



# PACE: A Probabilistic Atlas for Normal Tissue Complication Estimation in Radiation Oncology

Giuseppe Palma<sup>1\*</sup>, Serena Monti<sup>2</sup>, Amedeo Buonanno<sup>3</sup>, Roberto Pacelli<sup>4</sup> and Laura Cella<sup>1</sup>

<sup>1</sup> Institute of Biostructures and Bioimaging, National Research Council, Naples, Italy, <sup>2</sup> IRCCS SDN, Naples, Italy,

<sup>3</sup> Department of Engineering, University of Campania Luigi Vanvitelli, Aversa, Italy, <sup>4</sup> Department of Advanced Biomedical Sciences, Federico II University School of Medicine, Naples, Italy

## OPEN ACCESS

### Edited by:

Issam El Naqa,  
University of Michigan, United States

### Reviewed by:

Marianne Aznar,  
University of Manchester,  
United Kingdom  
Marco Durante,  
Istituto Nazionale di Fisica Nucleare,  
Italy

### \*Correspondence:

Giuseppe Palma  
giuseppe.palma@ibb.cnr.it

### Specialty section:

This article was submitted to  
Radiation Oncology,  
a section of the journal  
Frontiers in Oncology

**Received:** 15 October 2018

**Accepted:** 13 February 2019

**Published:** 13 March 2019

### Citation:

Palma G, Monti S, Buonanno A,  
Pacelli R and Cella L (2019) PACE: A  
Probabilistic Atlas for Normal Tissue  
Complication Estimation in Radiation  
Oncology. *Front. Oncol.* 9:130.  
doi: 10.3389/fonc.2019.00130

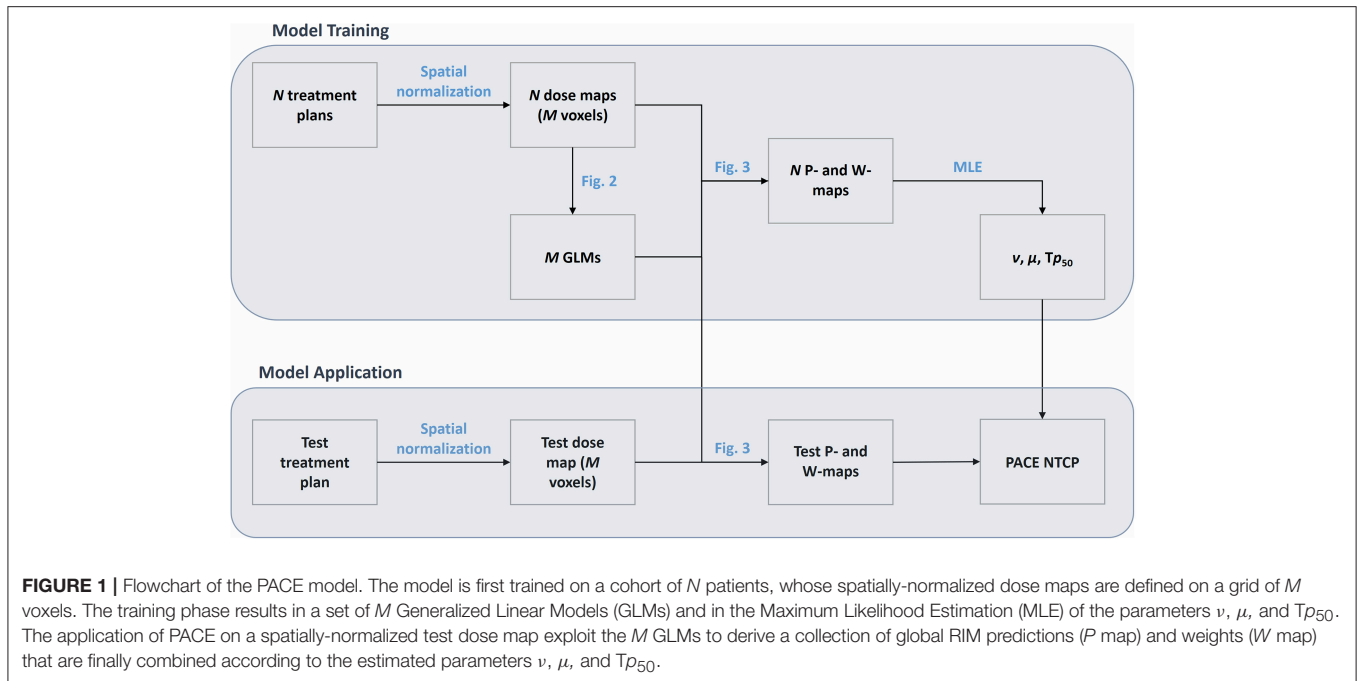
In radiation oncology, the need for a modern Normal Tissue Complication Probability (NTCP) philosophy to include voxel-based evidence on organ radio-sensitivity (RS) has been acknowledged. Here a new formalism (Probabilistic Atlas for Complication Estimation, PACE) to predict radiation-induced morbidity (RIM) is presented. The adopted strategy basically consists in keeping the structure of a classical, phenomenological NTCP model, such as the Lyman-Kutcher-Burman (LKB), and replacing the dose distribution with a collection of RIM odds, including also significant non-dosimetric covariates, as input of the model framework. The theory was first demonstrated *in silico* on synthetic dose maps, classified according to synthetic outcomes. PACE was then applied to a clinical dataset of thoracic cancer patients classified for lung fibrosis. LKB models were trained for comparison. Overall, the obtained learning curves showed that the PACE model outperformed the LKB and predicted synthetic outcomes with an accuracy >0.8. On the real patients, PACE performance, evaluated by both discrimination and calibration, was significantly higher than LKB. This trend was confirmed by cross-validation. Furthermore, the capability to infer the spatial pattern of underlying RS map for the analyzed RIM was successfully demonstrated, thus paving the way to new perspectives of NTCP models as learning tools.

**Keywords:** normal tissue complication probability, voxel-based analysis, radiation-induced morbidity, radio-sensitivity, radiation therapy

## INTRODUCTION

Early cancer detection and cancer treatment advances have contributed to improve local tumor control and overall survival. By January 1, 2024, it is estimated that in the USA the population of cancer survivors will increase to nearly 19 million individuals (1). This estimate supports research studies aimed at investigating the quality of life of these patients after the active phase of treatments, and the long-term effects of therapy.

Radiation induced damage to normal tissues is, indeed, the renowned Achilles' heel of radiation treatment of cancer. The risk of normal tissue complications associated with radiation therapy may even overshadow the benefits provided in terms of tumor control, and many cancer survivors must cope with long-term effects of radiation treatments that negatively affect their quality of life. Therefore, the development of mathematical models for the estimation of Normal Tissue Complication Probability (NTCP) has long been an active field of research in order to predict the



risk of radiation-induced morbidity (RIM) from the dose distribution released to critical organs (2, 3).

A major simplification traditionally applied for model-building was the introduction of the Dose-Volume Histograms (DVH) of organs at risk, designed to condense the full information of 3D dose distribution into an easy-to-handle mathematical tool, and exploited by the Lyman-Kutcher-Burman (LKB) (4) or Relative Seriality (5) models.

However, an increasing awareness has been triggered in the scientific community as to the value of including the spatial information of dose distributions within the analysis of RIM. The take-home messages of the studies that in the last 5 years have pioneered the voxel-based approach in RIM analysis fall into two categories. On the one hand, several studies highlighted an inhomogeneous radio-sensitivity of a specific organ related to the development of a given RIM (6–10); alternatively, a different application of regional analysis was proposed by (11) to allow a multi-organ study of dysphagia in head-and-neck cancer patients.

Overall, the statistical inference on spatial signature of RIM allowed to identify regions in which significant correlations exist between a global clinical outcome and the local dose release. While recognizing the urgency of including the regional inference in a planning optimization perspective, the above studies eventually resulted in identifying the mean dose to significant regions at an arbitrary  $\alpha$  level as a RIM predictor.

Relying on such predictors might be quite unsafe, and the definition of an avoidance region based on the significant clusters of the correlation  $p$ -maps seems at least as simplistic as setting a single constraint on the DVH of an organ at risk. In this context, a comprehensive NTCP model able to include full spatial info on dose distribution has not yet been achieved.

The purpose of this study is to propose a new formalism to fill this gap of knowledge and to develop a Probabilistic Atlas for normal tissue Complication Estimation in radiation therapy (PACE) in order to address the need for a modern NTCP philosophy that could fit the recent voxel-based evidences on organ radiobiology. The adopted strategy basically consists in keeping the general structure of a classical NTCP model approach, such as the LKB model, and replacing the dose distribution with a collection of RIM odds as input of the model framework. The theoretical structure was first demonstrated *in silico* and then applied to a clinical dataset of thoracic cancer patients.

