

HEAD AND NECK

Development of a prognostic model of overall survival in oropharyngeal cancer from real-world data: PRO.M.E.THE.O.

Sviluppo di un modello predittivo in pazienti affetti da neoplasia dell'orofaringe: PRO.M.E.THE.O.

Francesco Miccichè¹, Giuditta Chiloire¹, Silvia Longo¹, Rosa Autorino¹, Mariangela Massaccesi¹, Jacopo Lenkowicz¹, Pierluigi Bonomo², Isacco Desideri², Liliana Belgioia^{3,4}, Almalina Bacigalupo⁴, Elisa D'Angelo⁵, Federica Bertolini⁶, Anna Merlotti⁷, Nerina Denaro⁸, Pierfrancesco Franco⁹, Francesco Bussu¹⁰, Gaetano Paludetti^{11,12}, Umberto Ricardi⁹, Vincenzo Valentini^{1,12}

¹ Diagnostica per Immagini, Radioterapia Oncologica ed Ematologia, Fondazione Policlinico Universitario A. Gemelli IRCCS, Roma, Italy; ² Department of Radiation Oncology, Azienda-Ospedaliero-Universitaria Careggi, University of Florence, Italy; ³ Radiation Oncology Department, IRCCS Ospedale Policlinico San Martino, Genoa, Italy; ⁴ Department of Health Science (DISSAL), University of Genoa, Genoa, Italy; ⁵ Radiation Oncology, University Hospital of Modena, Italy; ⁶ Medical Oncology, University Hospital of Modena, Italy; ⁷ Department of Radiation Oncology, S. Croce and Carle Teaching Hospital, Cuneo, Italy; ⁸ Department of Oncology, S. Croce and Carle Teaching Hospital, Cuneo, Italy; ⁹ Department of Oncology, Radiation Oncology, University of Turin, Turin, Italy; ¹⁰ Otolaryngology Division, Sassari University Hospital, Sassari, Italy; ¹¹ Dipartimento Scienze dell'Invecchiamento, Neurologiche, Ortopediche e della Testa-collo, UOC di Otorinolaringoiatria, Fondazione Policlinico Universitario A. Gemelli IRCCS, Roma, Italy; ¹² Università Cattolica del Sacro Cuore, Roma, Italy

SUMMARY

Objective. The PRO.M.E.THE.O. study (PredictiOn Models in Ent cancer for anti-EGFR based THERapy Optimization) aimed to develop a predictive model (PM) of overall survival (OS) for patients with locally advanced oropharyngeal cancer (LAOC) treated with radiotherapy (RT) and cetuximab (Cet) from an Italian dataset.

Methods. We enrolled patients with LAOC from 6 centres treated with RT-Cet. Clinical and treatment variables were collected. Patients were randomly divided into training (TS) (80%) and validation (VS) (20%) sets. A binary logistic regression model was used on the TS with stepwise feature selection and then on VS. Timepoints of 2, 3 and 5 years were considered. The area under the curve (AUC) of receiver operating characteristic of 2, 3 and 5 year and confusion matrix statistics at 5-threshold were used as performance criteria.

Results. Overall, 218 patients were enrolled and 174 (79.8%) were analysed. Age at diagnosis, gender, ECOG performance, clinical stage, dose to high-risk volume, overall treatment time and day of RT interruption were considered in the final PMs. The PMs were developed and represented by nomograms with AUC of 0.75, 0.73 and 0.73 for TS and 0.713, 0.713, 0.775 for VS at 2, 3 and 5 years, respectively.

Conclusions. PRO.M.E.THE.O. allows the creation of a PM for OS in patients with LAOC treated with RT-Cet.

KEY WORDS: prediction model, personalised medicine, cetuximab, head and neck tumour

RIASSUNTO

Obiettivo. L'obiettivo di PRO.M.E.THE.O. è quello di sviluppare modelli predittivi (MP) di sopravvivenza globale (OS) per pazienti (pz) con neoplasia orofaringea localmente avanzata (LAOC) sottoposti a radioterapia (RT) e Cetuximab (Cet), provenienti da un dataset italiano multicentrico.

Metodi. Sono stati arruolati pazienti con LAOC trattati con RT-Cet provenienti da 6 centri italiani, dei quali sono state raccolte variabili cliniche e di trattamento. I pazienti sono stati suddivisi casualmente in un training set (80%) e in un gruppo di controllo (20%). Per la creazione del MP è stata considerata l'OS a 2, 3 e 5 anni. L'area sottesa alla curva

Received: May 6, 2021
Accepted: October 11, 2021
Published online: April 8, 2022

Correspondence

Silvia Longo

Diagnostica per Immagini, Radioterapia Oncologica ed Ematologia, Fondazione Policlinico Universitario A. Gemelli IRCCS, largo A. Gemelli 1, 00168 Rome, Italy
Tel. +39 06 3015498
E-mail: silvia.longo@guest.policlinicogemelli.it

How to cite this article: Miccichè F, Chiloire G, Longo S, et al. Development of a prognostic model of overall survival in oropharyngeal cancer from real-world data: PRO.M.E.THE.O. *Acta Otorhinolaryngol Ital* 2022;42:205-214. <https://doi.org/10.14639/0392-100X-N1672>

© Società Italiana di Otorinolaringoiatria e Chirurgia Cervico-Facciale



OPEN ACCESS

This is an open access article distributed in accordance with the CC-BY-NC-ND (Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International) license. The article can be used by giving appropriate credit and mentioning the license, but only for non-commercial purposes and only in the original version. For further information: <https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en>

(AUC) di ROC (receiver operating characteristic) a 2, 3 e 5 anni e la statistica della matrice di confusione a 5 classi rappresentano i criteri di performance.

Risultati. 218 pazienti sono stati arruolati e 174 (79,8%) analizzati. I parametri testati per la costruzione degli MP includevano: età alla diagnosi, sesso, ECOG, stadio clinico, dose del volume ad alto rischio, tempo complessivo di trattamento e giorni di interruzione della RT. I MP di OS sono stati rappresentati con dei nomogrammi.

Conclusioni. PRO.M.E.THE.O. ha consentito la creazione di MP di OS in pazienti LAOC trattati con RT-Cet.

