Clinical and Translational Radiation Oncology 6 (2017) 1–6

Contents lists available at ScienceDirect

Clinical and Translational Radiation Oncology

journal homepage: www.elsevier.com/locate/ctro

Original Research Article

A predictive model for residual disease after (chemo) radiotherapy in oropharyngeal carcinoma: Combined radiological and clinical evaluation of tumor response

Zeno A.R. Gouw [a,*], Bas Jasperse [b], Jan-Jakob Sonke [a], Wilma D. Heemsbergen [a], Arash Navran [a], Olga Hamming-Vrieze [a], Jan Paul de Boer [c], Michiel W.M. van den Brekel [d,e], Abrahim Al-Mamgani [a]

[a] Department of Radiation Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands
[b] Department of Radiology, The Netherlands Cancer Institute, Amsterdam, The Netherlands
[c] Department of Medical Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands
[d] Department of Head and Neck Oncology and Surgery, The Netherlands Cancer Institute, Amsterdam, The Netherlands
[e] Department of Oral and Maxillofacial Surgery, Academic Medical Center, University of Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 15 May 2017
Revised 5 July 2017
Accepted 12 July 2017
Available online 4 August 2017

Keywords:
Radiotherapy
Residual neoplasm
Oropharyngeal neoplasms
Radiology
Risk assessment
Magnetic resonance imaging

ABSTRACT

Background and purpose: Early detection of Residual disease (RD) is vital for salvage possibilities after (chemo) radiatiotherapy for oropharyngeal carcinoma (OPC). We standardized clinical investigation to test its added value to MRI response evaluation and investigated the benefit of FDG-PET/CT.
Materials and methods: Radiological response evaluation using Ojiri-score was done for 234 patients with OPC, using MRI 12 weeks after (chemo) radiotherapy between 2010 and 2014. The presence of mucosal lesions and/or major complaints (still completely tube feeding-dependent and/or opiate-dependent because of swallowing problems) was scored as clinical suspicion (CS). Retrospectively, the performance of Ojiri to predict RD was compared to CS and both combined using Pearson Chi-squared. Of the whole group, FDG-PET/CT metabolic response (MR) was available in 50 patients.
Results: Twelve out of 234 patients (5.1%) had RD. Ojiri and CS had excellent negative predictive value (NPV) (98% and 100% respectively). The combination of CS and Ojiri reduced false positives by 32% (38–26 patients) without lowering NPV (98%). No patients with complete MR ($n = 39$) at the FDG-PET/CT had RD compared to 5 (45%) with partial MR.
Conclusion: For response evaluation in OPC, the combination of CS and Ojiri-score improved the predictive accuracy by reducing false positives compared to them individually. FDG-PET/CT is promising to further reduce false positives.
© 2017 The Authors. Published by Elsevier Ireland Ltd on behalf of European Society for Radiotherapy and Oncology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Head and neck squamous cell carcinoma (HNSCC) ranks sixth among the most common cancers [1]. The incidence of oropharyngeal cancer (OPC) is increasing because of human papilloma virus (HPV) positive OPC [2,3].

Loco-regional control and overall survival (OS) in OPC have been impressively improved [4,5]. Nonetheless, local failure (LF) remains an important cause of morbidity and mortality [4,6]. Since a significant proportion of LFs arise from persisting local tumor after treatment, accurate response evaluation is vital. The early detection of residual disease (RD) can provide possibilities for salvage surgery [6].

Response evaluation is routinely done using radiological examination with magnetic resonance imaging (MRI) and/or computed tomography (CT) 2–3 months after (chemo) radiotherapy. In case of suspicion of RD, an examination under anesthesia (EUA) with biopsies is performed to confirm RD.