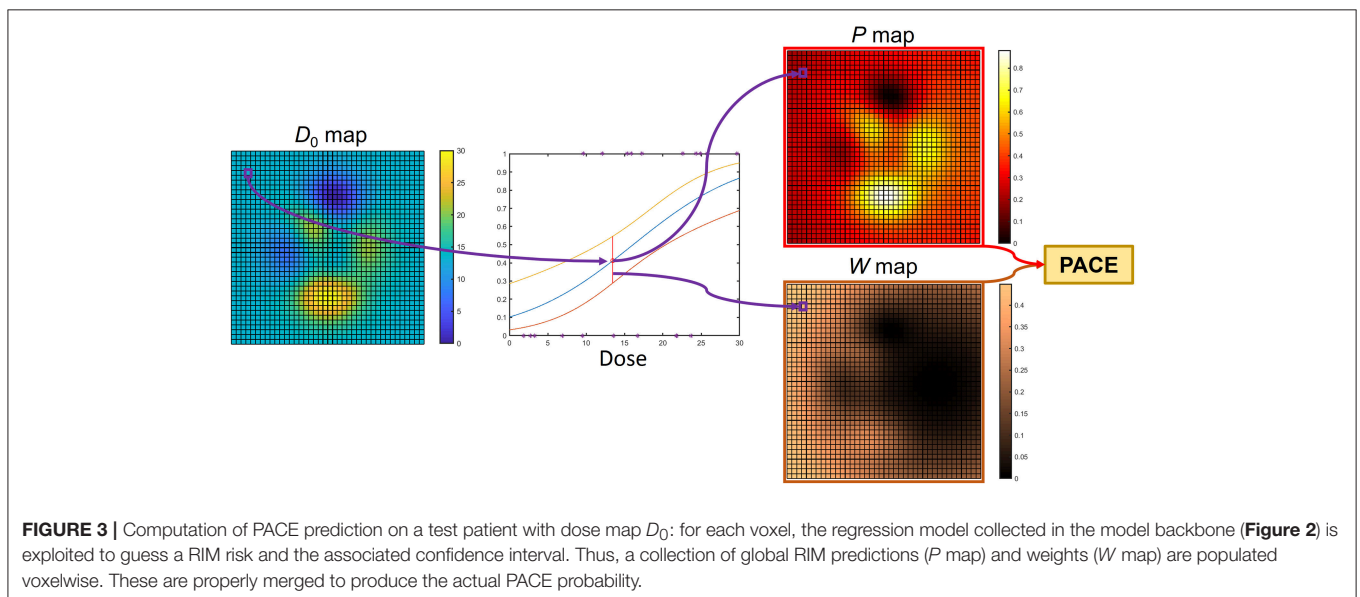
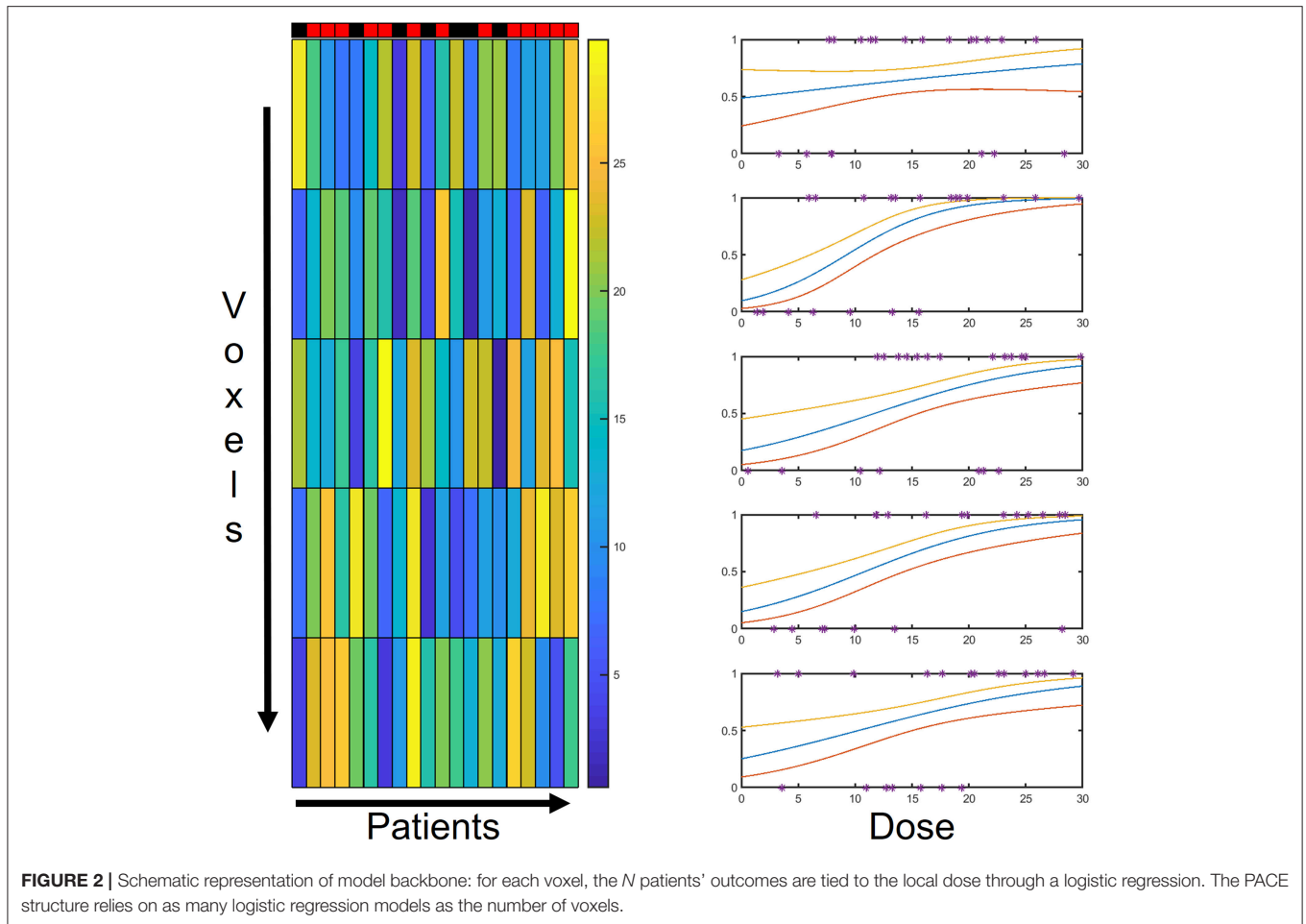
## MATERIALS AND METHODS

### Model Design

The proposed model (Figure 1) assumes the availability of a set of  $N$  3D dose distributions relative to as many patients classified according to a binary global outcome associated to the considered RIM. Each dose map  $D_i$  ( $i=1, \dots, N$ ) has to be spatially normalized to a common anatomical reference in a common coordinate system. In addition, a list of possible significant non-dosimetric variables  $\{V\}$  could be available, such as global covariates  $\{V_{i,k}\}$  (index  $k$  spans the non-dosimetric covariates like gender, age, medication, radiological features, etc.) or spatially normalized maps  $\{V_{i,k}(x_j)\}$  (CT, MRI, PET, etc.) defined on each voxel  $x_j$ .