PAROLE CHIAVE: modello predittivo, medicina personalizzata, cetuximab, tumori della testa e del collo

Introduction

Meta-analyses on chemotherapy (CT) in locally advanced head and neck squamous cell carcinoma have demonstrated an overall survival (OS) benefit for the addition of CT to radiotherapy (RT) ¹⁻⁵. Cetuximab is considered a viable treatment option for patients who are unfit for cisplatin, and has been shown to significantly improve OS when combined with RT compared with RT alone in a randomised phase III trial ³.

Even if squamous cell carcinomas (SCC) of the head and neck region share major risk factors and some clinical features, they still have specific characteristics, different treatment options and variable prognosis, depending on the tumour site and subsite ⁶.

As new strategies and therapies are being tested, it is becoming apparent that the magnitude of benefit derived from a specific treatment, and the corresponding toxicity profile, may vary in different patient groups ⁷.

Over the last decade, remarkable advances in cancer care have raised new challenges, leading clinical practice towards personalised medicine, although there remains a gap between evidence from clinical trials and real-world practice ⁸.

Moreover, studies investigating the physicians' performance in predicting radiosensitivity and oncological outcomes are currently lacking ⁹.

Since the clinical introduction of cetuximab, the lack of biomarkers to predict its efficacy have profoundly hampered its routine clinical use. The development of tools that allow physicians to individualise treatment will facilitate transformation from population-based strategies to personalised medicine with an essential role of decision supporting systems (DSSs) ^{10,11}. The DSS development process necessarily relies on an ontology that represents knowledge as a set of shared concepts within a domain and the relationships between them, using a large amount of data with proximity to daily clinical reality ¹².

The PRO.M.E.THE.O. (PredictiOn Models in Ent cancer for anti-EGFR based THERapy Optimization) project involved several Italian RT centres to implement a system that is able to analyse large heterogeneous datasets specific for oropharyngeal SCC (OPC). The aim of this project

is to develop predictive prognostic models (PPM) of OS at 2, 3 and 5 years for OPC patients treated with RT and cetuximab (bio-radiotherapy [b-RT]) based on a real-world data collection.

Materials and methods

The PRO.M.E.THE.O. project involved 6 Italian Radiation Oncology Departments. The promoting centre was Fondazione Policlinico Universitario A. Gemelli-IRCCS, Rome. A multidisciplinary team of physicians defined the project milestones and a teleconference was scheduled every two weeks for progress updates. The objective of this first phase was to implement a system that can analyse a large heterogeneous dataset which included all available data conforming to a standardised ontology collected from daily activity. The final aim was to develop reliable PPM for OS in OPC at 2, 3 and 5 years for use in clinical management DSS.

Ontology

CREATION OF AN ONTOLOGY

The first step was the creation of an ontology, as a defined data collection model, which is capable of collecting, standardising and organising features of patients with OPC treated with cetuximab. The ontology is fundamental for the data mining process, as it will explicitly declare the clinical variables involved and their mutual relationships. Each variable has four main properties: name, form, field type and level. The existing field types are text: some variables are described with a multiple choice option and some others with free text option; number: integer or decimal number; date; table; files: DICOM and.txt files are the standard file formats chosen for images and data on treatment, respectively.

Initially, a small multidisciplinary team based in the promoter centre identified clinical variables to be included in the ontology, and all the six centres validated them. Next, a technical committee consisting of an engineer, a physicist, a physician with experience in data storage, and a software expert was created. This multiprofessional group defined the characteristics required for the ontology to be accept-

ed. These requirements included defining the data type for each field, the possible values allowed, the cardinality of the entries (i.e., single- or multi-selection field) and the range allowed in case of numeric values.

After the formal definition of the ontology and its requirements, the working group and the technical committee were asked to define the tools to share the ontology among centres via a standardised form. To accomplish this task, the “Beyond Ontology Awareness” (BOA) software was developed to reproduce the ontology structure, manage the import of legacy data and coordinate data sharing activities^{13,14}.

ONTOLOGY IMPLEMENTATION

From January 2017 to January 2018, we selected more than 200 variables across 16 input forms related to OPC. The ontology was organised into three levels:

1. registry: with exclusive epidemiologic information (age, gender, ethnicity, height, etc.);
2. procedural: where treatment information and related toxicities were reported;
3. research: where dimensional data, such as radiomics and genomics were collected.

A BOA-Web service platform was created and the centers involved collected data using Case Report Forms (CRFs). A total of 16 CRFs were created for: registry and history; blood and serum test pre-treatment and follow-up; histology; staging c, yc, p, and yp; external beam RT; CT; brachytherapy; surgery; toxicity Common Terminology Criteria for Adverse Events (CTCAE) v4.0, and Radiation Therapy Oncology Group (RTOG); follow-up and outcomes.

Standardised data sharing

A network of private connections was implemented between the various centres and a computer system that automatically translated the various available information into extractable parameters that conformed to the ontology. The data was anonymised, encrypted and sent to the central repository via a secure https-based web service. The system allowed the data to be aggregated anonymously into a single ‘large database’ to proceed with analysis of selected patient data.

Patient characteristics

We retrospectively collected data on patients with OPC treated with b-RT from 2006 to 2018 with curative intent. Inclusion criteria were: age > 18 years, Eastern Cooperative Oncology Group-Performance Status (ECOG-PS Eastern Cooperative Oncology Group) 0-2, stage I-IVa, ineligible to be treated with cisplatin (due to clinical conditions known to contraindicate the use of cisplatin according to

international consensus recommendations¹⁵). The only exclusion criteria was the presence of metastatic disease.

Cetuximab was administered at an initial loading dose of 400 mg/m² one week before initiation of RT followed by a weekly dose of 250 mg/m² for 7 weeks concurrent to RT.

Patient underwent RT with different techniques such as 3D, Intensity-modulated radiotherapy (IMRT), tomotherapy, volumetric modulated arc therapy (VMAT) with simultaneous integrated boost (SIB), or sequential RT boost. Patients underwent weekly physical evaluation during treatment. Acute toxicities, including cutaneous side effects and onset of rash, were graded according to CTCAE version 4.0. Clinical follow-up consisted of physical examination every three months for 3 years from diagnosis and every 6 months thereafter, and was performed alternatively by the radiation oncologist, otolaryngologist and medical oncologist or in joint consultation.