Ojiri et al. [7] proposed a 4-point grading scale, comparing the primary tumor on pre- and post-treatment imaging. Patients with a higher score had a significantly increased risk of RD and LF. Van den Broek et al. [8] used MRI to determine the Ojiri-score 6–8 weeks posttreatment. They demonstrated Ojiri-score is

* Corresponding author at: Postbus 90203, 1006 BE Amsterdam, The Netherlands. Fax: +31 20 669 1101.
E-mail addresses: z.gouw@nki.nl (Z.A.R. Gouw), b.jasperse@nki.nl (B. Jasperse), j.sonke@nki.nl (J.-J. Sonke), w.heemsbergen@nki.nl (W.D. Heemsbergen), a.navran@nki.nl (A. Navran), o.vrieze@nki.nl (O. Hamming-Vrieze), j.d.boer@nki.nl (J. Paul de Boer), m.vd.brekel@nki.nl (M.W.M. van den Brekel), a.almamgani@nki.nl (A. Al-Mamgani).

http://dx.doi.org/10.1016/j.ctro.2017.07.002
2405-6308/© 2017 The Authors. Published by Elsevier Ireland Ltd on behalf of European Society for Radiotherapy and Oncology.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

excellent in identifying low-risk patients, obviating the need for an EUA in these patients. The relatively high false positives however can result in unnecessary and potentially harmful EUA's to exclude recurrences. Patients with OPC often have cardiovascular and pulmonary comorbidities raising the risk of anesthesia [9]. Post-radiotherapy invasive procedures, such as biopsies, was one of the significant predictive factors for persistent radionecrosis and mucosal ulceration [10].

The risk of these technical improvements in medical imaging is that it can cause physicians to disregard their clinical examination and judgment. In fact, articles on response evaluation often do not report symptoms and clinical examination at all [7,8,11,12]. Standardizing clinical investigation makes it possible to investigate it's performance and include it into a response evaluation model.

The study investigates the predictive performance of standardized clinical investigation: clinical suspicion (CS). It aims to develop a prediction model for RD in patients with OPC after (chemo) radiotherapy by combining Ojiri-score with CS. Patients will be stratified in RD risk groups to guide the need for further investigations. Finally, the predictive value of fluorodeoxyglucose (FDG)-positron emission tomography (PET)/CT for response evaluation will be investigated.

Patients and methods

Patients

Patients were included between January 2010 and April 2014. Inclusion criteria were histological proven squamous cell OPC scheduled for curative (chemo) radiotherapy. The research proposal was approved by the medical ethical committee and there was a waiver of the requirements for obtaining informed consent. All data were collected retrospectively from the electronic health record. The patient cohort was previously used to analyze the pattern of failure after (chemo) radiotherapy for OPC and to correlate the site of failure to the received radiation dose [6].

Pre-treatment evaluations consisted of complete medical history, physical examination and EUA with biopsies. Staging was completed with chest X-ray, ultrasound with FNA of the neck and head and neck MRI-scan. FDG-PET/CT was performed in patients with stage III and IV disease. All patients were presented at the weekly multidisciplinary meeting where the choice for radiotherapy alone or chemoradiotherapy was based on institutional guidelines.

Imaging

MRI was performed before treatment and the median interval after treatment was 11 weeks (range 8–14 weeks) for response evaluation. MRI scans were acquired at 1.5 Tesla on a Philips MRI scanner. The imaging protocol included T1 turbo spin echo weighted sequences before and after gadolinium injection (TR/TE: 538/10 ms, FA: 90 degrees, slice thickness: 3–4 mm with an 0.5 interslice gap), axial and coronal T2 SPIR weighted sequences (TR/TE: 3554/90 ms, FA: 90 degrees, slice thickness 3 mm with a 0.3 interslice gap) and 3D Thrive (T1 weighted) sequences after gadolinium injection (TR/TE: 4595/10 ms, FA: 10 degrees, slice thickness: 0.9 mm).

FDG-PET/CT was performed in case of a discrepancy between the clinical and radiological findings or in some cases according to physician discretion or patients' request. The median time between the end of treatment and the response evaluation by means of FDG-PET/CT was 15 weeks (range 13–20 weeks). The PET/CT is a Philips Gemini TF (with time of flight). The voxel size in the head- and neck area is 2 mm. The scanner fulfils the EARL accreditation. Osirix version 7.0.3 was used for the PET analysis. We defined complete metabolic response (CMR) as the absence of visually detectable uptake of FDG at prior tumor locations, above the surrounding background of FDG uptake in normal tissues in the head-neck area. Those with any residual metabolic activity above background inside the original tumor area were scored as partial metabolic response (PMR).