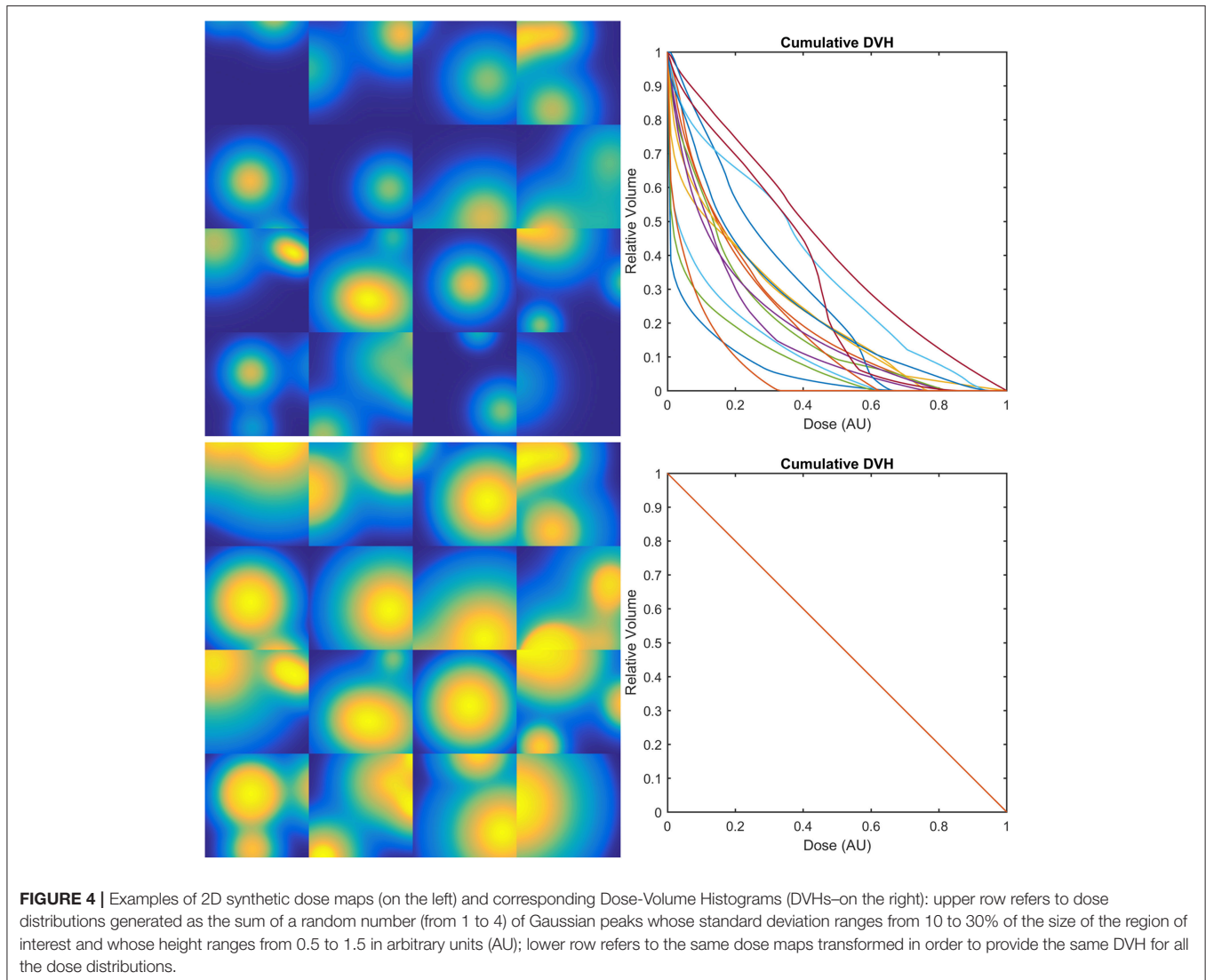
For each voxel, a logistic regression of the  $N$  outcomes is performed based on the  $N$  values of local dose  $\{D_i(x_j)\}_{i=1, \dots, N}$  (Figure 2) and of each covariate  $\{V_{i,k}(x_j)\}_{i=1, \dots, N}$ .

The computation of PACE prediction on a test patient with a dose map  $D_0$  is then obtained as follows (Figure 3). For each



voxel, the corresponding regression value  $P(x_j)$  and the associated 95% confidence interval  $CI(x_j)$  are evaluated based on the local  $D_0(x_j)$  and  $V_{0,k}(x_j)$ . Thus, a collection of global RIM predictions

$P(x_j)$  is populated voxelwise, along with the associated reliability scores given by  $W(x_j)=1/CI(x_j)$ , where the odds ratio for the dose is  $<1$  (i.e., dose appears to be protective),  $W(x_j)=0$ . The overall



**FIGURE 4 |** Examples of 2D synthetic dose maps (on the left) and corresponding Dose-Volume Histograms (DVHs—on the right): upper row refers to dose distributions generated as the sum of a random number (from 1 to 4) of Gaussian peaks whose standard deviation ranges from 10 to 30% of the size of the region of interest and whose height ranges from 0.5 to 1.5 in arbitrary units (AU); lower row refers to the same dose maps transformed in order to provide the same DVH for all the dose distributions.

PACE prediction is then summed up in a LKB fashion as

$$\text{PACE} = \frac{1 + \operatorname{erf} \frac{t}{\sqrt{2}}}{2}, \tag{1}$$

where

$$t = \frac{\text{gEUp} - Tp_{50}}{\mu \cdot Tp_{50}} \tag{2}$$

and gEUp (generalized equivalent uniform probability) is defined as

$$\text{gEUp} = \left[ \frac{\sum_j P_v^{\frac{1}{v}}(x_j) W(x_j)}{\sum_j W(x_j)} \right]^v. \tag{3}$$

The model contains three free parameters—namely  $v$  (equivalent to the volume effect parameter  $n$ ),  $\mu$  (controlling the slope of

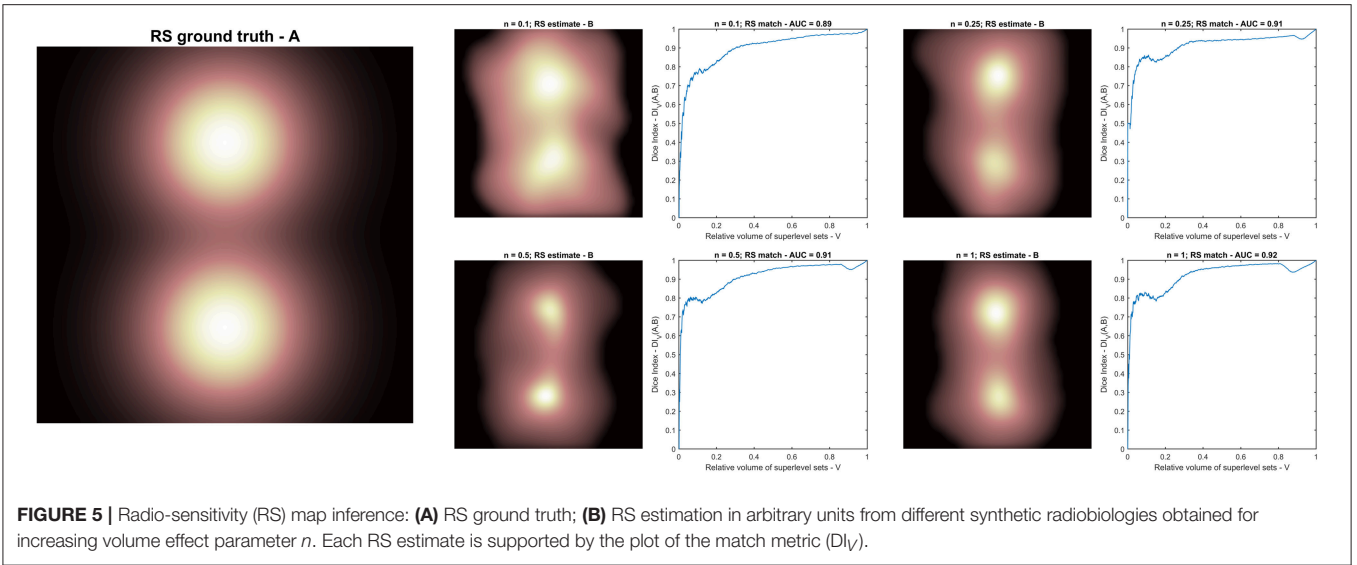
the response sigmoid curve) and the tolerance probability  $Tp_{50}$ —whose values can be customarily derived for a particular RIM by maximizing the likelihood function of the PACE model (12) [the related confidence interval are estimated according to Wilks’ theorem for nested models (13)].

