Different prognostic models for different patient populations: validation of a new prognostic model for patients with oropharyngeal cancer in Western Europe

M M Rietbergen[1], B I Witte[2], E R Velazquez[3], P J F Snijders[4], E Bloemena[5], E J Speel[6], R H Brakenhoff*,[1], B Kremer[7], P Lambin[3] and C R Leemans[1]

[1]Department of Otolaryngology/Head and Neck Surgery, VU University Medical Center, Amsterdam, The Netherlands; [2]Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, The Netherlands; [3]Department of Radiation Oncology (MAASTRO), GROW–School for Oncology and Developmental Biology, Maastricht University Medical Center (MUMC+), Maastricht, The Netherlands; [4]Department of Pathology, VU University Medical Center, Amsterdam, The Netherlands; [5]Department of Oral and Maxillofacial Surgery/Oral Pathology, VU University Medical Center/Academic Center for Dentistry Amsterdam (ACTA), Amsterdam, The Netherlands; [6]Department of Pathology, Maastricht University Medical Center, Maastricht, The Netherlands and [7]Department of Otolaryngology/Head and Neck Surgery, Maastricht University Medical Center, Maastricht, The Netherlands

Objective: The presence of human papillomavirus (HPV) infection in oropharyngeal squamous cell carcinoma (OPSCC) is a major determinant in prognostic risk modelling. Recently, a prognostic model was proposed in which HPV status, comorbidity and nodal stage were the most important prognostic factors to determine high-, intermediate- and low-risk survival groups. Here, we report on the validation of this model using an independent single-institutional cohort.

Methods: A total number of 235 patients curatively treated for OPSCC in the period 2000–2011 at the MUMC (Maastricht University Medical Center, The Netherlands) were included. The presence of an oncogenic HPV infection was determined by p16 immunostaining, followed by a high-risk HPV DNA PCR on the p16-positive cases. The model variables included were HPV status, comorbidity and nodal stage. As a measure of model performance, the Harrell's Concordance index (Harrell's C-index) was used.

Results: The 5-year overall survival (OS) estimates were 84.6%, 54.5% and 28.7% in the low-, intermediate- and high-risk group, respectively. The difference between the survival curves was highly significant ($P < 0.001$). The Harrell's C-index was 0.69 (95% confidence interval (CI): 0.63–0.75).

Conclusion: In this study a previously developed prognostic risk model was validated. This model will help to personalise treatment in OPSCC patients. This model is publicly available at www.predictcancer.org.

Over the past three decades, it has become clear that infection with high-risk human papillomavirus (HPV) is aetiologically linked to the development of head and neck squamous cell carcinomas, particularly those that arise in the oropharyngeal region.

Epidemiologic evidence has shown a rapid increase in the prevalence rates of HPV-induced oropharyngeal squamous cell carcinomas (OPSCCs) in Europe and the United States (Chaturvedi et al, 2011; Shaw and Robinson, 2011; Rietbergen

et al, 2013b). The HPV-associated OPSCCs are considered to be a different tumour entity, based on biological, epidemiological and clinical differences, compared with the HPV-negative OPSCCs. The most important clinical difference between patients with HPV-positive and HPV-negative OPSCCs is related to the prognosis. Several retrospective and prospective studies in the United States, Australia and Western Europe have consistently demonstrated that HPV-positive OPSCC is associated with a more favourable prognosis (Fakhry et al, 2008; Ang et al, 2010; Rischin et al, 2010; Posner et al, 2011). The presence of HPV in OPSCC now has become a major determinant in prognostic risk modelling that, in turn, is very important for individualised treatment (Lambin et al, 2013a,b).

Recently, a recursive partitioning model (RPA) for patients with OPSCC has been proposed, based on the Radiation Therapy Oncology Group study (RTOG 0129 study), with HPV as first stratification parameter next to smoking and stage (Ang et al, 2010). However, this model was based on a clinical trial population, and patients with severe comorbidity or early stage of disease were not included. Moreover, cigarette smoking habits in Europe differ from those in the United States. Therefore, we generated an adapted prognostic risk model that might be applicable for the entire population of patients who present with OPSCC (Rietbergen et al, 2013a). In this prognostic risk model, HPV status, comorbidity and nodal stage were the most important prognostic factors. However, it was not validated on an independent cohort. Here, we report on the validation of this model using an independent single-institutional cohort and compared model performance with that of the RTOG 0129 model.