PM development

Table I shows the clinical and treatment variables that were collected. Categorical covariates were dummy field, while numerical covariates were kept in their original version and binarised at different cut-off thresholds as reported in Table II.

Overall survival was considered as the time elapsing between the date of diagnosis of the neoplasm and last follow-up date or death from any cause; 2-, 3- and 5-year time points were considered. For PPM development and internal validation, the dataset was randomly split into 80% and 20%¹⁶ for the training and validation sets, respectively. The covariates selected and the ones engineered from those selected were included within a multivariable logistic regression model on the training set and further selected with step-wise regression based on the Akaike Information Criteria (AIC). The model was then applied on the testing set to assess the performance in terms of Receiver Operating Characteristic (ROC) Area Under the Curve (AUC). The models were retrained for each OS time point on the entire dataset to be represented in the form of nomograms. Statistical analysis was performed using R version 3.4.4.

Results

Patient characteristics and treatment

From February to December 2018, we collected data on 218 OPC patients. We considered clinical and treatment variables, as shown in Table II. The covariates analysed were: age at diagnosis, gender, ECOG score, smoking status, alcohol consumption, human papilloma virus (HPV)-status (p16 immunohistochemistry), cT, cN, tumour grading,

Table I. Clinical and treatment variables collected from February to December 2018.

Patient characteristics	Available (%)
Gender	174 (100%)
Analysed patients/collected patients per RT centre	
Rome	77/89 (86.5%)
Modena	17/22 (77.3%)
Florence	48/50 (96%)
Genoa	16/21 (76.2%)
Cuneo	12/16 (75%)
Turin	4/20 (20%)
HPV (human papillomavirus) DNA	Not analysed
HPV (human papillomavirus) RNA	Not analysed
cT	174 (100%)
cN	174 (100%)
ECOG PS	174 (100%)
RT start date	154 (88%)
RT end date	154 (88%)
Interruption days	144 (83%)
Prescription dose to CTV	166 (95%)
Dose reached (y/n)	166 (95%)
Censor death	139 (92%)
Last follow-up date	104 (94%)
Smoking status (Pack year)	109 (63%)
Alcohol consumption	41 (23%)
Acute toxicity	Not analysed
Relapse (y/n)	Not analysed
Cetuximab number of cycles	Not analysed

RT: radiotherapy; HPV: Human Papilloma Virus; ECOG: Eastern Cooperative Oncology Group; PS: Performance Status; CTV: clinical target volume.

dose to high risk clinical target volume (CTV), fractionation, overall treatment time (OTT), days of RT interruption (DRTI). DRTI was taken as surrogate of treatment tolerance. We considered only variables with at least 75% of available values (Tab. II).

A total of 174 (79.8%) patients were analysed. All patients underwent combined b-RT with a median OTT of 51 days (range 6-101) with a median dose of 69.9 Gy (17.6-79.2 Gy). This wide dose range is due to the discontinuation of RT treatment for patients receiving less than 50 Gy.

PM development

The analysis was performed considering several variables. Categorical covariates were dummy field, whereas numerical covariates were kept in their original version and binarised at different cut-off thresholds (such as total RT dose to high risk CTV ≥ 66 Gy, age at diagnosis $>$ median value, DRTI yes/no, DRTI > 2 , > 3 , > 5 , > 7 , > 10 , > 20). HPV

status was not considered because of the not available rate $> 25\%$ (Tab. I). Covariates tested that showed a negative impact on 2-, 3- 5-year OS were cN+, ECOG ≥ 1 , age, RT dose < 66 Gy and DRTI > 3 . At a median follow-up of 57.6 months (1.5-142.0), the OS rate at 2, 3, 5 years was 71.5%, 67.4% and 64.7%, respectively.

No significant difference was found between the distributions of training and validation set covariates. The logistic regression models trained with the covariates selected from the previous stepwise selection are summarised in Table III for the three different OS time points. The performance of these models in terms of AUC is reported in Table IV, along with their respective ROC curves (Fig. 1). The nomograms built from these models are reported in Figures 2, 3 and 4 and two clinical examples are reported in Table V.

Discussion

In this study, we collected registry and procedural level variables and prognostic models were developed to predict 2-, 3- 5-year OS for patients with OPC treated with b-RT. Specifically, nomograms allowed the integration of clinical, treatment-related and epidemiological risk factors, assessing their interactions and estimating the final effect on survival. Most importantly, the PPMs provide personalised, patient-specific estimates of OS that can be used for risk-stratification and prognosis discussions with patients. Fundamental to the creation and validation of any PPM is the generation of an ontology that in this study was used as a fundamental tool to collect, standardise and organise data from OPC patients treated with cetuximab.

The collection of data from several Italian centres selected for accrual and expertise, allowed us to perform an analysis on a large population of OPC based on real-world experience. Our results are consistent with the literature, showing better survival rates than those reported in the Bonner studies^{17,18}.

The creation of a large database gave us the ability to predict which variables impact clinical outcomes. In our analyses, clinically negative nodal status ($p < 0.05$), good performance status with ECOG 0 and RT dose > 66 Gy were protective factors, while DRTI ≥ 3 appeared to be a detrimental factor for OS at 2, 3 and 5 year.

The strength of our project is related to the multicentre origin of the data and their quality due to the relative homogeneity of the population, including patients with oropharynx cancer treated with cetuximab-based RT protocols.

PRO.M.E.THE.O also has the benefit of taking into account treatment compliance, which is a key parameter in patients with OPC, where discontinuation of RT may affect the overall efficacy of treatment. It indirectly shows

Table II. Variables analysed (with at least 75% of available values).