Treatment

Patients were immobilized in supine treatment position in a custom-made head-and-neck mask. A contrast-enhanced planning CT-scan was performed in all patients. The gross tumor volume of the primary tumor (GTVp) and the involved node(s) (GTVn) were delineated on the contrast-enhanced CT-scan. The high-dose clinical target volumes ($CTVp_{70Gy}$ and $CTVn_{70Gy}$) was generated by adding 1 cm margin to the delineated GTVp and GTVn. The elective CTV of the primary tumor ($CTVp_{46Gy}$) was generated by adding 5 mm margin to the generated $CTVp_{70Gy}$. The elective CTV of the neck ($CTVn_{46}$) consisted of level I–V in case of node-positive and level II–IV in case of node-negative neck. The planning target volume (PTV) included a margin of 5 mm beyond all CTVs to account for different targeting uncertainties. Treatments were delivered using Intensity-Modulated Radiation Therapy or Volumetric Modulated Arc Therapy. The radiation treatment consisted of elective radiotherapy to one or both sides of the neck; 46 Gy in 23 fractions in case of sequential boost and 54.25 Gy in 35 fractions in case of concomitant boost. The primary tumor and involved nodes received a total dose 70 Gy in 35 fractions. Set-up verification and correction of the patients was done with an online or offline Cone-Beam CT-guided correction protocol.

Patients with T3–4 tumors, N2c-3 or extracapsular extension of lymph node metastases, based on the MRI findings, received concomitant cisplatin (100 mg/m^2), at day 1, 22 and 43 of treatment.

Response evaluation and follow-up

Following completion of treatment, patients were seen at the out-patient clinic every 2 weeks until the acute radiation-induced toxicities subsided. Response evaluation was performed 11 weeks (range 8–14) after treatment with MRI-scans. For the primary tumor, the posttreatment MRI was compared to the pre-treatment MRI by a dedicated head and neck radiologist. The response was scored according to a 4-point grading system introduced by Ojiri et al.[7]. Ojiri 0 was scored if no detectable focal abnormalities other than post-radiation changes remained. Ojiri 1 represented anatomical asymmetry or discrete mass $\leq$ 10 mm. Ojiri 2a meant the presence of a discrete mass > 10 mm and grade 2b was scored when the greatest diameter reduced by less than 50% [7,8]. In 50 patients, FDG-PET/CT was also performed for the response evaluation, 15 weeks (range 13–20) after completion of the (chemo) radiotherapy. Of those patients, 30 patients belong to the high-risk group of RD, 18 to the intermediate risk, and 2 to the low-risk group. The FDG-PET/CT was done in the 2 low-risk group because of patients' request or doctor discretion and in high- and intermediate-risk groups because of clinical and/or radiological suspicion for RD (major complaints, mucosal lesion and/or Ojiri 2a or 2b).

One week after the response MRI, patients attended the outpatient clinic to discuss the results and for clinical evaluation. The presence of mucosal lesion (ulcer, mass or irregular aspect) and/or major complaints (still completely tube feeding-dependent and/or opiate-dependent because of swallowing pain) was defined as highly suspicious for RD, those patients scored positive for clinical suspicion (CS). RD was defined as LF within the first 6 months

after the start of treatment. LF's were confirmed by biopsy after radiological- or clinical suspicion.

After the response evaluation, medical histroy and clinical examination (including flexible laryngoscopy) continued once every 3 months the first year, once every 4 months the second year and thereafter every 6 months.

End points and statistical analysis

The performance of Ojiri-score and CS to predict RD was evaluated using Pearson Chi-squared test. This was compared with the combination of both Ojiri-score and CS (a positive result consisted of patients with both Ojiri-score 2a + b and CS). Subsequently the sensitivity, specificity, accuracy, positive predictive value (PPV) and negative predictive value (NPV) were calculated. In a subgroup of patients, this was also performed for FDG-PET/CT CMR and PMR.

A Kaplan-Meier estimator was used to investigate Local failure free survival (LFFS) for patients grouped to Ojiri-score and CS and curves were compared using a log-rank test.

All *p*-values were 2-sided using a significance threshold of 0.05. All analyses were done using SPSS statistics version 22, 2013.

Results

The current study included 234 patients with OPC with MRI pre-treatment and 3 months after completion of (chemo) radiotherapy. Out of the original cohort of 240 patients, 6 were excluded because no response MRI was present. The median follow-up was 33 months (range, 5–70). Table 1 shows patient's demographics.

