = 0.91$ ;  $AUC_{late} = 0.89$ ;  $AUC_{permanent} = 0.89$ ), BS and calibration (Table 3, Figs. 1-3).

The hierarchical regression approach identified a further significant improvement of prediction performance when acute G2 RIA was included in the stepwise logistic model of late G2 RIA ( $p=0.015$ ) and when late G2 RIA was included in the stepwise logistic model of permanent effect ( $p<0.001$ ).

## Discussion

In the present study we modeled temporary and permanent G2 RIA in a cohort of brain cancer patients treated by active scanning PT. Two different approaches were applied: a purely dosimetric LKB NTCP and a multivariable logistic regression model including both dosimetric and clinical covariates. To this end, we devised a new algorithm for the automated extraction of the scalp, based on the patient outline (i.e. body) and the brain structure which are commonly contoured in the clinical practice of brain tumor RT.

Limited literature reported on dose–response relationship for the occurrence of temporary or permanent alopecia after cranial irradiation and analyzed potential confounding variables that might contribute to predict this often unavoidable side effect of RT [4].

In the present study, we proposed the first comprehensive modelling analysis of both temporary (acute and late) and permanent G2 RIA on a large cohort of brain cancer patients. The obtained results support and expand the previously reported indications or findings.

From the average DSHs, a clear separation can be observed between patients who developed G2 RIA and those who did not (Figs. 1-3), throughout a large range of scalp doses. Therefore, we first adopted the LKB approach to toxicity modelling which could improve the robustness of the prediction by exploiting the average gEUD of different dose level parameters. The training of the 3 LKB models resulted in a very good prediction performance and did not reveal any effect of “calibration-in-the-large” while slope and intercept of cross validation calibration plots are close to ideal [28]. In particular, this behavior is confirmed even for acute and late G2 RIA, despite the scattered appearance of points around the best fit of the calibration curve.

Passing from acute to permanent alopecia, the consistent increase of TD50 values from 22 Gy to 44 Gy appears clinically sound and also confirms previous findings on dose value thresholds [4-6]. The values for the surface parameter  $n$  suggest a weak dose surface effect. Moreover, the estimation of

the  $n$  parameter can be exploited for RT plan optimization by constraining the gEUD, the most common empirical model implemented in several treatment planning systems.

Several clinical factors turned out to be strongly correlated with the analyzed outcomes. We therefore investigated their potential impact on the G2 RIA predictive power by a multivariable logistic regression approach in view of a personalized treatment strategy. In addition to DSH metrics, younger age at time of RT was selected as a risk factor for acute G2 RIA while surgery was found to be a risk factor for late RIA. There is conflicting evidence concerning the effect of age on the prevalence and severity of radiation induced skin toxicity [29, 30]. Older age in adults is usually associated with a higher normal tissue radiosensitivity due to the depletion of tissue stem cells, to age-related accumulation of mutations and to a subsequent decline in DNA repair capacity [31]. However, a correlation between age and acute RIA was not reported in previous literature, except for a trend reported in [6]. Further studies are needed to clarify the influence of age on acute skin toxicity and, in particular, on acute RIA.

On the contrary, clinical variables did not significantly add prediction power to the logistic model for permanent G2 RIA which was based only on the near maximum scalp dose.

It is worth noting that hierarchical regressions highlighted the presence of a consequential component for the scalp in the development of late and permanent alopecia, thus confirming that the improvement of an early response to irradiation (through the proper NTCP models) may be a useful approach to minimize later side effects [27].

The risks of a potential increase of radiation induced skin or scalp toxicity has long been considered a peculiar drawback in the clinical use of protons and in particular for its early conventional dose delivery method, using passive scattering PT. Indeed, while in photon RT, the skin sparing is obtained thanks to the initial dose buildup effect; this is not the case for the directly and locally ionizing features of the particle beam radiation used in hadron therapy, where the skin lies in the dose profile

plateau region. The latter plateau, in an extended spread out Bragg peak, may become considerably high. However, proton pencil beam active scanning, through energy and spot selection features, offers higher flexibility and enhanced modulation capability in treatment planning [32]. Therefore, if the scalp is included in the optimization strategy, a reduction in the dose released to the skin and, thus, a potential reduction in the incidence rates of alopecia can be obtained.

The main limitation of the present study is that the derived models could not be validated on an independent cohort from an external institution, thus hindering their straightforward inclusion in the clinical practice. Nonetheless, the reported findings were obtained by a fully automated implementation of both the extraction of the scalp structure and the following statistical processing, thus strengthening the performed analysis and easing the process of unbiased validation on external cohorts of patients. In addition, we did not take into account the possible impact of a variable RBE for protons which was set to the widely accepted constant value of 1.1 [33, 34]. Nonetheless, our result on proton TD50 value for permanent alopecia is in agreement with the TD50 of 43 Gy derived for photon cranial irradiation and reported in [4].