Covariates tested:	Training set 139 patients (%)	Validation set 35 patients (%)	P-value
Age at diagnosis > 65 years	70 (50%)	13 (37%)	0.187 ^a
Gender			
Male	106 (76%)	26 (74%)	0.826 [*]
Female	33 (24%)	9 (26%)	
Nodal status			
N0	116 (83%)	29 (83%)	1 [*]
N+	23 (17%)	6 (17%)	
ECOG			
ECOG = 0	55 (39%)	20 (57%)	0.084 [*]
ECOG = 1	48 (34%)	2 (5%)	
Treatment interruption			
Yes	110 (79%)	32 (91%)	0.359 [*]
No	29 (21%)	3 (9%)	
Treatment interrupted			
Interruption days ≥ 2	95 (68%)	26 (74%)	0.544 [*]
Interruption days ≥ 3	85 (61%)	22 (62%)	1 [*]
Interruption days ≥ 5	52 (37%)	19 (54%)	0.084 [*]
Interruption days ≥ 8	40 (29%)	16 (46%)	0.068 [*]
Interruption days ≥ 10	25 (18%)	12 (34%)	0.062 [*]
Interruption days ≥ 21	8 (5%)	5 (14%)	0.140 [*]
Median OTT in days (range)	51 (6-101)	52 (41-81)	0.284 ^a
Total RT dose ≥ 66 Gy	126 (90%)	34 (97%)	0.306 ^a

OTT: Overall Treatment Time; RT: radiotherapy; ECOG: Eastern Cooperative Oncology Group. ^a Mann-Whitney test; ^{*} Fisher's Exact Test.

Table III. Logistic regression models trained with the selected covariates for the three different OS time points.

Covariate	OS model								
	2 years			3 years			5 years		
	Coefficient	Standard error	P-value	Coefficient	Standard error	P-value	Coefficient	Standard error	P-value
N = 0	-1.0439	0.6260	0.095	-0.8863	0.5702	0.120	-0.8863	0.5702	0.120
ECOG PS = 0	-1.3410	0.4634	0.003 ^{**}	-0.9209	0.4129	0.025 [*]	-0.9209	0.4129	0.025 [*]
RT Dose ≥ 66 Gray	-2.0091	0.6714	0.002 ^{**}	-1.7503	0.6499	0.007 ^{**}	-1.7503	0.6499	0.007 ^{**}
RT Interruption days ≥ 3	0.9058	0.4405	0.039 [*]	1.0223	0.4165	0.014 [*]	1.0223	0.4165	0.014 [*]

OS: overall survival; ECOG PS: Eastern Cooperative Oncology Group Performance Status; RT: radiotherapy. ^{*}: p-value < 0.05; ^{**}: p-value < 0.01.

Table IV. Performance of logistic regression models in terms of receiver operating characteristic (ROC) of area under the curve (AUC).

OS time point	AUC training set	AUC validation set
2 years	0.750	0.713
3 years	0.731	0.713
5 years	0.729	0.775

OS: overall survival.

compliance with cetuximab, underscoring its importance³. Unfortunately, we do not have clear data about the main reasons for treatment interruptions, which would have been interesting and could be a subject of future research. HPV status is a major determinant of the prognosis of patients with OPC with a 60% reduction in the risk of death. Data from randomised trials highlight the impact of HPV status on survival, demonstrating a 60% improvement in HPV-positive patients.