Table 1
Patient's demographics.

		n = 234 (%)
Age at start treatment (years)		
	Range	38–88
	Median	62
Gender		
	Male	152 (65)
	Female	82 (35)
Follow-up (months)		
	Range	5–70
	Median	33
T-stage		
	T1	46 (20)
	T2	88 (38)
	T3	53 (23)
	T4	47 (20)
N-stage		
	N0	54 (23)
	N1–3	180 (77)
AJCC-stage		
	I	4 (2)
	II	55 (24)
	III	120 (51)
	IV	55 (24)
HPV status		
	Positive	107 (46)
	Negative	100 (43)
	Unknown	27 (11)
Chemotherapy		
	No	96 (41)
	Yes	138 (59)
Ojiri-score		
	0 + 1	188 (80)
	2a + b	46 (20)
Clinical suspicion		
	No	163 (70)
	Yes	71 (30)

Patient, tumor, and treatment characteristics of all patients. Follow-up range is depicted between parentheses. *Abbreviations*: FU = follow-up, HPV = human papilloma virus.

Out of 234 patients, 12 had RD (5.1%). The incidence of RD for patients with Ojiri-score 0 + 1 and 2a + b was 2.1% (4 out of 188) and 17.4% (8 out of 46), respectively ($p < 1e{-}4$) and for patients with and without CS was 16.9% (12 out of 71) and 0% (0 out of 163), respectively ($p < 1e{-}7$). Of the 34 patients who had both CS and Ojiri 2a + b, 23.5% (8 out of 34) had RD compared to 2% (4 out of 200) for all other patients ($p < 1e{-}6$). The combination of CS and Ojiri reduced the false positives by 32% (38–26) without increasing the false negatives (Table 2).

Table 3 demonstrates the performance of the Ojiri-score, CS and the combination of the two for separating patients with and without RD. Sensitivities of Ojiri-score, CS and the combination were 67%, 100% and 67%, respectively. Specificities were 83%, 73% and 88%, respectively. PPVs were 17%, 17% and 24%, respectively. NPVs were 98%, 100% and 98%, respectively. In all the HPV-positive patients ($n = 107$) no RD was reported (0%), while 12% of patients (12 out of 100) had RD in the HPV-negative group ($p < 1e{-}3$). Supplementary Table 1 shows the performance of Ojiri-score, CS and the combination of both stratified by HPV-status. In the 100 patients with an HPV-negative tumor, the sensitivities were 67%, 100% and 67%, respectively, the specificities were 80%, 65% and 84%, respectively, the accuracies were 78%, 69% and 82%, respectively, the PPVs were 31%, 28% and 36%, respectively, and the NPVs were 95%, 100%, 95%, respectively. In the 107 patients with an HPV-positive tumor, there were no patients with RD so the sensitivities could not be calculated. The specificity and accuracy were 84%, 79% and 90%, respectively, the PPV was 0%, and the NPV was 100% for all three.

The long-term prognostic values of Ojiri-score and CS are illustrated in Figs. 1 and 2 and Supplementary Figs. 1 and 2. Overall LFFS was 86.8% (203 out of 234 patients). Stratified by Ojiri, LFFS in Ojiri 0 + 1 was 89.4% (168 out of 188 patients) and in Ojiri 2a + b was 76.1% (35 out of 46 patients). Fig. 1 shows the

Table 2
Crosstabs Ojiri and CS vs RD.

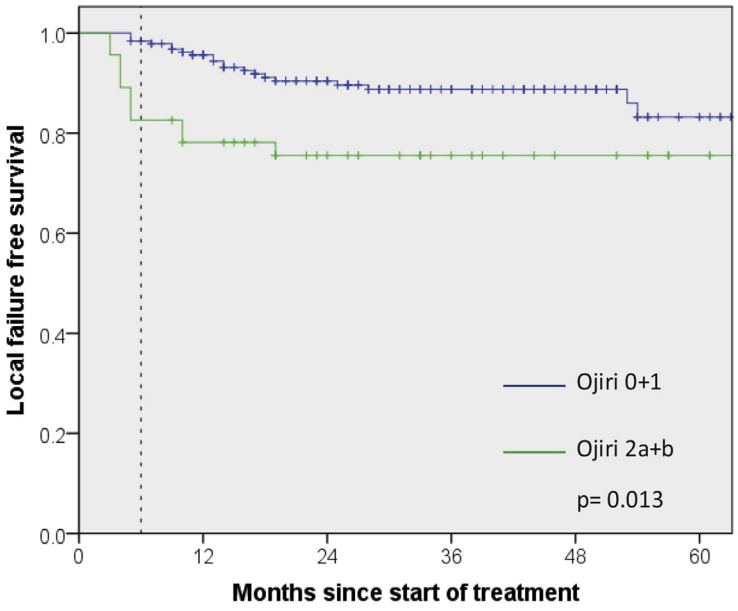
		RD		Total
		No	Yes	
Ojiri	0 + 1	184	4	188
	2 + b	38	8	46
CS	No	163	0	163
	Yes	59	12	71
Ojiri + CS	No	196	4	200
	Yes	26	8	34
Total		222	12	234