In conclusion, following different modelling approaches, we derived six NTCP models describing the different types of G2 RIA after active scanning PT, for which prior literature failed to offer a comprehensive modelling framework of both temporary and permanent occurrences. From LKB modelling, the TD50 values confirms previous findings on dose value thresholds obtained after photon cranial irradiation while the surface parameter  $n$  suggests a weak dose surface effect for the scalp. All the above findings provide a coherent picture of the analyzed morbidities and, once externally validated, could be easily exploited for individualized treatment planning and for scalp sparing RT in brain tumor patients. Active scanning PT offers indeed a large flexibility for the optimization and the inclusion of the scalp-specific model parameters in the planning strategy may

minimize the occurrence of radiation induced hair-loss, and in particular of permanent one with a consequent improvement of patient's quality of life.

## References

- [1] Dorr VJ. A practitioner's guide to cancer-related alopecia. *Semin Oncol*. 1998;25:562-70.
- [2] Rubin P, Michael T, Milano and Louis S. Constine. *ALERT Adverse Late Effects of Cancer Treatment: Volume 2: Normal Tissue Specific Sites and Systems*. 2013:189.
- [3] Mahadevan A, Sampson C, LaRosa S, et al. Dosimetric analysis of the alopecia preventing effect of hippocampus sparing whole brain radiation therapy. *Radiat Oncol*. 2015;10:245.
- [4] Lawenda BD, Gagne HM, Gierga DP, et al. Permanent alopecia after cranial irradiation: dose-response relationship. *Int J Radiat Oncol Biol Phys*. 2004;60:879-87.
- [5] Min CH, Paganetti H, Winey BA, et al. Evaluation of permanent alopecia in pediatric medulloblastoma patients treated with proton radiation. *Radiat Oncol*. 2014;9:220.
- [6] Dutz A, Luhr A, Agolli L, et al. Development and validation of NTCP models for acute side-effects resulting from proton beam therapy of brain tumours. *Radiother Oncol*. 2019;130:164-71.
- [7] Ahmed SK, Brown PD, Foote RL. Protons vs Photons for Brain and Skull Base Tumors. *Semin Radiat Oncol*. 2018;28:97-107.
- [8] Indelicato DJ, Rotondo RL, Uezono H, et al. Outcomes Following Proton Therapy for Pediatric Low-Grade Glioma. *Int J Radiat Oncol Biol Phys*. 2019, 104:149-156.
- [9] Hauswald H, Rieken S, Ecker S, et al. First experiences in treatment of low-grade glioma grade I and II with proton therapy. *Radiation Oncology*. 2012;7:189.
- [10] Palma G, Monti S, Conson M, et al. Normal tissue complication probability (NTCP) models for modern radiation therapy. *Seminars in Oncology: Elsevier*; 2019.
- [11] Paganetti H, Blakely E, Carabe-Fernandez A, et al. Report of the AAPM TG-256 on the relative biological effectiveness of proton beams in radiation therapy. *Med Phys*. 2019.
- [12] ICRU Report 78 - PRESCRIBING, RECORDING, AND REPORTING PROTON-BEAM THERAPY: CONTENTS. *Journal of the International Commission on Radiation Units and Measurements* 2007. p. NP-NP.
- [13] Deasy JO, Blanco AI, Clark VH. CERR: a computational environment for radiotherapy research. *Med Phys*. 2003;30:979-85.

- [14] Kyoizumi S, Suzuki T, Teraoka S, et al. Radiation sensitivity of human hair follicles in SCID-hu mice. *Radiat Res.* 1998;149:11-8.
- [15] Jimenez F, Izeta A, Poblet E. Morphometric analysis of the human scalp hair follicle: practical implications for the hair transplant surgeon and hair regeneration studies. *Dermatol Surg.* 2011;37:58-64.
- [16] Shih FY. *Image Processing and Mathematical Morphology: Fundamentals and Applications*: CRC Press; 2017.
- [17] Pastore F, Conson M, D'Avino V, et al. Dose-surface analysis for prediction of severe acute radio-induced skin toxicity in breast cancer patients. *Acta Oncol.* 2016;55:466-73.
- [18] Palma G, Cella L. A new formalism of Dose Surface Histograms for robust modeling of skin toxicity in radiation therapy. *Phys Med.* 2019;59:75-8.
- [19] Youden WJ. Index for rating diagnostic tests. *Cancer.* 1950;3:32-5.
- [20] Niemierko A. A generalized concept of equivalent uniform dose (EUD). *Med Phys* 1999;26.
- [21] Luxton G, Hancock SL, Boyer AL. Dosimetry and radiobiologic model comparison of IMRT and 3D conformal radiotherapy in treatment of carcinoma of the prostate. *International Journal of Radiation Oncology\* Biology\* Physics.* 2004;59:267-84.
- [22] Cella L, Palma G, Deasy JO, et al. Complication probability models for radiation-induced heart valvular dysfunction: do heart-lung interactions play a role? *PLoS One.* 2014;9:e111753.
- [23] Cella L, D'Avino V, Liuzzi R, et al. Multivariate normal tissue complication probability modeling of gastrointestinal toxicity after external beam radiotherapy for localized prostate cancer. *Radiat Oncol.* 2013;8:221.
- [24] Cella L, Liuzzi R, Conson M, et al. Multivariate normal tissue complication probability modeling of heart valve dysfunction in Hodgkin lymphoma survivors. *Int J Radiat Oncol Biol Phys.* 2013;87:304-10.
- [25] Cella L, D'Avino V, Palma G, et al. Modeling the risk of radiation-induced lung fibrosis: Irradiated heart tissue is as important as irradiated lung. *Radiother Oncol.* 2015;117:36-43.
- [26] Cohen J, Cohen P, West SG, et al. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*: CRC Press; 2002.
- [27] Dorr W, Hendry JH. Consequential late effects in normal tissues. *Radiother Oncol.* 2001;61:223-31.
- [28] Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology.* 2010;21:128-38.