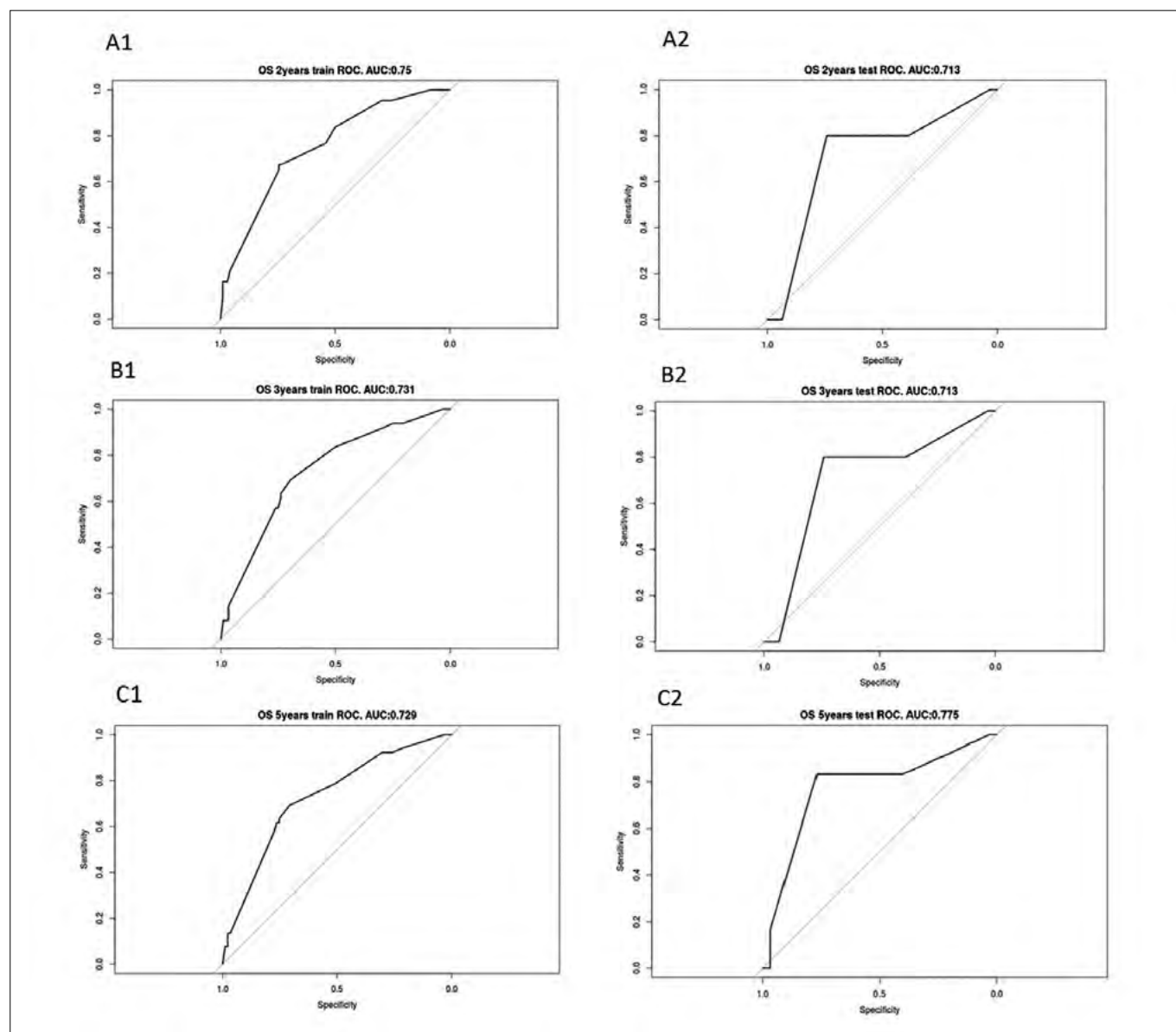


Figure 1. The performances of training (1) and validation (2) model at 2 (A), 3 (B) and 5 (C) year OS in terms of receiver operating characteristic (ROC) curves.

Our retrospective observational case series on a large amount of data lacks HPV status because in the period analysed HPV determination was not part of routine clinical practice; nevertheless, the model provided has highlighted prognostic factors for survival that in our opinion may represent an essential integration in the prognostic framing of patients with OPC.

There is heterogeneity in HPV positive patients due to tumour stage, smoking status and other prognostic factors such as radiological extranodal extension, matted nodes and PIK3CA¹⁹. Therefore, even if the integration of HPV status in the nomogram for the clinical use is absolutely

recommended, in this heterogeneous setting, our models, which overcome HPV-status, could be useful as well. In addition, the benefit of the present work is that it is based on a real-world data collection and reflects the fact that in many Italian centres the evaluation of HPV in OPC is still not standardised nor routinely performed.

Nonetheless, our results provide evidence of the feasibility of model building on real-world data also considering that they are in line with the literature, even without considering HPV-status. In fact, our models showed a good predictive power with an AUC between 0.73 and 0.75.

Some sources of big data already exist in the literature^{10,20},

Table V. Clinical examples of the nomograms prediction in terms of overall survival modulated in relation to risk factors.

	Age at diagnosis	N0	ECOG 0	RT Dose ≥ 66 Gy	Interruption RT days ≥ 3	Death risk at 2 y (%)	
Very Good OS > 90% at 2 y	65	N-	Yes	Yes	No	2	
	65	N-	No	No	No	3	
	75	N-	Yes	Yes	No	3	
	65	N-	No	Yes	No	4	
	65	N-	Yes	Yes	Yes	5	
	75	N-	No	Yes	No	6	
	65	N+	Yes	Yes	No	7	
	75	N-	Yes	Yes	Yes	7	
Good OS > 75-85% at 2 y	65	N-	No	Yes	Yes	11	
	75	N+	Yes	Yes	No	11	
	75	N-	No	Yes	Yes	15	
	65	N+	No	Yes	No	16	
	65	N-	Yes	No	No	16	
	65	N+	Yes	Yes	Yes	17	
	75	N+	No	Yes	No	21	
	75	N-	Yes	No	No	22	
	75	N+	Yes	Yes	Yes	24	
	Poor OS > 50-70% at 2 y	65	N+	No	Yes	Yes	33
65		N-	Yes	No	Yes	33	
75		N-	No	No	No	39	
75		N+	No	Yes	Yes	42	
75		N-	Yes	No	Yes	42	
65		N+	Yes	No	No	43	
65		N-	No	No	Yes	53	
75		N+	Yes	No	No	53	
Very Poor OS > 15-40% at 2 y		65	N+	No	No	No	63
		75	N-	No	No	Yes	63
	65	N+	Yes	No	Yes	66	
	75	N+	No	No	No	72	
	75	N+	Yes	No	Yes	74	
	65	N+	No	No	Yes	82	
	75	N+	No	No	Yes	87	

OS: overall survival; ECOG PS: Eastern Cooperative Oncology Group Performance Status; RT: radiotherapy. Covariates with a negative impact are in bold.

but are constrained by important limitations, including low granularity (i.e. lack of detailed information on RT). Available big data sources are usually at least 2-3 years behind current practice due to the time required to collect and assemble the data and perform quality control. Adding the time for data analysis and publication, current studies examining quality of care and comparative effectiveness usually report data that is at least 5 years old^{22,23}. However, to our knowledge, the

reliability, consistency and accuracy of these calculators in predicting outcomes for individual patients in different populations remain unclear. These uncertainties are critical in order to optimally implement these tools in clinical practice²⁴.

The present study has some limitations. First, the absence of external validation and analysis for variables such as HPV-status, smoking and alcohol habits that may have an impact on OS. However, considering the results of De-Es-