Crosstabs showing the ability of Ojiri, CS and the combination of the two to identify patients with RD and those without RD. Ojiri, CS and the two combined were all 3 able to significantly predict RD (Pearson Chi-squared was $p < 1e{-}4$, $p < 1e{-}7$ and $p < 1e{-}6$ respectively). There was a modest decrease of false-positives for the combination. *Abbreviations:* CS = clinical suspicion, RD = residual disease.

Table 3
Response prediction performances.

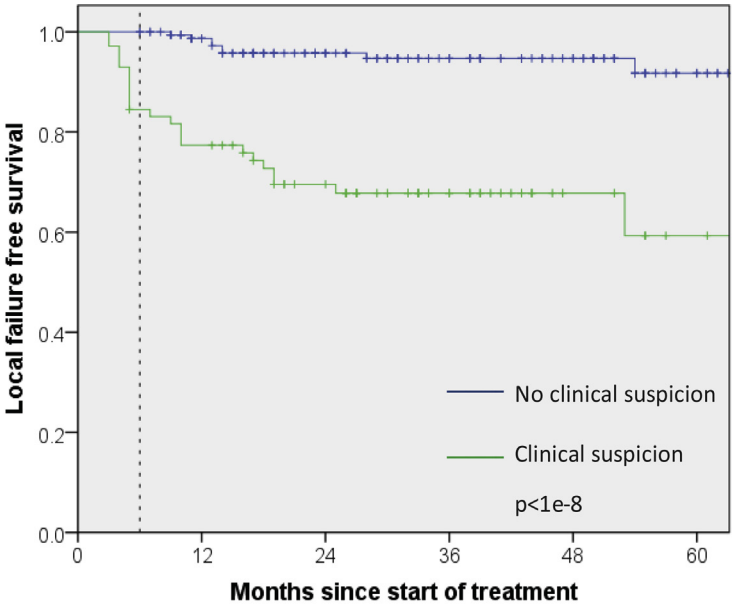
	Ojiri	CS	Ojiri + CS
Sensitivity	67% (8/12)	100% (12/12)	67% (8/12)
Specificity	83% (184/222)	73% (163/222)	88% (196/222)
Accuracy	82% (192/234)	75% (175/234)	87% (204/234)
PPV	17% (8/46)	17% (12/71)	24% (8/34)
NPV	98% (184/188)	100% (163/163)	98% (196/200)

Performances of Ojiri, CS and the two combined to predict the presence of Residual disease. The numerators and denominators of the tests are depicted between parentheses. The combination shows a modest improvement in PPV while maintaining the excellent NPV. *Abbreviations:* CS = clinical suspicion, PPV = positive predictive value, NPV = negative predictive value.



Months	0	12	24	36	48	60
Ojiri 0+1	188	159	114	77	47	19
Ojiri 2a+b	46	34	23	14	9	4

Fig. 1. Local failure free survival for patients with Ojiri-score 0 + 1 vs patients with Ojiri-score 2a + b. The dotted line represents the RD threshold of 6 months. All LF's before this threshold are categorized as RD. Twenty out of 188 patients with Ojiri 0 + 1 had LF (10.6%), of which 4 were a RD. Eleven out of 46 patients with Ojiri 2a + b (23.9%) had LF and 8 out of 11 LF's were a RD. Patients with Ojiri 0 + 1 had a significant better local failure free survival ($p = 0.013$). *Abbreviations*: RD = residual disease, LF = local failure.