PATIENTS AND METHODS

Study design and HPV testing. The independent external validation cohort consisted of 235 patients curatively treated for OPSCC in the period 2000–2011 at Maastricht University Medical Center, The Netherlands. The presence of HPV was detected using pretreatment formalin-fixed, paraffin-embedded (FFPE) biopsies. Eligible samples included histopathologically confirmed invasive squamous cell carcinoma of the oropharynx (International classification of diseases for Oncology [ICD-10] codes C019, C051, C052, C090–C099 and C100–C109). A sample was scored as HPV-positive based on a positive p16 immunostaining and a subsequent positive GP $5+/6+$ HPV DNA PCR, according to our test algorithm that has a validated accuracy of 98% (Smeets et al, 2007; Rietbergen et al, 2013b).

Patient characteristics, information on smoking behaviour (one pack year = 20 cigarettes a day during 1 year) and alcohol consumption (one unit year = one drink a day during 1 year) as well as clinical outcome were obtained from the patient files. Comorbidity was classified according to the Adult Comorbidity Evaluation 27 (ACE-27) index calculator (http://oto2.wustl.edu/clinepi/calc.html) that divides comorbidity into three categories: mild, moderate and severe. The ACE-27 index is a comorbidity classification system based on the Kaplan–Feinstein Comorbidity index (Kaplan and Feinstein, 1974) and was proven to be of prognostic value (Kallogjeri et al, 2012).

Patient characteristics of the validation cohort are shown in Table 1.

Statistical methods. Differences in patient characteristics between the external validation cohort and the internal data set on which our prognostic model was based (the VUmc/EMC cohort as described in Rietbergen et al, 2013a) were assessed using Pearson's χ^2 test or Student's t-test. Bonferroni correction was used to compare subgroups for specific variables, and Mantel–Haenszel test was used to compare both cohorts stratified by HPV status.

Table 1. Patient characteristics

	Cohort VUmc/EMC	Validation cohort	P-value
	Number (%)	Number (%)	
Patient characteristics			
No. of cases	723	235	
Age at diagnosis			$P=0.93$[a]
Mean	60.54	60.24	
Median	60.0	58.96	
Gender			$P=0.047$[b]
Male	482 (66.7%)	173 (73.6%)	
Female	241 (33.3%)	62 (26.4%)	
HPV			$P=0.004$[b]
Positive	152 (21.0%)	71 (30.2%)	
Negative	571 (79.0)	164 (69.8%)	
Comorbidity (ACE score)			$P=0.18$[b]
0–1	511 (70.7%)	169 (71.9%)	
2–3	210 (29.0%)	57 (24.2%)	
Unknown	2 (0.3%)	9 (3.8%)	
Smoking			$P=0.03$[b] $P=0.33$[c]
0–10 Pack years	92 (12.7%)	44 (18.7%)	
>10 Pack years	625 (86.4%)	191 (81.3%)	
Unknown	6 (0.8%)	0	
T-stage			$P=0.44$[b]
T1–2	344 (47.6%)	121 (51.5%)	
T3–4	377 (52.1%)	114 (48.5%)	
Tx	2 (0.3%)	0	
N-stage			$P=0.03$[b] $P=0.05$[c]
N0–N2a	416 (57.5%)	113 (48.1%)	
N2b–N3	306 (42.3%)	122 (51.9%)	
Nx	1 (0.1%)	0	
Stage of disease			$P=0.29$[b]
I–II	171 (23.8%)	48 (20.4%)	
III–IV	549 (76.2%)	187 (79.6%)	
Treatment modalities			$P<0.001$[b]
SURG ± RT	215 (29.9%)	71 (30.2%)	
RT	208 (28.9%)	108 (46.0%)	
CRT	202 (28.1%)	45 (19.1%)	
Other	94 (13.1%)	11 (4.7%)	

Abbreviations: ACE = Adult Comorbidity Evaluation; CRT = chemoradiotherapy; EMC = Erasmus University Medical Center; HPV = human papillomavirus; RT = radiotherapy; SURG = surgery; VUmc = VU University Medical Center.
[a]As defined by the independent t-test.
[b]As defined by the χ^2 test.
[c]As defined by the Mantel–Haenszel test.

All 235 patients were stratified into three risk groups according to the prognostic model described by Rietbergen et al (2013a) (i.e., HPV status, comorbidity, nodal stage and tumour stage). With the purpose of externally validating our prognostic model, the Kaplan–Meier group-stratified overall survival (OS) and progression-free survival (PFS) curves were estimated. Comparison between the curves was carried out using the log-rank test. Overall survival was defined as the time from date of incidence (defined as the date on which the squamous cell carcinoma was histologically confirmed) to death (any cause). Progression-free survival was defined as the time period from date of incidence to death or the first documented relapse that was categorised as local–regional recurrence or distant metastases. Patients who developed a second primary tumour were censored at the incidence date of the second primary tumour.