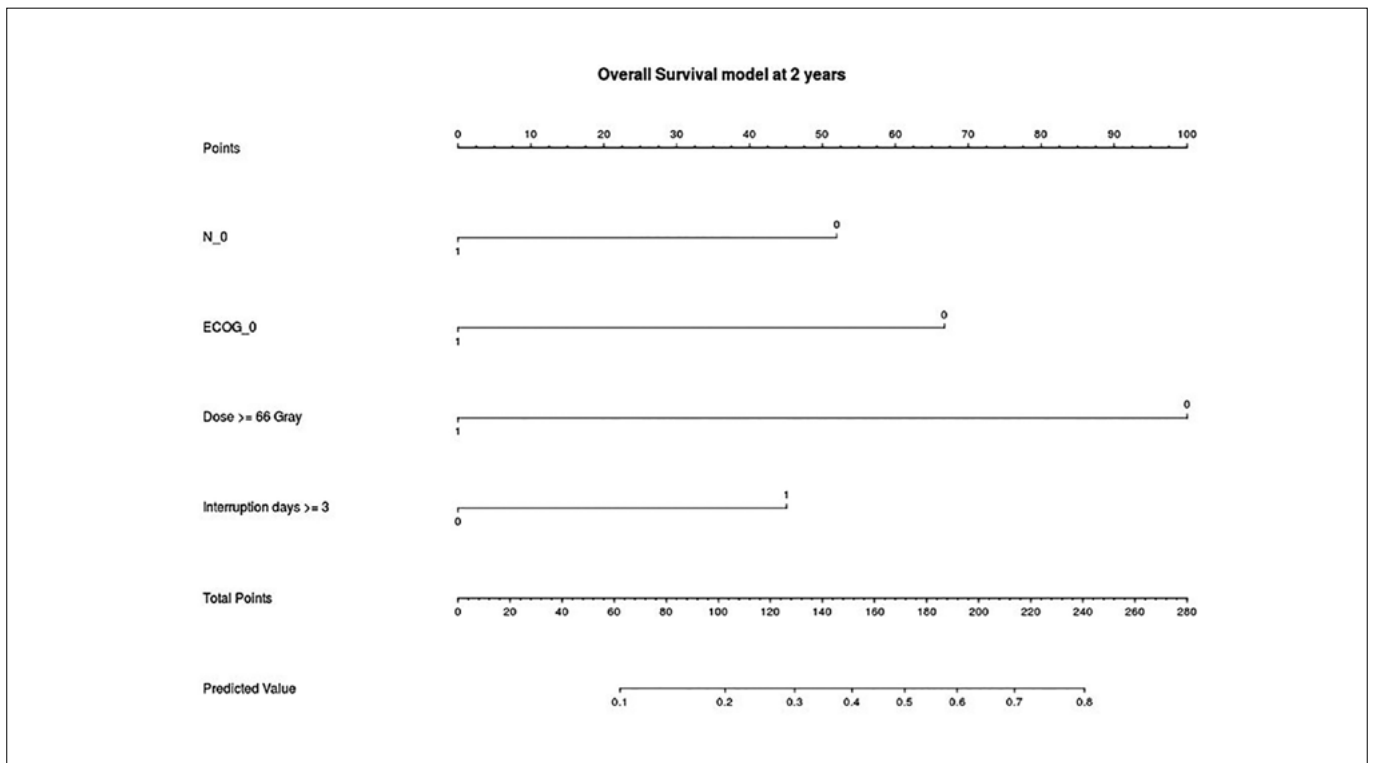


Figure 2. The nomogram built from the OS model at 2 years.

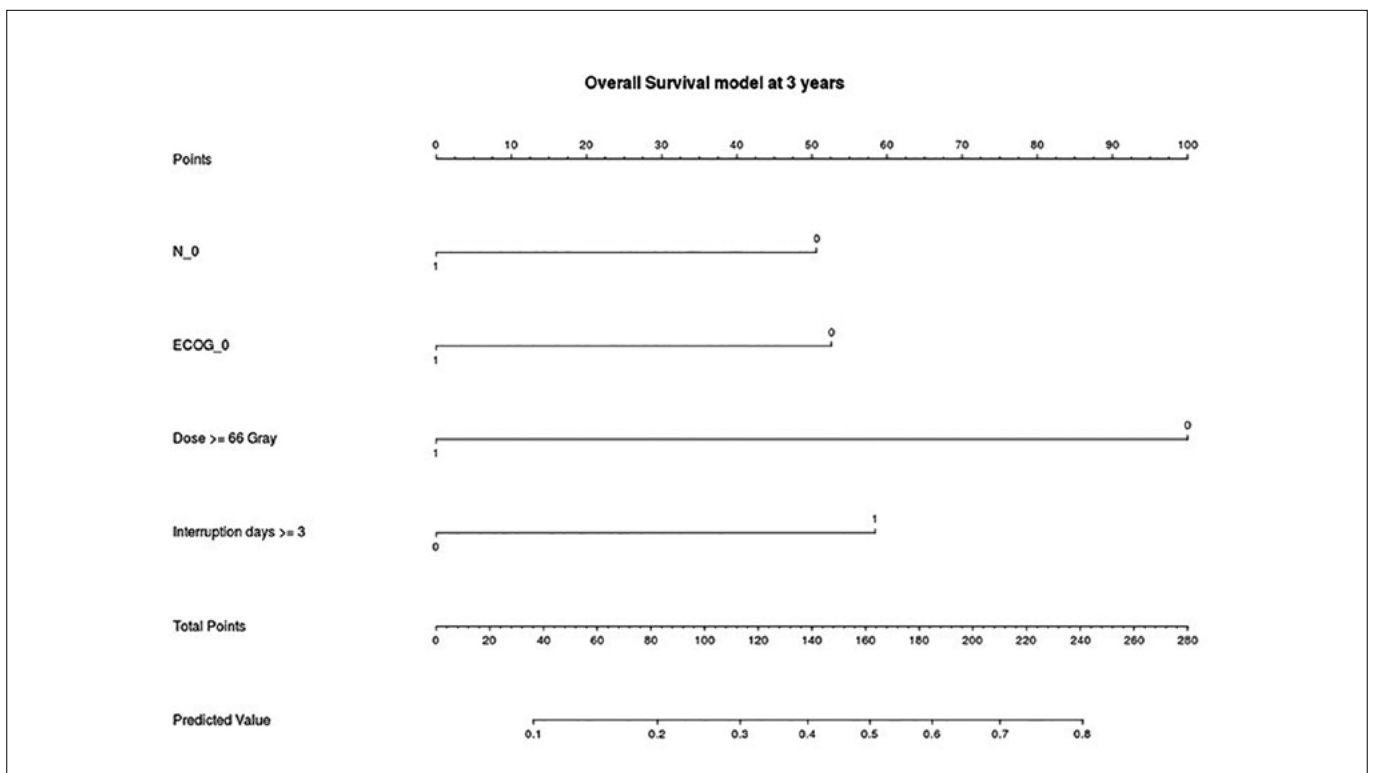


Figure 3. The nomogram built from the OS model at 3 years.