### In silico Validation

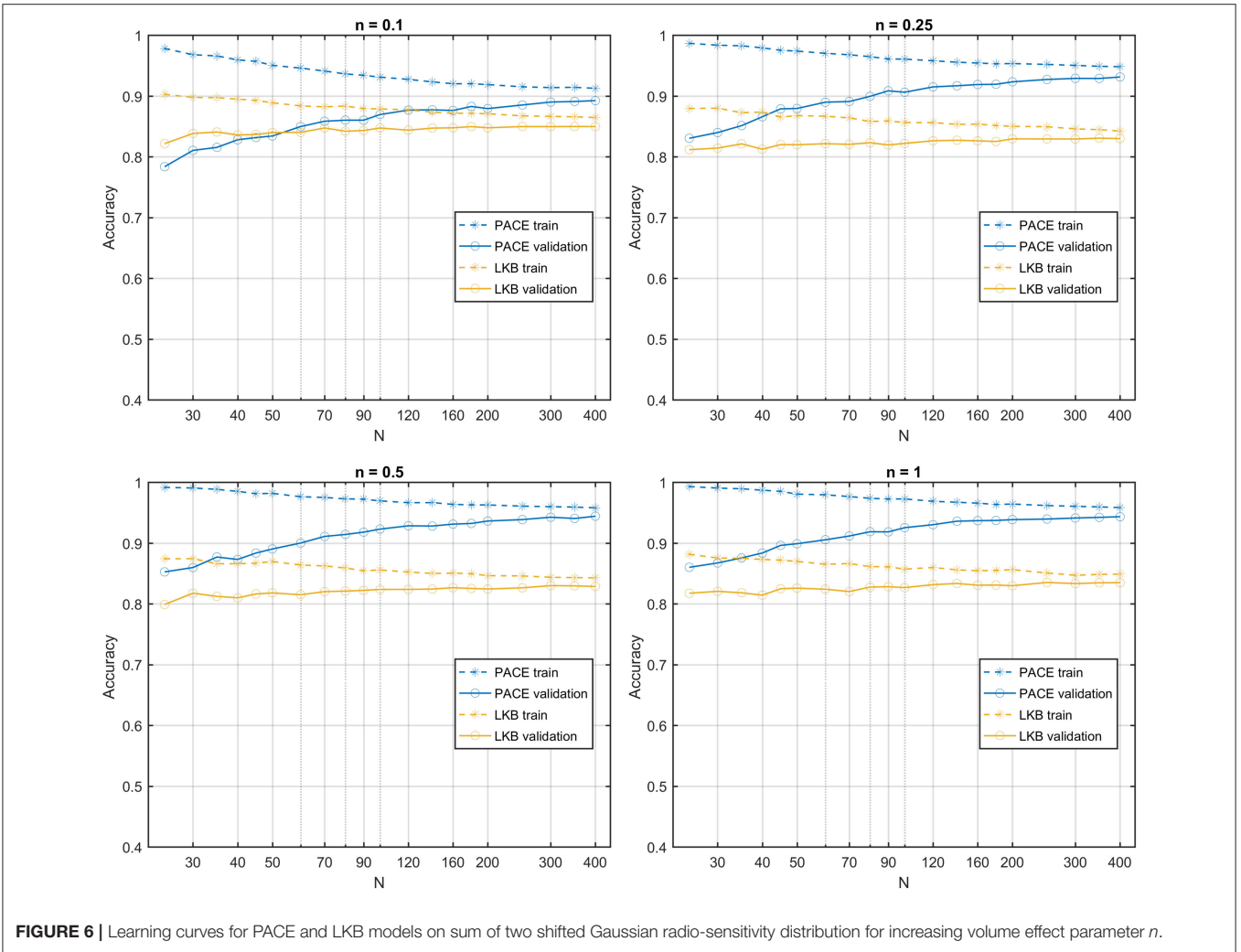
PACE model has been first demonstrated on two classes of sets (of increasing cardinality  $N$ ) of 2D synthetic dose maps defined on a square region of interest.

The dose distributions  $D_i(x_j)$  of the first class of sets were generated as the sum of a random number (from 1 to 4) of Gaussian peaks whose standard deviation ranges from 10 to 30% of the size of the region of interest and whose height ranges from 0.5 to 1.5 in arbitrary units (Figure 4, upper row).

The second class was obtained from the first one by transforming the original dose maps as  $E_i(x_j) = g_i[D_i(x_j)]$ , where each  $g_i$  is defined in order to provide the same constant



**FIGURE 5 |** Radio-sensitivity (RS) map inference: **(A)** RS ground truth; **(B)** RS estimation in arbitrary units from different synthetic radiologies obtained for increasing volume effect parameter  $n$ . Each RS estimate is supported by the plot of the match metric ( $Dl_V$ ).



**FIGURE 6 |** Learning curves for PACE and LKB models on sum of two shifted Gaussian radio-sensitivity distribution for increasing volume effect parameter  $n$ .

differential DVH on a compact support for all the  $E_i$  (Figure 4, lower row).

To classify the dose maps according to a realistic synthetic outcome, a new summary statistic, derived from the generalized Equivalent Uniform Dose (14), was designed for a given dose distribution  $D(x_j)$  as

$$g^2\text{EUD} = \left[ \frac{\sum_j D_j^{\frac{1}{n}}(x_j) \text{RS}(x_j)}{\sum_j \text{RS}(x_j)} \right]^n \quad (4)$$

in order to account for possible inhomogeneities in organ radio-sensitivity (RS), besides the usual volume effect parametrized by the organ dose-volume effect parameter  $n$  (2). For each class of dose maps, two RS maps were considered: a homogeneous one and one given by the sum of two shifted Gaussian distributions (S2G—Figure 5A). The  $g^2\text{EUDs}$  were then compared to a threshold value  $D_{th}$ , thus deriving a synthetic RIM outcome as  $O_i = [g^2\text{EUD}_i > D_{th}]$ . Of note, the second class of dose sets was

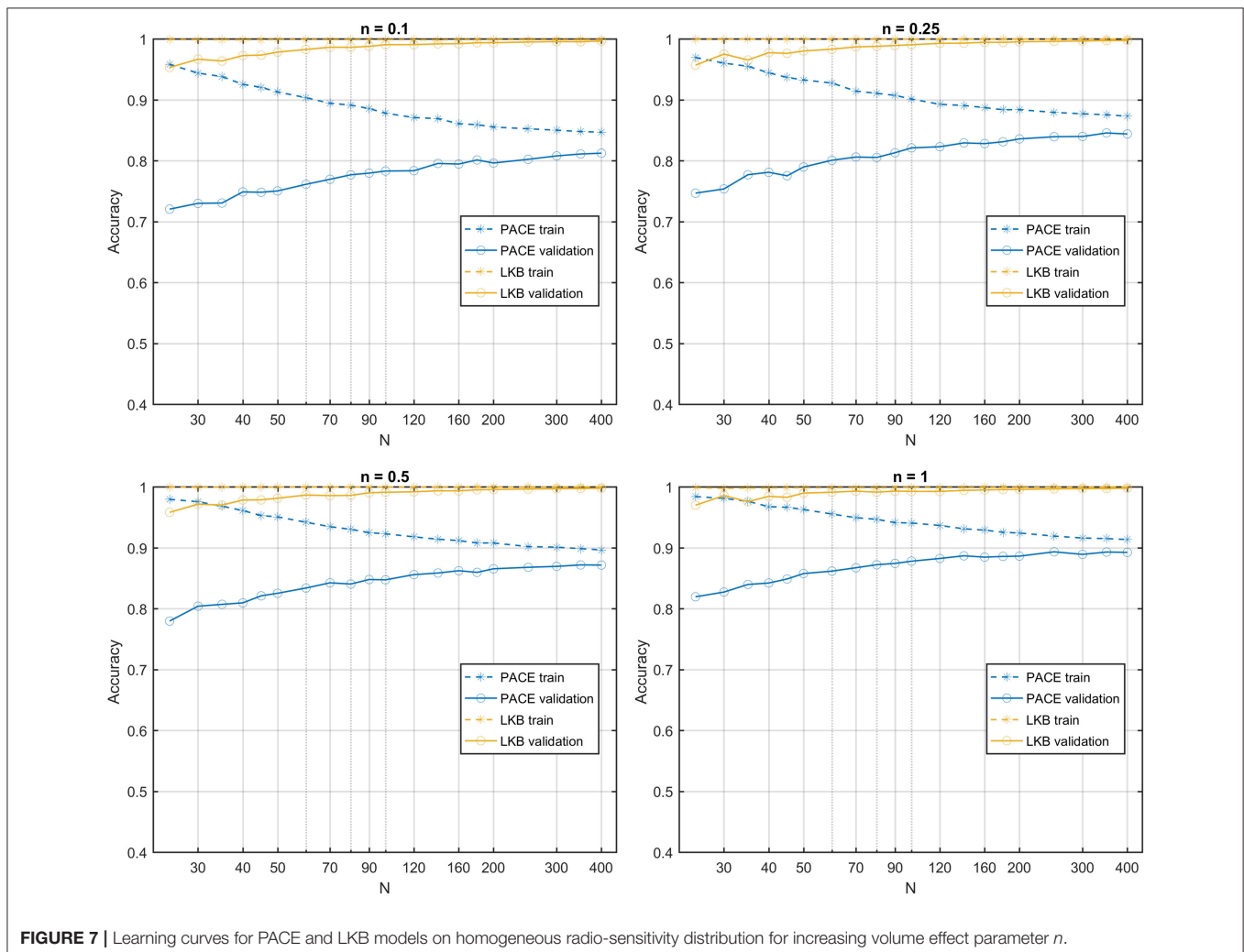
not evaluated according to a homogeneous RS, since the maps would have been indistinguishable.