As a measure of model performance, the Harrell's Concordance index (Harrell's C-index) was used (Harrell et al, 1996). The maximum value of the Harrell's C-index is 1.0, indicating a perfect prediction model. A value of 0.5 indicates that 50% of the patients

are correctly classified (i.e., as good as a chance). The RPA model based on the RTOG 0129 study (the RTOG RPA model) (Ang et al, 2010) was independently validated as well, including model performance analysis using the Harrell's C-index.

RESULTS

The baseline patient and tumour characteristics of the independent external validation cohort compared with the VUmc/EMC cohort are shown in Table 1. The groups differed significantly regarding the proportion of HPV-positive cases ($P = 0.004$), smoking and nodal stage (both $P = 0.03$). However, after stratification for HPV status (HPV-positive fraction was higher in the independent validation series; 30.2% vs 21.0%), the smoking distribution did not differ significantly, and comorbidity distribution remained only slightly significant ($P = 0.049$). There was no significant difference in survival between the different treatment groups for either HPV-negative patients ($P = 0.37$) or HPV-positive patients ($P = 0.1$).

Figure 1 shows the distribution of the external validation cohort according to our prognostic risk model (Rietbergen et al, 2013a). Survival curves of the low-risk group ($n = 78$), intermediate-risk group ($n = 86$) and high-risk group ($n = 103$) are depicted. Group-stratified Kaplan–Meier OS and PFS curves were estimated and compared using the log-rank test. The 5-year OS estimates were 84.6%, 54.3% and 28.7%, respectively. The difference between the survival curves was highly significant ($P < 0.001$). The Harrell's C-index was 0.69 (95% confidence interval (CI): 0.63–0.75). For comparison, the Harrell's C-index as measurement of model performance when using HPV status as single variable was 0.37 (95% CI: 0.32–0.42) and for nodal stage was 0.55 (95% CI: 0.49–0.60). The 5-year PFS estimates were 86.9%, 78.0% and 56.1%, respectively, with the Harrell's C-index of 0.66 (95% CI: 0.59–0.74).

The patients in the external validation cohort were also stratified according to the RPA model based on the RTOG 0129 study (Figure 2). This yielded Harrell's C-index of 0.65 (95% CI: 0.59–0.70) for OS and 0.61 (95% CI: 0.54–0.68) for PFS. The shift in distribution of case numbers in the different risk groups is remarkable (see also Figure 2); 43 of 235 (18%) patients belonged to the low-risk group according to the RTOG 0129 model as compared with 73 of 235 (31%) patients according to our prognostic model. The patient shift between the RTOG 0129 model and our prognostic model is depicted in Table 2. The group of HPV-positive patients having a low ACE-27 score but smoking more than 10 pack years ($n = 21$) would have belonged to the

'intermediate' risk group according to the RTOG 0129 model, while they actually had a favorable 5-year survival rate (89.7%).

DISCUSSION

The survival of patients with OPSCC may differ significantly between individuals, mainly related to HPV status (Ang et al, 2010; Rietbergen et al, 2013a). Consequently, the ability to estimate survival probability of OPSCC patients before any type of treatment would be very valuable for decision making, especially for patients who might be enrolled in treatment de-escalation trials (the 'low-risk' group). The prognostic model we previously developed and validated in this study seems to be applicable for the entire patient population who presents in daily practice, also including patients with early stage of disease and severe comorbidity. Moreover, this model might be very suitable for a patient population with a high percentage of heavy smokers, as is the case in most European patient populations. This is partially shown by the shift in distribution between the two low-risk groups in the RTOG 0129 model and our prognostic model; the group of HPV-positive patients with a low comorbidity score but smoking more than 10 pack years would have belonged to the 'intermediate' risk group according to the RTOG 0129 model, while they actually had a good 5-year survival rate. This finding also demonstrates that eventually more patients could be included in clinical de-escalation trials (i.e., HPV patients who smoked > 10 pack years but without comorbidity). The difference between the models and the shift in case distribution also poses a warning when using these models in clinical trial designs. The population on which the model has been developed should be more or less identical with the population that is enrolled in clinical trials, or the model should be separately
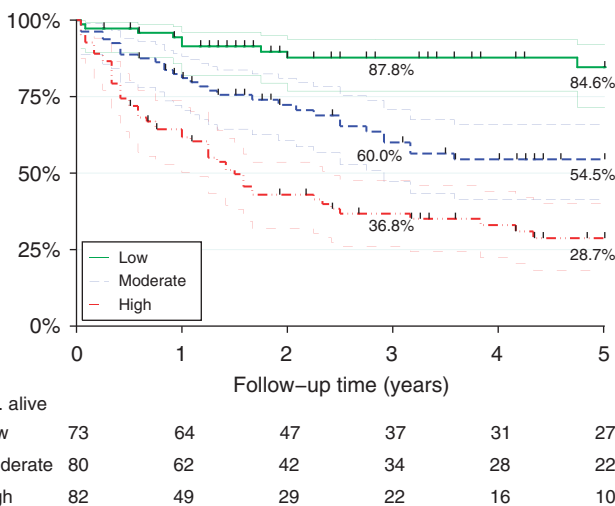