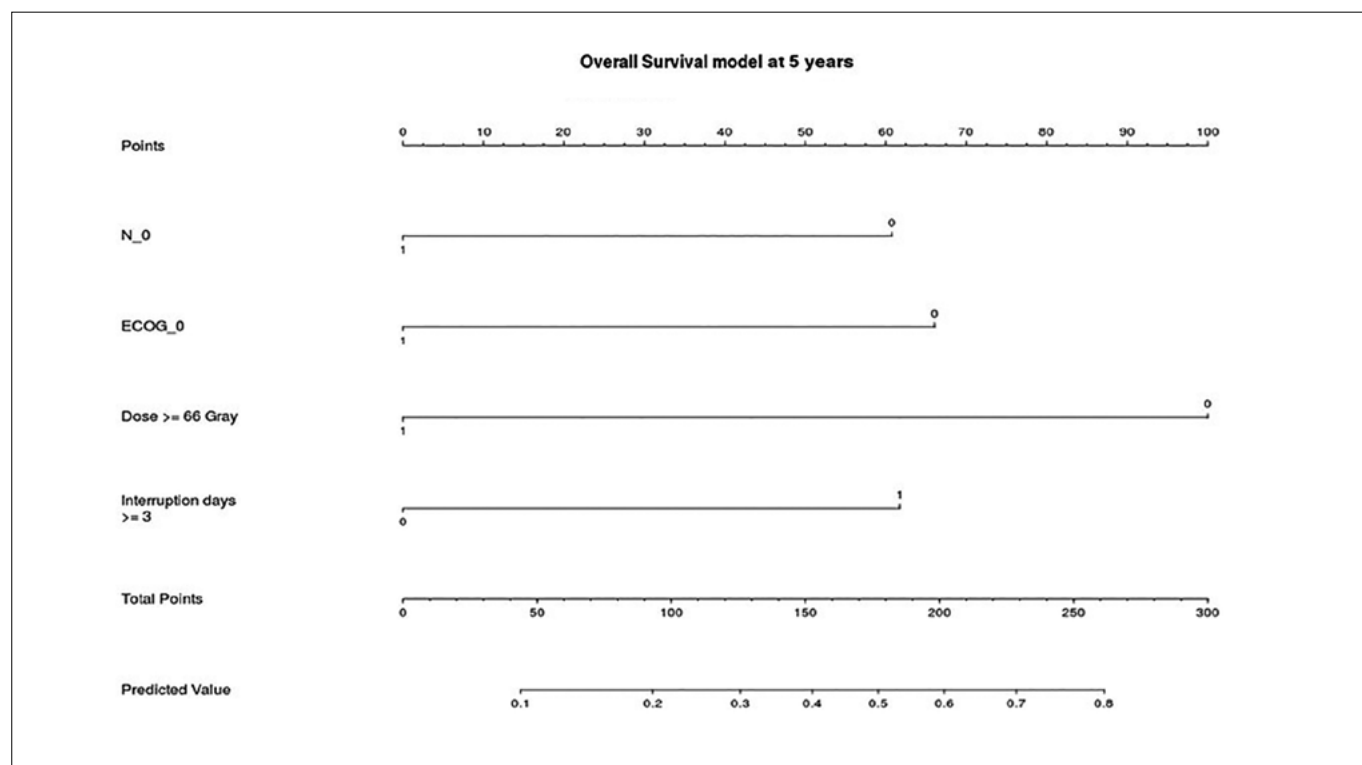


Figure 4. The nomogram built from the OS model at 5 years.

calate⁴ and RTOG 1016⁵ trials, it is important to create a model that bypasses HPV-status and investigate the impact of other factors on survival. Another limitation is that we did not study adherence to cetuximab therapy, as we did not consider cetuximab dose intensity as a potential variable influencing outcome. We also did not specifically address the impact of toxicity, as we assumed that toxicity-induced RT discontinuations were too rare to have an impact on OS. However, we are planning to investigate this aspect in future analyses, where it may be of interest to evaluate the impact of cetuximab dose intensity and cumulative dose on skin toxicity and OS.

Conclusions

This project represents the first example of a PPM created specifically for OPC patients treated with b-RT based on real-world data. The next step of this project will be external validation of the PPM. We would also like to implement our analysis with HPV, smoking status, alcohol consumption, details about toxicity (and in particular compliance to cetuximab treatment) and radiomics data. Schematisations have been shown to significantly enhance physician-patient communication, and nomograms provide a visual picture of prognostic factors and their relative influence to support

personalised treatment. This could facilitate active participation by patients in decision-making.

Conflict of interest statement

The authors declare no conflict of interest.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' contributions

All authors have equally contributed to the manuscript.

Ethical consideration

This study was approved by the Institutional Ethics Committee of Catholic University of the Sacred Heart (approval number/protocol number 0049309/16). The research was conducted ethically, with all study procedures being performed in accordance with the requirements of the World Medical Association's Declaration of Helsinki. Written informed consent was obtained from each participant/patient for study participation and data publication.

The local ethics committees of each centre approved the protocol before patient enrolment according to Italian legislation. A Collaboration and Data Transfer Agreement was signed

to define the type of data, permitted use and data protection policies.