Each set of classified dose distributions was then split in a training and a validation set, and the learning curves (15) of PACE and LKB (2) models were finally compared.

### Application to Clinical Data

The PACE model was applied to a cohort of  $N = 98$  thoracic cancer patients classified for lung fibrosis of any grade according to RTOG late pulmonary toxicity scoring system (18 events) (16). The age was the only non-dosimetric variable significantly correlated ( $p = 0.019$ ) with the considered RIM (17). All participants gave their written informed consent and patient data were analyzed anonymously. This retrospective study was approved by the local Ethics Committee (Comitato Etico per le Attività Biomediche, Università “Federico II,” Napoli, n. 222–10). All experimental protocols and procedures were performed in accordance with the guidelines of the Università “Federico II,” Napoli.

To satisfy PACE assumptions, each dose map was normalized to a common anatomical reference via a log-diffeomorphic



demons registration tool, as described in Palma et al. (8) and Monti et al. (17).

The full PACE model including dose maps and age, the purely dosimetric PACE (18) and the LKB models were trained for comparison.

Model performances were evaluated by the Area Under the Receiver Operating Characteristic (ROC) Curves (AUCs) and via the calibration plots. ROC-AUCs were compared by a Z-test according to the standard error estimates provided by (19). A Leave-One-Out (LOO) cross validation of the models was performed. The accuracy (i.e.,  $[True\ Positive + True\ Negative]/[Positive + Negative]$ ), the balanced accuracy (i.e.,  $[True\ Positive/Positive + True\ Negative/Negative]/2$ ) and the  $F_1$  score (i.e.,  $2 * True\ Positive/[2 * True\ Positive + False\ Positive + False\ Negative]$ ) of predictions were computed.

### Radio-Sensitivity Mapping

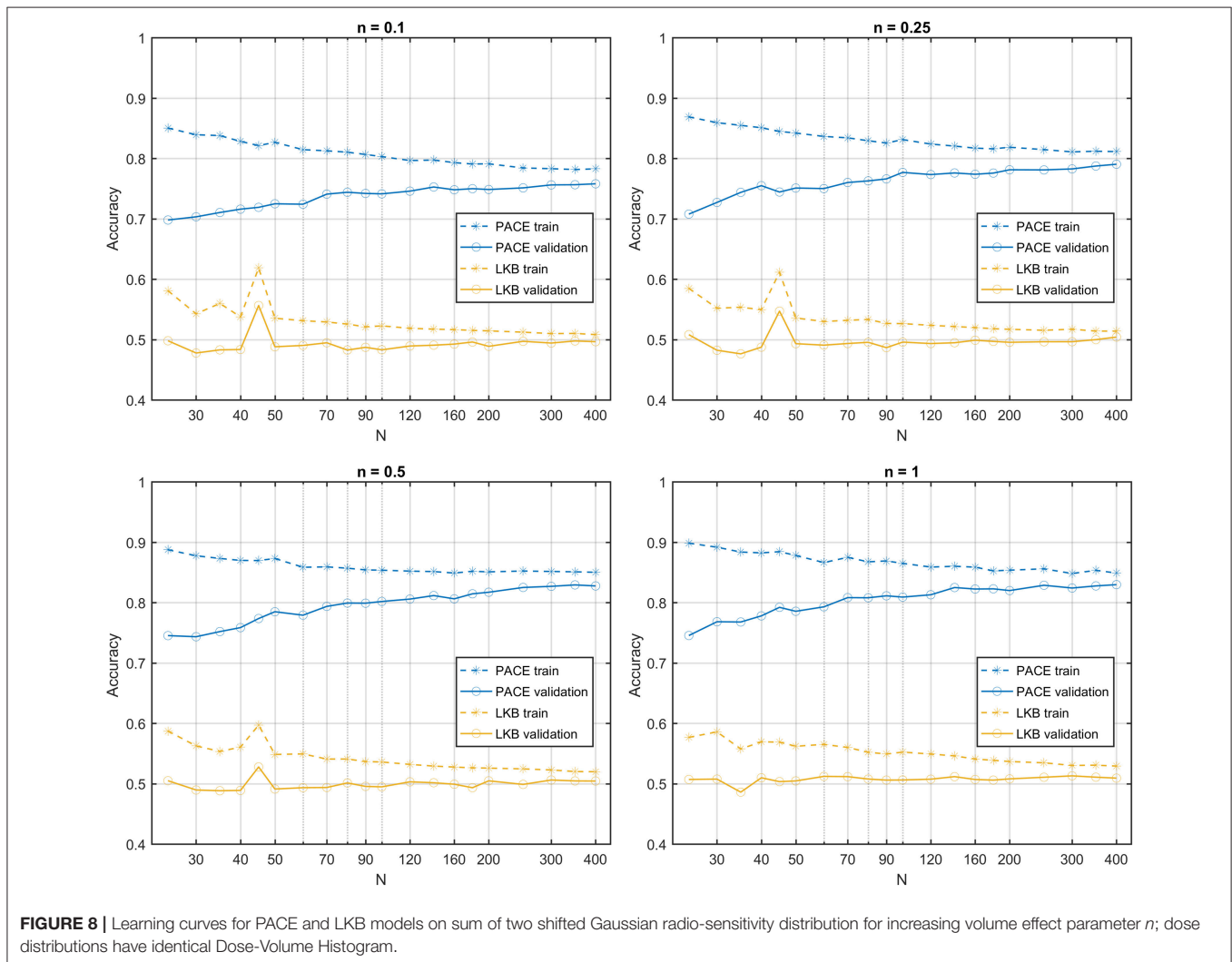
A spinoff application of the proposed approach was evaluated. Such application can lead beyond the strict sense of NTCP and potentially discloses new insights on RS maps. In particular,

once a PACE model has been trained on a given dataset, the level sets of the RS map can be inferred by computing PACE predictions on new purposely generated probe-set of synthetic dose distributions. The latter are designed with a single, narrow hot spot, whose position shifts along the set to span the entire region of interest. The PACE probability obtained for a given dose distribution of the probe-set can be then related to the position of the associated hot spot, thus getting a voxelwise measure of the underlying RS map.

The RS map inference was therefore evaluated on four PACE models; each model was trained on a dataset of 1,000 dose maps classified according to Equation 4 with a different value of the volume effect parameter  $n$  and using the same S2G as RS map (Figure 5).

In order to quantify the match between the RS ground truth (A) and each *a posteriori* estimate (B), the following metrics is defined as the Dice Index (DI) (20) between the superlevel sets of A ( $S_V[A]$ ) and B ( $S_V[B]$ ) with same size  $V$  (17):

$$DI_V(A, B) = DI(S_V[A], S_V[B]) \tag{5}$$



**TABLE 1** | Training and validation of PACE, PACE<sub>NoAge</sub>, and LKB models on clinical data: Parameters and performance scores.

	PACE		PACE <sub>NoAge</sub>		LKB	
<b>Parameters (95% CI)</b>	$\nu$	0.05 ([0.01–0.10])	$\nu$	0.12 ([0.01–0.25])	$n$	0.06 ([0.01–1])
	$\mu$	0.39 ([0.29–0.55])	$\mu$	0.34 ([0.20–0.54])	$m$	0.25 ([0.13–0.98])
	$Tp_{50}$	0.55 ([0.42–0.76])	$Tp_{50}$	0.38 ([0.25–0.75])	$TD_{50}$	36 ([18–300]) Gy
Discrimination value	0.37		0.10		0.20	
<b>Performance</b>						
AUC (95% CI)	0.85 ([0.76–0.91])		0.79 ([0.70–0.87])		0.66 ([0.56–0.76])	
Calibration slope $\pm$ SE	1.06 $\pm$ 0.13		0.92 $\pm$ 0.14		0.76 $\pm$ 0.39	
Calibration intercept $\pm$ SE	–0.010 $\pm$ 0.037		–0.012 $\pm$ 0.039		0.012 $\pm$ 0.078	
Calibration $R^2$	0.93		0.90		0.50	
LOO Accuracy	0.74		0.82		0.70	
LOO Balanced accuracy	0.67		0.65		0.62	
LOO $F_1$	0.44		0.44		0.38	
LOO AUC (95% CI)	0.75 ([0.65–0.83])		0.67 ([0.57–0.76])		0.59 ([0.48–0.68])	
LOO Calibration slope $\pm$ SE	0.56 $\pm$ 0.11		0.55 $\pm$ 0.14		0.29 $\pm$ 0.24	
LOO Calibration intercept $\pm$ SE	0.063 $\pm$ 0.036		–0.061 $\pm$ 0.045		0.103 $\pm$ 0.047	
LOO Calibration $R^2$	0.83		0.77		0.43	

CI, confidence interval; AUC, Area Under the ROC Curve; SE, Standard Error; LOO, Leave One Out.