Figure 2. Kaplan–Meier overall survival curves with 95% CIs of the three risk groups of the RTOG 0129 study.

	No. alive					
Low	43	38	26	19	15	11
Moderate	40	35	27	22	19	18
High	152	102	65	52	41	30

Figure 1. Kaplan–Meier overall survival curves with 95% CIs of the three risk groups according to our prognostic model.
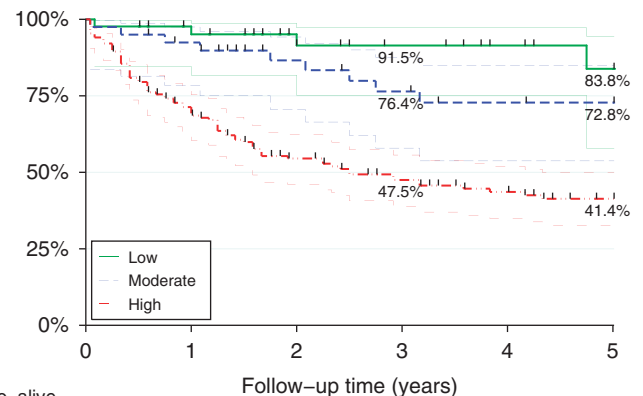
No. alive						
Low	73	64	47	37	31	27
Moderate	80	62	42	34	28	22
High	82	49	29	22	16	10

Table 2. Number of cases in the three risk catergories according to the RTOG 0129 model and the prognostic model as defined by Rietbergen et al (2013a)

	Rietbergen et al prognostic model		
RTOG 0129	Low risk	Intermediate risk	High risk
Low risk	37	6	0
Intermediate risk	21	16	3
High risk	15	58	79
Abbreviation: RTOG = Radiation Therapy Oncology Group study.			

validated. Whether our model or the RTOG 0129 model is most suited for patient selection in clinical de-escalation trials remains to be determined. Patients selected for these trials are typically stage III/IV, scheduled for chemoradiation and with good performance status. The current group of 235 patients was not large enough to focus analyses on these specific cases.

Moreover, in the context of broadening the HPV-positive population suitable for de-escalation trials, the group of HPV-positive patients with N2c nodal stage should also be considered. O'Sullivan et al (2013) showed that this group of patients have a reduced distant control when treated with radiotherapy alone and seem less suited for deintensification strategies that omit chemotherapy. As the number of HPV-positive patients with a high nodal stage in this study was small ($n = 9$) and most patients received chemoradiotherapy, we cannot draw definite conclusions on this. In future studies we hope to further evaluate this important group of HPV-positive patients.

The difference in proportion of HPV-positive cases between the two Dutch cohorts might relate to the time period from which cases have been selected; the patients of the independent validation cohort were selected from the period 2000–2011, whereas the patients of the previous VUmc/EMC cohort were selected from the period 2000–2006. HPV prevalence rates in oropharyngeal tumours have been increasing significantly in the past two decades. We showed a continuous increase in the proportion of HPV-positive OPSCC from 5% in 1990 to 29% in 2010 at our institute (Rietbergen et al, 2013b).

This study has certain limitations. First, the retrospective data collection might have caused a possible lack of accuracy in the smoking and comorbidity data. Second, the heterogeneity of the treatment given, and the high proportion of patients treated with radiation alone, might have affected outcomes. However, there was no significant difference in survival between the different treatment groups for either HPV-negative patients ($P = 0.37$) or HPV-positive patients ($P = 0.1$).

In conclusion, we validated our previously defined prognostic model with an independent series of patients. The group consisted of patients with all different stages of disease and patients with a low comorbidity as well as a high comorbidity score. This model could be used for stratification in clinical trials, especially in patient populations with a high percentage of heavy smokers. The model will be publicly available on the website www.predictcancer.org after publication.

REFERENCES

Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tan PF, Westra WH, Chung CH, Jordan RC, Lu C, Kim H, Axelrod R, Silverman CC, Redmond KP, Gillison ML (2010) Human papillomavirus and survival of patients with oropharyngeal cancer. N Engl J Med 363(1): 24–35.