References

- 1 Lacas B, Carmel A, Landais C, et al. Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): an update on 107 randomized trials and 19,805 patients, on behalf of MACH-NC Group. *Radiother Oncol* 2021;156:281-293. <https://doi.org/10.1016/j.radonc.2021.01.013>
- 2 Blanchard P, Baujat B, Holostenco V, et al. Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): a comprehensive analysis by tumour site. *Radiother Oncol* 2011;100:33-40. <https://doi.org/10.1016/j.radonc.2011.05.036>
- 3 Magrini SM, Buglione M, Corvò R, et al. Cetuximab and radiotherapy versus cisplatin and radiotherapy for locally advanced head and neck cancer: a randomized phase II trial. *J Clin Oncol* 2016;34:427-435. <https://doi.org/10.1200/JCO.2015.63.1671>
- 4 Mehanna H, Robinson M, Hartley A, et al. Radiotherapy plus cisplatin or cetuximab in low-risk human papillomavirus-positive oropharyngeal cancer (De-ESCALaTE HPV): an open-label randomised controlled phase 3 trial. *Lancet* 2019;393:51-60. [https://doi.org/10.1016/S0140-6736\(18\)32752-1](https://doi.org/10.1016/S0140-6736(18)32752-1)
- 5 Gillison ML, Trotti AM, Harris J, et al. Radiotherapy plus cetuximab or cisplatin in human papillomavirus-positive oropharyngeal cancer (NRG Oncology RTOG 1016): a randomised, multicentre, non-inferiority trial. *Lancet* 2019;393:40-50. [https://doi.org/10.1016/S0140-6736\(18\)32779-X](https://doi.org/10.1016/S0140-6736(18)32779-X)
- 6 Rios Velazquez E, Hoebbers F, Aerts HJWL, et al. Externally validated HPV-based prognostic nomogram for oropharyngeal carcinoma patients yields more accurate predictions than TNM staging. *Radiother Oncol* 2014;113:324-330. <https://doi.org/10.1016/j.radonc.2014.09.005>
- 7 Bentzen SM, Hendry JH. Variability in the radiosensitivity of normal cells and tissues. Report from a workshop organised by the European Society for Therapeutic Radiology and Oncology in Edinburgh, UK, 19 September 1998. *Int J Radiat Biol* 1999;75:513-517. <https://doi.org/10.1080/095530099140447>
- 8 van Baardwijk A, Wanders S, Boersma L, et al. Mature results of an individualized radiation dose prescription study based on normal tissue constraints in stages I to III non-small-cell lung cancer. *J Clin Oncol* 2010;28:1380-1386. <https://doi.org/10.1200/JCO.2009.24.7221>
- 9 Scott JG, Berglund A, Schell MJ, et al. A genome-based model for adjusting radiotherapy dose (GARD): a retrospective, cohort-based study. *Lancet Oncol* 2017;18:202-211. [https://doi.org/10.1016/S1470-2045\(16\)30648-9](https://doi.org/10.1016/S1470-2045(16)30648-9)
- 10 Meldolesi E, Van Soest J, Dinapoli N, et al. An umbrella protocol for standardized data collection (SDC) in rectal cancer: a prospective uniform naming and procedure convention to support personalized medicine. *Radiother Oncol* 2014;112:59-62. <https://doi.org/10.1016/j.radonc.2014.04.008>
- 11 Gambacorta MA, Valentini C, Dinapoli N, et al. Clinical validation of atlas-based auto-segmentation of pelvic volumes and normal tissue in rectal tumors using auto-segmentation computed system. *Acta Oncol* 2013;52:1676-1681. <https://doi.org/10.3109/0284186X.2012.754989>
- 12 van Stiphout RGPM, Lammering G, Buijsen J, et al. Development and external validation of a predictive model for pathological complete response of rectal cancer patients including sequential PET-CT imaging. *Radiother Oncol* 2011;98:126-133. <https://doi.org/10.1016/j.radonc.2010.12.002>
- 13 Tagliaferri L, Kovács G, Autorino R, et al. ENT COBRA (Consortium for Brachytherapy Data Analysis): interdisciplinary standardized data collection system for head and neck patients treated with interventional radiotherapy (brachytherapy). *J Contemp Brachytherapy* 2016;8:336-343. <https://doi.org/10.5114/jcb.2016.61958>
- 14 Meldolesi E, van Soest J, Dinapoli N, et al. Medicine is a science of uncertainty and an art of probability (Sir W. Osler). *Radiother Oncol* 2015;114:132-134. <https://doi.org/10.1016/j.radonc.2014.12.013>
- 15 Porceddu S V, Scotté F, Aapro M, et al. Treating patients with locally advanced squamous cell carcinoma of the head and neck unsuitable to receive cisplatin-based therapy. *Front Oncol* 2020;9:1522. <https://doi.org/10.3389/fonc.2019.01522>
- 16 Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ* 2015;350:g7594. <https://doi.org/10.1136/bmj.g7594>
- 17 Bonner JA, Harari PM, Giralt J, et al. Radiotherapy plus cetuximab for squamous-cell carcinoma of the head and neck. *N Engl J Med* 2006;354:567-578. <https://doi.org/10.1056/NEJMoa053422>
- 18 Bonner JA, Harari PM, Giralt J, et al. Radiotherapy plus cetuximab for locoregionally advanced head and neck cancer: 5-year survival data from a phase 3 randomised trial, and relation between cetuximab-induced rash and survival. *Lancet Oncol* 2010;11:21-28. [https://doi.org/10.1016/S1470-2045\(09\)70311-0](https://doi.org/10.1016/S1470-2045(09)70311-0)
- 19 Beaty BT, Moon DH, Shen CJ, et al. PIK3CA mutation in HPV-associated OPSCC patients receiving deintensified chemoradiation. *J Natl Cancer Inst* 2020;112:855-858. <https://doi.org/10.1093/jnci/djz224>
- 20 Tagliaferri L, Budrukkar A, Lenkovicz J, et al. Review papers ENT COBRA ONTOLOGY: the covariates classification system proposed by the Head & Neck and Skin GEC-ESTRO Working Group for interdisciplinary standardized data collection in head and neck patient cohorts treated with interventional radiotherapy (brachytherapy). *J Contemp Brachytherapy* 2018;10:260-266. <https://doi.org/10.5114/jcb.2018.76982>
- 21 Fakhry C, Zhang Q, Nguyen-Tân PF, et al. Development and validation of nomograms predictive of overall and progression-free survival in patients with oropharyngeal cancer. *J Clin Oncol* 2017;35:4057-4065. <https://doi.org/10.1200/JCO.2016.72.0748>
- 22 Pagedar NA, Chioreso C, Schlichting JA, et al. Treatment selection in oropharyngeal cancer: a surveillance, epidemiology, and end results (SEER) patterns of care analysis. *Cancer Causes Control* 2017;28:1085-1093. <https://doi.org/10.1007/s10552-017-0938-3>
- 23 Hararah MK, Stokes WA, Jones BL, et al. Nomogram for preoperative prediction of nodal extracapsular extension or positive surgical margins in oropharyngeal squamous cell carcinoma. *Oral Oncol* 2018;83:73-80. <https://doi.org/10.1016/j.oraloncology.2018.06.005>
- 24 Beesley LJ, Hawkins PG, Amlani LM, et al. Individualized survival prediction for patients with oropharyngeal cancer in the human papillomavirus era. *Cancer* 2019;125:68-78. <https://doi.org/10.1002/cncr.31739>