The NTCP models as well as the evaluation steps were implemented in house in Matlab (MATLAB® Release 2016b, The MathWorks, Inc., Natick, MA, USA).

## RESULTS

### In silico Performances

The learning curves obtained on the first class of datasets classified according to a S2G RS map show that the PACE model outperforms the LKB model (Figure 6), with a slightly increasing model bias at lower  $n$  values. When the same class of datasets is ranked according to a homogeneous RS map, by construction, the LKB model turns out to be, the perfect classifier (Equation 4 coincides with standard gEUD definition). Nonetheless, PACE proves still able to provide an accuracy around the 0.8 level (Figure 7).

In the second class of datasets, all dose maps are summarized by the same DVH. Therefore, the LKB model turns out to be a random classifier (expected accuracy = 0.5), while the PACE model still predicts the outcome with an accuracy >0.8 (Figure 8).

### Clinical Data Demonstration

The training process of the considered models on the real patients' data was summarized in Table 1.

The models consistently highlighted a similar volume effect for the organ response to irradiation (see  $\nu$  and  $n$  parameters); however, the confidence intervals of the three free parameters estimated for PACE models were systematically smaller than the LKB ones.

On the training set, both PACE models outperformed the LKB model in terms of discrimination (Figure 9A): at a pairwise comparison with the LKB ROC curve, the full PACE model showed an AUC significantly higher ( $p = 0.018$ ), while a trend

was found for the purely dosimetric PACE model ( $p = 0.054$ ). The calibration curves of both PACE models show higher  $R^2$  compared to LKB, with calibration-in-the-large  $a$  and calibration slope  $b$  close to the ideal (i.e.,  $a = 0$  and  $b = 1$ ), in contrast to LKB (Figures 9B–D).

A LOO cross validation confirmed the better performances of PACE models compared to LKB model (Figure 10A), as shown by the accuracy and balanced accuracy of outcome predictions. Furthermore, the above described patterns of discrimination and calibration survived at validation (Figures 10B–D).

### Radio-Sensitivity Inference

The RS maps probed with the PACE models trained on the synthetic datasets suggest a trend in the spatial pattern for increasing  $n$ . Nonetheless, the AUC of  $DI_V$  between inferred and ground truth RS maps (overall close to 0.9) reveals that the shape of the level sets is only mildly dependent on  $n$  (Figure 5).

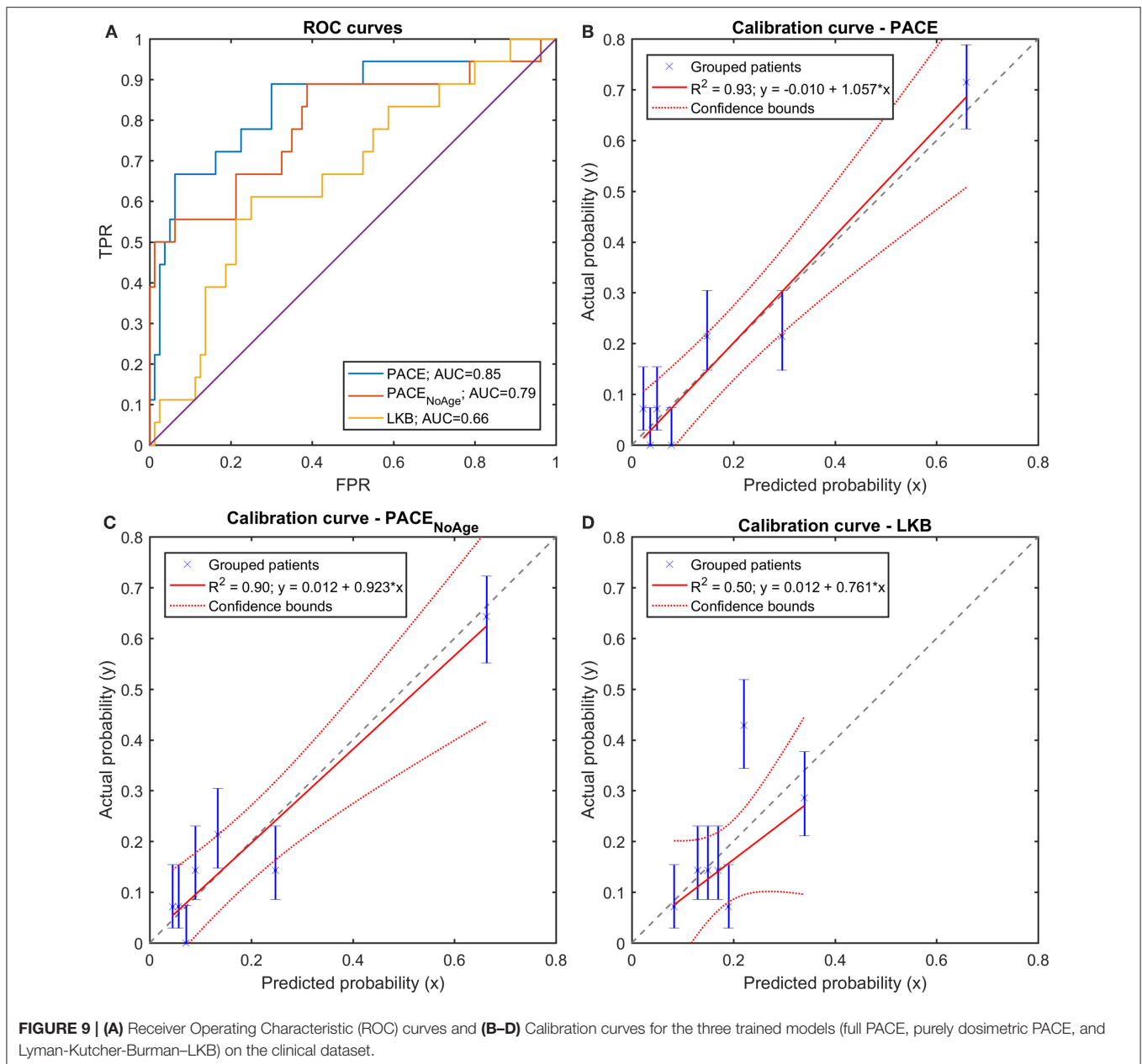
## DISCUSSION

We devised a mathematical solution for the statistical modeling of NTCP that allows to deal with inhomogeneous organ susceptibility to radiation exposure. This was solicited by the recent evidences from voxel-based analyses and by two-decade studies about non-local dose effects on tissue damages (21).

In the last ten years several attempts have been made to include heterogeneous dose-response relationships into NTCP models.

Some studies highlighted a spatially dependent point-to-point correlation between dose and induced damage, which, in the context of lung toxicity, could be indexed by the variation of CT density (accounting for fibrosis) (22, 23) or of SPECT activity (measuring the perfusion) (24). Despite the value of such findings, the above studies were not suggesting an actual NTCP