Months	0	12	24	36	48	60
without CS	163	135	96	67	46	18
with CS	71	54	39	21	9	4

Fig. 2. Local failure free survival for patients with CS vs patients without CS. The dotted line represents the RD threshold of 6 months. All LF's before this threshold are categorized as RD. Of the 163 patients without CS, 8 had a LF (4.9%) and none of them were a RD. For patients with CS, 23 out of 71 had a LF (32.4%) of which 12 were a RD. Patients without CS had a significant better local failure free survival ($p < 1e{-}8$). *Abbreviations*: RD = residual disease, LF = local failure.

significant effect of Ojiri on LFFS ($p = 0.013$). CS has even a stronger effect on LFFS. In patients without CS, LFFS was 95.1% (155 out of 163 patients). For patients with CS, LFFS was 67.6% (48 out of 71 patients). The effect of CS on LFFS is shown in Fig. 2 ($p < 1e{-}8$). Furthermore, CS shows an additive prognostic value on top of Ojiri-score. This is reflected for Ojiri 0 + 1 patients in Supplementary Fig. 1 ($p < 1e{-}6$) and for Ojiri 2a + b patients in Supplementary Fig. 2 ($p = 0.034$).

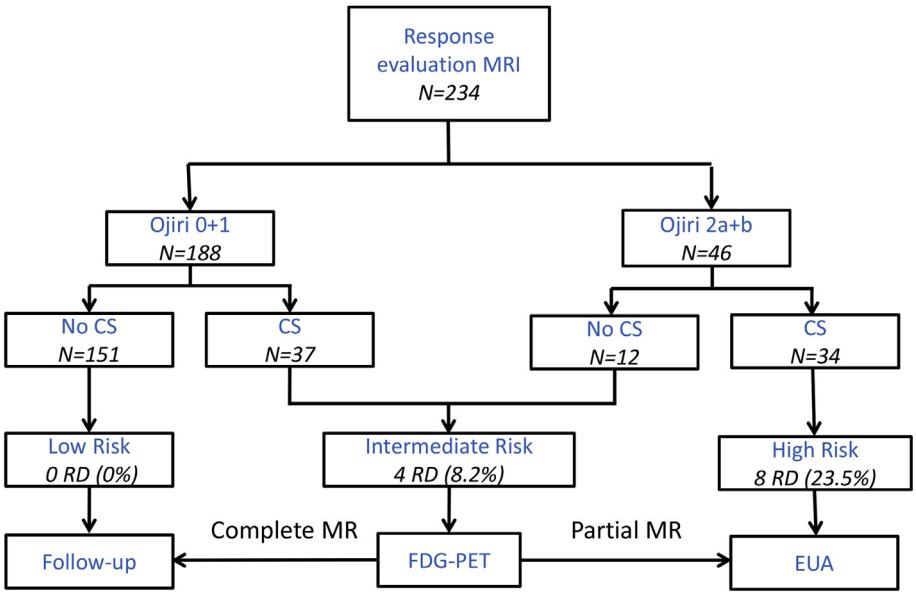


Fig. 3. Risk stratification based on Ojiri-score and CS. Three risk groups are created with 0% RD for low risk-patients, 8.2% RD in intermediate-risk and 23.5% RD in the high-risk group. We propose to further differentiate patients with- and without RD in the intermediate-risk group, Referring patients with CMR to clinical follow up and PMR to EUA with biopsies. *Abbreviations*: RD = residual disease, CS = clinical suspicion, *N* = number of patients, MR = metabolic response, EUA = examination under anesthesia.

Risk group stratification

Based on this data, patients with OPC were stratified into 3 risk groups with regard to the risk of RD, using Ojiri-score and CS (Fig. 3). The low-risk group contained patients with Ojiri 0 + 1 without CS. The high-risk group of patients with Ojiri 2a + b and CS. All other patients (Ojiri 0 + 1 with CS and Ojiri 2a + b without CS) were classified as intermediate-risk. Low-risk patients had no RD. Intermediate-risk patients had 8.2% RD (4 out of 49 patients). High-risk patients had 23.5% RD (8 out of 34 patients).

