Radiotherapy and Oncology xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original article

Independent component analysis for rectal bleeding prediction following prostate cancer radiotherapy

Auréline Fargeas [a,b,*], Oscar Acosta [a,b], Juan David Ospina Arrango [a,b], Amine Ferhat [a,b], Nathalie Costet [a,b], Laurent Albera [a,b], David Azria [d], Pascal Fenoglietto [d], Gilles Créhange [e], Véronique Beckendorf [f], Mathieu Hatt [g], Amar Kachenoura [a,b], Renaud de Crevoisier [a,b,c]

[a] *INSERM, UMR1099, Rennes;* [b] *LTSI, Université de Rennes 1;* [c] *Département de radiothérapie, Centre Eugène Marquis, Rennes;* [d] *Département d'oncologie radiothérapie, INSERM U860, centre de recherche en cancérologie de Montpellier, CRLC Val-d'Aurelle Paul-Lamarque;* [e] *Centre Georges François Leclerc, Dijon;* [f] *Centre Alexis Vautrin, Vandoeuvre les Nancy; and* [g] *LaTIM, INSERM UMR1101, IBSAM, CHRU Morvan, Brest, France*

ARTICLE INFO

Article history:
Received 29 July 2016
Received in revised form 18 November 2017
Accepted 20 November 2017
Available online xxxx

Keywords:
Prostate cancer radiotherapy
Toxicity
Rectal bleeding
Predictive model
Independent component analysis

ABSTRACT

Background and purpose: To evaluate the benefit of independent component analysis (ICA)-based models for predicting rectal bleeding (RB) following prostate cancer radiotherapy.
Materials and methods: A total of 593 irradiated prostate cancer patients were prospectively analyzed for Grade ≥2 RB. ICA was used to extract two informative subspaces (presenting RB or not) from the rectal DVHs, enabling a set of new *pICA* parameters to be estimated. These DVH-based parameters, along with others from the principal component analysis (PCA) and functional PCA, were compared to "standard" features (patient/treatment characteristics and DVH bins) using the Cox proportional hazards model for RB prediction. The whole cohort was divided into: (i) training (N = 339) for ICA-based subspace identification and Cox regression model identification and (ii) validation (N = 254) for RB prediction capability evaluation using the C-index and the area under the receiving operating curve (AUC), by comparing predicted and observed toxicity probabilities.
Results: In the training cohort, multivariate Cox analysis retained *pICA* and *PC* as significant parameters of RB with 0.65 C-index. For the validation cohort, the C-index increased from 0.64 when *pICA* was not included in the Cox model to 0.78 when including *pICA* parameters. When *pICA* was not included, the AUC for 3-, 5-, and 8-year RB prediction were 0.68, 0.66, and 0.64, respectively. When included, the AUC increased to 0.83, 0.80, and 0.78, respectively.
Conclusion: Among the many various extracted or calculated features, ICA parameters improved RB prediction following prostate cancer radiotherapy.
 Radiotherapy and Oncology xxx (2017) xxx–xxx

Toxicity prediction following radiotherapy requires the integration of many heterogeneous variables, including patient (clinical history, age, *etc.*) and treatment (DVH, treatment techniques, *etc.*) characteristics, into predictive models. As predictive models require a large amount of data, overfitting issues may occur, such as having too many parameters for the number of events. Improving toxicity thus becomes a trade-off between including a large amount of data (gathering as much as information as possible) and not too much to cause overfitting. To overcome this issue, feature extraction/reduction strategies have recently emerged with advances in machine-learning methods. Principal component analysis (PCA) and functional PCA (FPCA) can resolve the issues of dimensionality reduction and have demonstrated effective predictive capacity for rectal toxicity in prostate cancer radiotherapy [1,2]. FPCA enables a functional DVH representation that overcomes the issues of correlation between neighboring DVH bins [2]. More specifically, PCA decomposes the data into several orthogonal bases, yielding a set of features with maximized variance. The orthogonality constraints imposed by PCA can, however, be relaxed by using more statistical information, such as mutual independence [3]. These relaxed constraints lead to the concept of independent component analysis (ICA) [4] enabling a specific observed multidimensional vector (*i.e.,* rectal DVH) to be decomposed into several components, which should be as statistically independent as possible [5]. For predicting rectal bleeding (RB) following prostate cancer radiotherapy, we thus propose using ICA to estimate two informative subspaces of patients, one with RB, one without, enabling a normalized distance (*pICA*) to be computed

* Corresponding author at: Laboratoire Traitement du Signal et de l'Image (LTSI), Université de Rennes 1, Campus de Beaulieu, Bât. 22, 35042 Cedex Rennes, France.
E-mail address: aureline.fargeas@univ-rennes1.fr (A. Fargeas).

https://doi.org/10.1016/j.radonc.2017.11.011
0167-8140/

for each of the two identified subspaces. In this way, this parameter ranging from 0 to 1 enables us to determine whether a patient is more likely to be within the RB-present group or not. The RB risk for a patient increases as the *pICA* values decrease. To the best of our knowledge, the *pICA* concept has never before been investigated as a predictor of toxicity following radiotherapy.