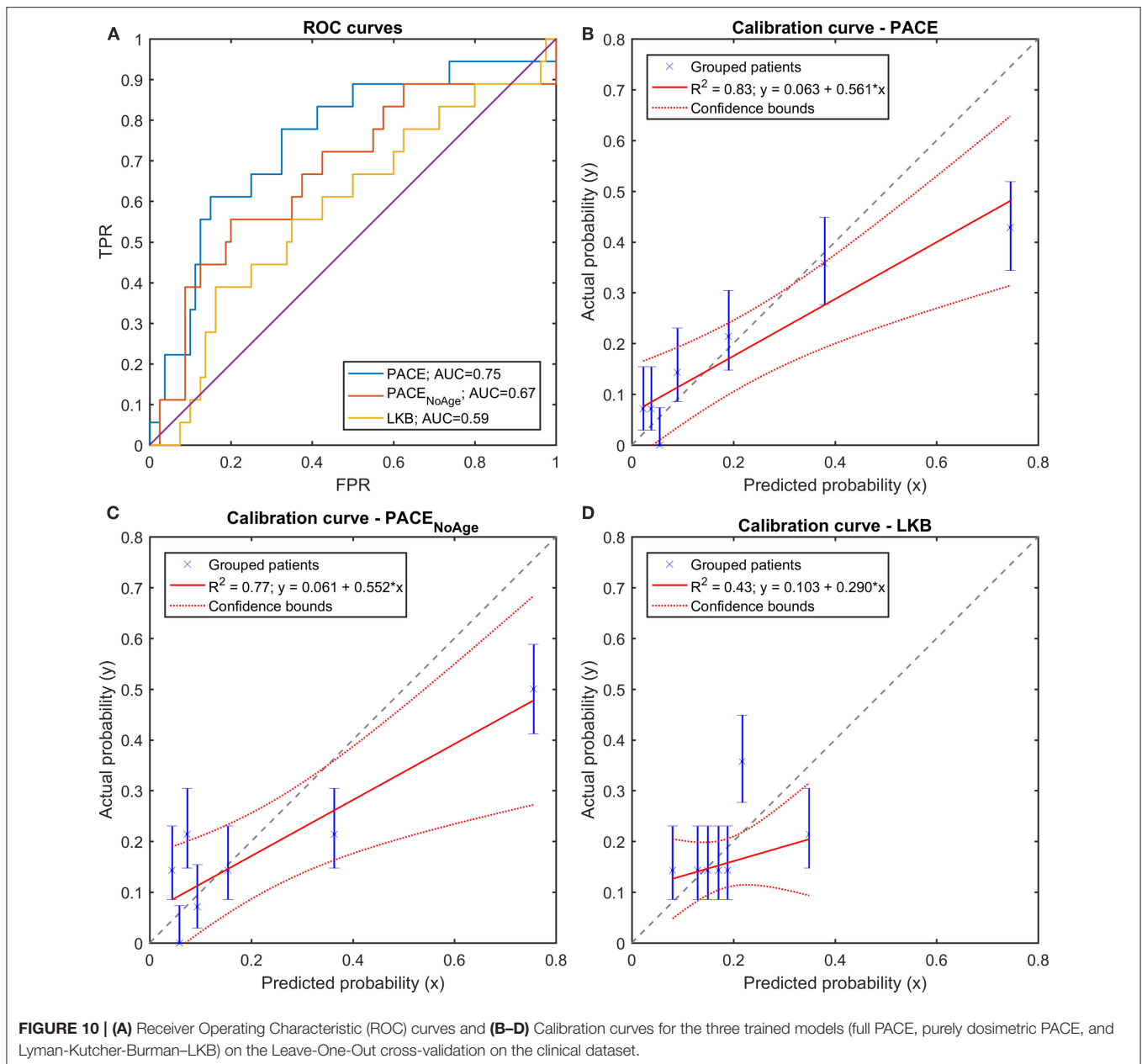


model, since no global outcome was related to the predicted local radiological variations.

Differently, Dean and coworkers (25, 26) included spatial dose metrics (3D moment invariants) of the oral mucosa within their multivariable analysis of severe acute mucositis, thus accounting for possible regional variations in RS through some summary parameters. In a following study on dysphagia resulting from head-and-neck radiation therapy (27), they found that such approach did not add any significant predictive power to the DVH analysis, possibly due to the high correlation between such spatial metrics and the ordinary DVH metrics. Limited to the prediction of rectal bleeding following prostate cancer radiation therapy, an ensemble of neural networks was trained

in (28) to find the relevant features in a 2D dose-surface map of the rectum. Similarly, in (29) it is shown that the predictability of gastro-intestinal toxicity increases using spatial metrics compared to DVH metrics. Furthermore, some studies on dose map population comparison highlighted a somehow strong correlation between the RIM outcome and the mean dose to a sublevel set of the significance  $p$ -map (7, 8, 11). However, such statistics could be barely robust predictors of toxicity since they are intrinsically unable to account for differently significant spatial dose signature of RIM.

The proposed model falls within the large framework of the phenomenological models, which, in opposition to mechanistic models (30), aim at a substantial consistency with the



available data without necessarily relying on a fully established radiobiological background (31).

The functional form of the PACE model was designed to take into account some key points. First, a huge body of literature on NTCP suggests that the radiation toxicity is a function more or less related to the extent of the dose release, depending on structure and physiology of the organ and on the considered endpoint. Second, same doses to different subregions could in principle result in different odds of global damage. Finally, the relationship between dose and outcome can be more or less strong within the considered anatomical district.

The first point suggested to adopt a theoretical framework including a parameter for the description of the dose-volume

effect, such as the largely exploited DVH-based LKB model (Equation 1). The second point led to replace the dose distribution with a collection of odds specifically estimated for the spatial position of local dose release (Equation 2). Finally, the third point was addressed by weighting the odds according to their uncertainty (Equation 3).

The PACE framework, by construction, allows to naturally model a multivariable phenomenon such a RIM, by including possible non-dosimetric covariates (potentially point functions) within the assessment of NTCP.

The model was first demonstrated *in silico*. To this purpose, we devised and implemented several synthetic plausible radiobiologies, in order to simulate diverse RS maps and

several dose-volume effects. PACE scheme was able to learn the responses assigned to the dose distributions with high accuracy and robustness, consistently outperforming the LKB approach in a wide range of synthetic physiopathologies. Not surprisingly, the reported learning curves revealed that, compared to LKB, the PACE model requires a higher number of dose distributions to reduce the deviation between training and validation accuracies, as a result of the wider spectrum of information it has to take into account. Nevertheless, the overall lower model bias exhibited by PACE allows to obtain more accurate predictions in the validation set even at moderately low cardinality ( $N < 50$ ) of the training set.

This is reflected in the application to a real clinical dataset of about 100 patients with a relatively low number of RIM events (18%). The estimates of the volume-effect parameters ( $\nu$  for PACE or  $n$  for LKB) highlighted the relevance of the value of the highest doses in the RIM prediction, which is consistent with some previous findings on lung toxicity (16, 32). Interestingly, the same dataset permits to obtain much sharper estimates of the PACE model parameters than those provided by the LKB.

PACE performance, evaluated by both discrimination and calibration scores, was significantly higher than the LKB benchmark. This trend was confirmed by the cross-validation, which ruled out severe overfitting issues, despite the complexity of the PACE structure.

It is worth mentioning that the better accuracy achieved by the full PACE model is not to be ascribed just to the inclusion of a significant clinical covariate (i.e., the age). Indeed, while the PACE superiority bears witness to the opportunity to incorporate non-dosimetric variables, the scores obtained by the simpler PACE<sub>NoAge</sub> still prove a sensible improvement over a conventional NTCP scheme.

Remarkably, a further application of this new NTCP philosophy concerns the capability to infer the spatial pattern of underlying RS map for the analyzed RIM. The fulfillment of such task was successfully demonstrated for several synthetic radiobiologies, showing that the RS mapping tool applies pretty well over a wide spectrum of dose-volume effect behavior driven by  $n$ . This enables a change in perspective for the usage of NTCP models, which turns out to be a valuable learning tool for knowledge building if properly queried. A key point in this

context, strictly related to the generalizability of the PACE model, are the characteristics of the learning cohort of patients, which—similarly to what happens in the training of standard DVH-based NTCP models—is warranted to show highly inhomogeneous and mutually uncorrelated dose distributions (11).

At the same time, the very nature of this new NTCP scheme, which takes into account the full 3D distribution of dose, makes it particularly suitable for studying toxicity outcomes related to modern radiation therapy techniques, such as stereotactic body radiation therapy and particle therapy, characterized by greatly heterogeneous dose distributions.

We are aware that the increased complexity of a PACE model, compared to the corresponding LKB function, may hinder the diffusion of this scheme in the clinical practice. Nonetheless, an increasing consciousness in the radiation oncology community of the inhomogeneous RS of several organs justifies and actually calls for an adequate and modern NTCP approach. On the other hand, PACE scheme can be considered in all respects within the large family of machine learning algorithms, which find more and more room within the recent releases of treatment planning systems.

In conclusion, we designed and demonstrated, both *in silico* and in clinical datasets, an original, phenomenological NTCP model, which we nicknamed PACE, to blend the need for RIM prediction with the awareness of inhomogeneous dose susceptibility. It is conceived to work for different RIM outcomes from arbitrary district irradiation, with minimal research cognitive bias. We believe this could open new perspectives for clinical radiobiology.

## DATA AVAILABILITY

The datasets analyzed for this study can be found in <http://www.ibb.cnr.it/?command=viewcms&and2=32&id=226>.

## AUTHOR CONTRIBUTIONS

GP and LC conceived and designed the study. RP participated in patient recruitment and collected the clinical data. GP wrote the code. GP, SM, and LC processed the data. GP and AB analyzed the results. GP and LC wrote the manuscript. All authors reviewed the final manuscript.