- [29] Ginot A, Doyen J, Hannoun-Levi JM, et al. [Normal tissue tolerance to external beam radiation therapy: skin]. *Cancer Radiother.* 2010;14:379-85.
- [30] Palazzi M, Tomatis S, Orlandi E, et al. Effects of treatment intensification on acute local toxicity during radiotherapy for head and neck cancer: prospective observational study validating CTCAE, version 3.0, scoring system. *Int J Radiat Oncol Biol Phys.* 2008;70:330-7.
- [31] Rube CE, Fricke A, Widmann TA, et al. Accumulation of DNA damage in hematopoietic stem and progenitor cells during human aging. *PLoS One.* 2011;6:e17487.
- [32] Cella L, Lomax A, Miralbell R. New techniques in hadrontherapy: intensity modulated proton beams. *Phys Med.* 2001;17 Suppl 1:100-2.
- [33] Durante M, Orecchia R, Loeffler JS. Charged-particle therapy in cancer: clinical uses and future perspectives. *Nat Rev Clin Oncol.* 2017;14:483-95.
- [34] Paganetti H. Relative biological effectiveness (RBE) values for proton beam therapy. Variations as a function of biological endpoint, dose, and linear energy transfer. *Phys Med Biol.* 2014;59:R419-72.

## Figure legend

**Figure 1.** Acute Radiation Induced Alopecia (RIA) modelling. a) Average Dose-Surface Histograms (DSHs)  $\pm$  SEM (Standard Error of the mean) of patients with and without RIA; b) cross-validated ROC curves of Lyman-Kutcher-Burman (LKB) and multivariable logistic regression models (FPR: False Positive Rate, TPR: True Positive Rate); c) cross-validated calibration plot of LKB model; d) cross-validated calibration plot of logistic model; e) risk curve by LKB model with the observed fraction of complications from the data grouped in bins; f) risk curve by logistic regression model evaluated at the 1<sup>st</sup>, 3<sup>rd</sup> and 5<sup>th</sup> sextiles of the age distribution with the observed fraction of complications from the data grouped in bins for each tertile. In c)-f) the error bars for the reported values represent the 68% confidence intervals.

**Figure 2.** Late Radiation Induced Alopecia (RIA) modelling. a) Average Dose-Surface Histograms (DSHs) of patients with and without RIA; b) cross-validated ROC curves of Lyman-Kutcher-Burman (LKB) and multivariable logistic regression models (FPR: False Positive Rate, TPR: True Positive Rate); c) cross-validated calibration plot of LKB model; d) cross-validated calibration plot of logistic model; e) risk curve by LKB model with the observed fraction of complications from the data grouped in bins; f) risk curve by logistic regression model evaluated at the observed fraction of complications from the data grouped in bins for patients who underwent surgery and who did not. In c)-f) the error bars for the reported values represent the 68% confidence intervals.

**Figure 3.** Permanent Radiation Induced Alopecia (RIA) modelling. a) Average Dose-Surface Histograms (DSHs) of patients with and without RIA; b) cross-validated ROC curves of Lyman-Kutcher-Burman (LKB) and multivariable logistic regression models (FPR: False Positive Rate, TPR: True Positive Rate); c) cross-validated calibration plot of LKB model; d) cross-validated calibration plot of logistic model; e) risk curve by LKB model with the observed fraction of complications from the data grouped in bins;

f) risk curve by logistic regression model evaluated at the observed fraction of complications from the data grouped in bins. In c)-f) the error bars for the reported values represent the 68% confidence intervals.