This study sought to evaluate the benefit of adding the new *pICA* parameters to both "standard" variables (patient/treatment characteristics) and computed variables from PCA and FPCA into a Cox proportional hazards model to predict RB following prostate cancer radiotherapy.

Materials and methods

Patients, treatment, and toxicity grading

All patient, treatment, and tumor characteristics are summarized in Table 1.

The records of 593 patients who underwent radiotherapy for localized prostate cancer were prospectively registered in three French academic institutions as part of three prospective studies: the STIC-IGRT study (clinical trial NCT00433706, accrual between June 2007 and November 2012; cut-off date: January 2015) [6], the GETUG 14 study (clinical trial NCT00104741, accrual between December 2003 and August 2008; cut-off date: March 2016), and the BDD IMRT-cancer prostate ICM study (accrual between February 2003 and October 2008; cut-off date: January 2014). Approval was granted for these studies by the appropriate ethics committee, and all patients provided informed consent in accordance with the latest Helsinki declaration. The target volume was defined as the prostate and seminal vesicles. The pelvic lymph nodes were not irradiated. The mean dose delivered to the prostate was 79.3 Gy (range: 76–80) at 2 Gy per fraction, with 46 Gy delivered to the seminal vesicles. Patients were placed in the supine position for all simulations and treatments. The target volume and organs at risk (bladder, rectum, and femoral heads) were delineated on CT slices according to French GETUG recommendations [7]. The planning target volume was calculated from a 10-mm margin all around the prostate and seminal vesicles, except in the posterior direction, where a 5-mm margin was applied. The rectal DVH complied with GETUG recommendations [7], namely fixing the volume receiving 72 Gy (V72) at < 25% and the maximum dose (within 1.8 cc) at <76 Gy.

All patients were prospectively evaluated 2 months following radiotherapy, then every 6 months thereafter. Late rectal toxicity was defined as any event occurring over 6 months after beginning radiotherapy. Rectal toxicity was scored according to the common terminology criteria for adverse events (CTCAE), Version 3.0. Patients with history of hemorrhoids were not included in the analysis. Mean follow-up was 67 months (range: 32–152). The 5-year risk of ≥Grade 2 RB was 11% (95% CI: 8.4–13.6%), calculated using the Kaplan–Meier method.

Study workflow

The overall study workflow is depicted in Fig. 1. The cohort was split into a training group, comprising 339 patients from two treatment centers, and a validation group of 254 patients treated in another center (see Table 1 for characteristics). The training group was used to compute the ICA-based subspaces that enable further specific-feature calculation from both cohorts. The validation cohort was used for evaluating the RB-prediction capability of both the "standard" parameters and the original calculated features, using the Cox proportional hazards model generated during the training step.

Parameters for toxicity prediction

The "standard" parameters considered for toxicity prediction were characteristics related to patients (age, diabetes, and anticoagulant treatment), tumors (T-stage and D'Amico risk group), and treatments (androgen-deprivation therapy [ADT], intensity-modulated radiation therapy [IMRT], image-guided radiation therapy [IGRT], and DVH). Full rectal DVHs were bin-wise analyzed with a 1-Gy step size. In addition to these variables, the following specific calculated features were extracted from rectal DVHs: our proposed parameters using ICA (*pICA*, detailed below and in the Appendix), principal components (PC) from PCA, and functional

Table 1
Patients, tumors and treatments characteristics, follow-up and rectal bleeding toxicity.

Characteristics	Training cohort ($N = 339$)	Validation cohort ($N = 254$)	Whole cohort ($N = 593$)
Patient characteristics			
Age (mean, range in year)	68 (40–89)	68 (52–82)	68 (40–89)
Anticoagulant treatment	18.9%	18.1%	18.6%
Diabetes	9.1%	6.3%	8%
Tumor characteristics			
T stages (AJCC 1992 (14))			
T1	46.6%	20.5%	35.4%
T2	42.7%	63.8%	51.8%
T3	10.7%	15.7%	12.8%
Prognostic risk groups (D'Amico (15))			
Low	0.3%	2.7%	1.4%
Intermediate	45.4%	70.9%	56.3%
High	54.3%	26.4%	42.3%
Treatment characteristics			
Total dose to the prostate (mean, rang in Gy)	79.3 (76–80)	79.4 (76–80)	79.3 (76–80)
IMRT	99.1%	62.2%	83.3%
IGRT	14.2%	54.3%	31.4%
Rectal V72 (median, range)	8.2% (0–20.7)	9.2% (0–65.2)	8.5% (0–65.2)
Androgen-deprivation therapy	70.8%	30.3%	53.5%
Follow-up and rectal bleeding toxicity			
Follow-up (mean, range in month)	71 (48–122)	63 (32–152)	67 (32–152)
5-Year Grade ≥ 2 RB rate (CI 95%)	11.7 (8.1–15.3)	9.9 (6.1–13.7)	11 (8.4–13.6)

AJCC, American Joint Committee on Cancer Staging System; IMRT, intensity-modulated radiotherapy; IGRT, image-guided radiotherapy; RB, rectal bleeding; CI, confidence interval.

Please cite this article in press as: Fargeas A et al. Independent component analysis for rectal bleeding prediction following prostate cancer radiotherapy. Radiother Oncol (2017), https://doi.org/10.1016/j.radonc.2017.11.011