Chaturvedi AK, Engels EA, Pfeiffer RM, Hernandez BY, Xiao W, Kim E, Jiang B, Goodman MT, Sibug-Saber M, Cozen W, Liu L, Lynch CF, Wentzensen N, Jordan RC, Altekruse S, Anderson WF, Rosenberg PS, Gillison ML (2011) Human papillomavirus and rising oropharyngeal cancer incidence in the United States. J Clin Oncol 29(32): 4294–4301.

Fakhry C, Westra WH, Li S, Cmelak A, Ridge JA, Pinto H, Forastiere A, Gillison ML (2008) Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. J Natl Cancer Inst 100(4): 261–269.

Harrell Jr FE, Lee KL, Mark DB (1996) Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 15(4): 361–387.

Kallogjeri D, Piccirillo JF, Spitznagel Jr EL, Steyerberg EW (2012) Comparison of scoring methods for ACE-27: simpler is better. J Geriatr Oncol 3(3): 238–245.

Kaplan MH, Feinstein AR (1974) The importance of classifying initial co-morbidity in evaluatin the outcome of diabetes mellitus. J Chronic Dis 27(7-8): 387–404.

Lambin P, Roelofs E, Reymen B, Velazquez ER, Buijsen J, Zegers CM, Carvalho S, Leijenaar RT, Nalbantov G, Oberije C, Scott MM, Hoebers F, Troost EG, van Stiphout RG, van EW, van der Weijden T, Boersma L, Valentini V, Dekker A (2013a) 'Rapid Learning health care in oncology' - an approach towards decision support systems enabling customised radiotherapy'. Radiother Oncol 109(1): 159–164.

Lambin P, van Stiphout RG, Starmans MH, Rios-Velazquez E, Nalbantov G, Aerts HJ, Roelofs E, van EW, Boutros PC, Granone P, Valentini V, Begg AC, De RD, Dekker A (2013b) Predicting outcomes in radiation oncology–multifactorial decision support systems. Nat Rev Clin Oncol 10(1): 27–40.

O'Sullivan B, Huang SH, Siu LL, Waldron J, Zhao H, Perez-Ordonez B, Weinreb I, Kim J, Ringash J, Bayley A, Dawson LA, Hope A, Cho J, Irish J, Gilbert R, Gullane P, Hui A, Liu FF, Chen E, Xu W (2013) Deintensification candidate subgroups in human papillomavirus-related oropharyngeal cancer according to minimal risk of distant metastasis. J Clin Oncol 31(5): 543–550.

Posner MR, Lorch JH, Goloubeva O, Tan M, Schumaker LM, Sarlis NJ, Haddad RI, Cullen KJ (2011) Survival and human papillomavirus in oropharynx cancer in TAX 324: a subset analysis from an international phase III trial. Ann Oncol 22(5): 1071–1077.

Rietbergen MM, Brakenhoff RH, Bloemena E, Witte BI, Snijders PJ, Heideman DA, Boon D, Koljenovic S, Baatenburg-de Jong RJ, Leemans CR (2013a) Human papillomavirus detection and comorbidity: critical issues in selection of patients with oropharyngeal cancer for treatment de-escalation trials. Ann Oncol 24(11): 2740–2745.

Rietbergen MM, Leemans CR, Bloemena E, Heideman DA, Braakhuis BJ, Hesselink AT, Witte BI, Baatenburg de Jong RJ, Meijer CJ, Snijders PJ, Brakenhoff RH (2013b) Increasing prevalence rates of HPV attributable oropharyngeal squamous cell carcinomas in the Netherlands as assessed by a validated test algorithm. Int J Cancer 132(7): 1565–1571.

Rischin D, Young RJ, Fisher R, Fox SB, Le QT, Peters LJ, Solomon B, Choi J, O'Sullivan B, Kenny LM, McArthur GA (2010) Prognostic significance of p16INK4A and human papillomavirus in patients with oropharyngeal cancer treated on TROG 02.02 phase III trial. J Clin Oncol 28(27): 4142–4148.

Shaw R, Robinson M (2011) The increasing clinical relevance of human papillomavirus type 16 (HPV-16) infection in oropharyngeal cancer. Br J Oral Maxillofac Surg 49(6): 423–429.

Smeets SJ, Hesselink AT, Speel EJ, Haesevoets A, Snijders PJ, Pawlita M, Meijer CJ, Braakhuis BJ, Leemans CR, Brakenhoff RH (2007) A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen. Int J Cancer 121(11): 2465–2472.