## REFERENCES

- DeSantis CE, Lin CC, Mariotto AB, Siegel RL, Stein KD, Kramer JL, et al. Cancer treatment and survivorship statistics, 2014. *CA Cancer J Clin.* (2014) 64:252–71. doi: 10.3322/caac.21235
- Marks LB, Yorke ED, Jackson A, Ten Haken RK, Constine LS, Eisbruch A, et al. Use of normal tissue complication probability models in the clinic. *Int J Radiat Oncol Biol Phys.* (2010) 76:S10–19. doi: 10.1016/j.ijrobp.2009.07.1754
- Rubin P, Constine LS, Marks LB. *ALERT—Adverse Late Effects of Cancer Treatment. Volume 1: General Concepts and Specific Precepts.* (2014) Berlin; London: Springer.
- Kutcher GJ, Burman C, Brewster L, Goitein M, Mohan R. Histogram reduction method for calculating complication probabilities for three-dimensional treatment planning evaluations. *Int J Radiat Oncol Biol Phys.* (1991) 21:137–46. doi: 10.1016/0360-3016(91)90173-2
- Kallman P, Agren A, Brahme A. Tumour and normal tissue responses to fractionated non-uniform dose delivery. *Int J Radiat Biol.* (1992) 62:249–62. doi: 10.1080/09553009214552071
- Acosta O, Drean G, Ospina JD, Simon A, Haigron P, Lafond C, et al. Voxel-based population analysis for correlating local dose and rectal toxicity in prostate cancer radiotherapy. *Phys Med Biol.* (2013) 58:2581–95. doi: 10.1088/0031-9155/58/8/2581
- Drean G, Acosta O, Ospina JD, Fargeas A, Lafond C, Corregge G, et al. Identification of a rectal subregion highly predictive of rectal bleeding in prostate cancer IMRT. *Radiother Oncol.* (2016) 119:388–97. doi: 10.1016/j.radonc.2016.04.023

8. Palma G, Monti S, D'Avino V, Conson M, Liuzzi R, Pressello MC, et al. A voxel-based approach to explore local dose differences associated with radiation-induced lung damage. *Int J Radiat Oncol Biol Phys.* (2016) 96:127–33. doi: 10.1016/j.ijrobp.2016.04.033
9. Palorini F, Cozzarini C, Gianolini S, Botti A, Carillo V, Iotti C, et al. First application of a pixel-wise analysis on bladder dose-surface maps in prostate cancer radiotherapy. *Radiother Oncol.* (2016) 119:123–8. doi: 10.1016/j.radonc.2016.02.025
10. Yahya N, Ebert MA, House MJ, Kennedy A, Matthews J, Joseph DJ, et al. Modeling urinary dysfunction after external beam radiation therapy of the prostate using bladder dose-surface maps: evidence of spatially variable response of the bladder surface. *Int J Radiat Oncol Biol Phys.* (2017) 97:420–6. doi: 10.1016/j.ijrobp.2016.10.024
11. Monti S, Palma G, D'Avino V, Gerardi M, Marvaso G, Ciardo D, et al. Voxel-based analysis unveils regional dose differences associated with radiation-induced morbidity in head and neck cancer patients. *Sci Rep.* (2017) 7:7220. doi: 10.1038/s41598-017-07586-x
12. Cella L, Palma G, Deasy JO, Oh JH, Liuzzi R, D'Avino V, et al. Complication probability models for radiation-induced heart valvular dysfunction: do heart-lung interactions play a role? *PLoS ONE.* (2014) 9:e111753. doi: 10.1371/journal.pone.0111753
13. Behnke O, Kröniger K, Schott GG, Schörner-Sadenius T. *Data Analysis in High Energy Physics: a Practical Guide to Statistical Methods.* WILEY-VCH Verlag GmbH & Co. KGaA. doi: 10.1002/9783527653416
14. Niemierko A. A generalized concept of equivalent uniform dose (EUD). *Med Phys.* (1999) 26:1100.
15. Duda RO, Hart PE, Stork DG. *Pattern classification.* (2001). New York, NY: Wiley.
16. Cella L, D'Avino V, Palma G, Conson M, Liuzzi R, Picardi M, 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. doi: 10.1016/j.radonc.2015.07.051
17. Monti S, Pacelli R, Cella L, Palma G. Inter-patient image registration algorithms to disentangle regional dose bioeffects. *Sci Rep.* (2018) 8:4915. doi: 10.1038/s41598-018-23327-0
18. Palma G, Buonanno A, Monti S, Pacelli R, Cella L. Space based normal tissue complication probability modeling. *Radiother Oncol.* (2018) 127:S267–8. doi: 10.1016/S0167-8140(18)30822-3
19. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* (1988) 44:837–45. doi: 10.2307/2531595
20. Dice LR. Measures of the amount of ecologic association between species. *Ecology.* (1945) 26:297–302. doi: 10.2307/1932409
21. Trott KR, Doerr W, Facoetti A, Hopewell J, Langendijk J, van Luijk P, et al. Biological mechanisms of normal tissue damage: importance for the design of NTCP models. *Radiother Oncol.* (2012) 105:79–85. doi: 10.1016/j.radonc.2012.05.008
22. Avanzo M, Barbiero S, Trovo M, Bissonnette JP, Jena R, Stancanello J, et al. Voxel-by-voxel correlation between radiologically radiation induced lung injury and dose after image-guided, intensity modulated radiotherapy for lung tumors. *Phys Med.* (2017) 42:150–6. doi: 10.1016/j.ejmp.2017.09.127
23. Defraene G, van Elmpt W, Crijns W, De Ruyscher D. Regional variability in radiation-induced lung damage can be predicted by baseline CT numbers. *Radiother Oncol.* (2017) 122:300–6. doi: 10.1016/j.radonc.2016.11.021
24. Scheenstra AE, Rossi MM, Belderbos JS, Damen EM, Lebesque JV, Sonke JJ. Local dose-effect relations for lung perfusion post stereotactic body radiotherapy. *Radiother Oncol.* (2013) 107:398–402. doi: 10.1016/j.radonc.2013.04.003
25. Dean JA, Wong KH, Welsh LC, Jones AB, Schick U, Newbold KL, et al. Normal tissue complication probability (NTCP) modelling using spatial dose metrics and machine learning methods for severe acute oral mucositis resulting from head and neck radiotherapy. *Radiother Oncol.* (2016) 120:21–7. doi: 10.1016/j.radonc.2016.05.015
26. Dean JA, Welsh LC, Wong KH, Aleksic A, Dunne E, Islam MR, et al. Normal tissue complication probability (NTCP) modelling of severe acute mucositis using a novel oral mucosal surface organ at risk. *Clin Oncol.* (2017) 29:263–73. doi: 10.1016/j.clon.2016.12.001
27. Dean J, Wong K, Gay H, Welsh L, Jones AB, Schick U, et al. Incorporating spatial dose metrics in machine learning-based normal tissue complication probability (NTCP) models of severe acute dysphagia resulting from head and neck radiotherapy. *Clin Transl Radiat Oncol.* (2018) 8:27–39. doi: 10.1016/j.ctro.2017.11.009
28. Buettner F, Gulliford SL, Webb S, Partridge M. Using dose-surface maps to predict radiation-induced rectal bleeding: a neural network approach. *Phys Med Biol.* (2009) 54:5139–53. doi: 10.1088/0031-9155/54/17/005
29. Casares-Magaz O, Muren LP, Moiseenko V, Petersen SE, Pettersson NJ, Hoyer M, et al. Spatial rectal dose/volume metrics predict patient-reported gastrointestinal symptoms after radiotherapy for prostate cancer. *Acta Oncol.* (2017) 56:1507–13. doi: 10.1080/0284186X.2017.1370130
30. Rutkowska E, Baker C, Nahum A. Mechanistic simulation of normal-tissue damage in radiotherapy—implications for dose-volume analyses. *Phys Med Biol.* (2010) 55:2121–36. doi: 10.1088/0031-9155/55/8/001
31. van der Schaaf A, Langendijk JA, Fiorino C, Rancati T. Embracing phenomenological approaches to normal tissue complication probability modeling: a question of method. *Int J Radiat Oncol Biol Phys.* (2015) 91:468–71. doi: 10.1016/j.ijrobp.2014.10.017
32. Hope AJ, Lindsay PE, El Naqa I, Alaly JR, Vicic M, Bradley JD, et al. Modeling radiation pneumonitis risk with clinical, dosimetric, and spatial parameters. *Int J Radiat Oncol Biol Phys.* (2006) 65:112–24. doi: 10.1016/j.ijrobp.2005.11.046

**Conflict of Interest Statement:** The reviewer MD declared a past co-authorship with one of the authors LC to the handling Editor.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Palma, Monti, Buonanno, Pacelli and Cella. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.