Response evaluation using FDG-PET/CT was done in 50 patients. Of the whole group, 39 patients had a CMR (78%). None of these patients had RD translating to an excellent sensitivity (100%) and NPV (100%). Of the 11 patients with PMR, 5 had RD, translating to a PPV of 45%. The specificity and accuracy were 87% and 88%, respectively. Subgroup analysis by risk stratification showed that both patients from the low-risk group had CMR (100%). The figures for the intermediate- and high-risk groups were 17 (94%), and 20 (67%), respectively.

Discussion

Response evaluation after (chemo) radiotherapy for OPC is routinely done using clinical examination with MRI and/or CT 2–3 months after (chemo) radiotherapy.

The current study aimed to standardize the clinical investigation and examine its predictive value for RD and to construct a minimally invasive prediction model for RD, thereby decreasing unnecessary EUA's and biopsies with subsequent reduction of the associated morbidity of these investigations in this vulnerable group of patients [9,10,13].

We defined CS as the presence of mucosal lesions (ulcers, irregular aspect or masses) and/or major complaints (still completely tube feeding-dependent and/or opiate-dependent because of serious swallowing problems). This definition is in concordance with daily clinical practice in our institution. The combination of CS and Ojiri-score resulted in a reduction of false positives by 32% without increasing the false negatives. Our results show why standardized clinical investigation should be an important part of the response assessment since it had a similar predictive value compared to MRI (Ojiri-score). This makes it very remarkable that the role of clinical judgement is barely mentioned in the response assessment studies [7,8,11,12].

Based on the results of the current study, we could use predefined criteria for CS and Ojiri-score to stratify patients into three risk groups with regard to RD to guide the need for additional investigations. In the low-risk group, additional investigations can safely be omitted without jeopardizing salvage possibilities. No patients had a RD in this group (Fig. 3). In the high-risk group, 23.5% of patients had RD, making an EUA with biopsies mandatory. This leaves an intermediate group of patients with a relatively low rate of RD (8.2%) but too high to omit further investigations altogether. In order to reduce the false positivity of this model we add FDG-PET/CT to the response evaluation in 50 patients. This demonstrated excellent sensitivity and NPV and, therefore, FDG-PET/CT might be promising.

The HPV-status was highly predictive for RD. When we repeated the analysis according to the HPV-status, the incidence of RD in HPV-positive and HPV-negative patients were 0% and 12%, respectively ($p < 1e{-}3$). Nevertheless, we aimed to construct a response prediction model, applicable for all patients with OPC, independent of HPV-status. Even in patients with HPV-positive disease with a very low chance of having RD, radiation oncologist will still perform additional investigation in these patients when there is clinical (having major complaints or mucosal ulcer) or radiological (Ojiri 2a and 2b) suspicion of RD.

Patients with OPC are often treated primarily with (chemo) radiotherapy aiming at organ-preservation. Accurate response evaluation is mandatory to offer salvage surgery for patients with RD. Our group recently analyzed the impact of salvage treatment on OS in patients with recurrent or residual OPC after organ preserving treatment. Two-year OS after local recurrence was significantly better in patients with salvage treatment possibilities compared to those without (82% vs. 21% OS after 2 years; $p < 0.001$, HR for salvage: 0.18, 95% CI: 0.07–0.51) [6].

Different imaging modalities are used to evaluate the response of HNSCC to the organ-preservation strategies. The survey conducted by the Dutch Head and Neck Oncology Cooperative Group

in eight head and neck cancer centers concerning response evaluation after chemoradiotherapy for advanced OPC showed a substantial variation in the diagnostic policy concerning response evaluation after CRT in the Netherlands [14]. Similar to our study, King et al. [15] investigated the role of MRI in early posttreatment assessment of the primary tumor by dividing patients according to the pattern of the residual masses on the T2-weighted MRI. In that study, the incidence of RD in pattern 1 (only scar tissue), pattern 2 (no signs of pattern 1 or 3) and pattern 3 (expansile mass $\geq$ 1 cm) were 0%, 55%, and 100%, respectively. Abdel Razek et al. [16] identified a cut off ADC value of 1.3×10^{-3} mm^2/s for the differentiation of RD from post-therapeutic changes in cohort of 30 patients with an accuracy of 87%. Vandecaveye et al. [17] compared the pretreatment ADC value to the values after chemoradiotherapy for HNSCC. The increase in ADC was significantly lower in patients with a later recurrence.

FDG-PET/CT is nowadays increasingly used for the response evaluation. Mehanna et al. [18] showed that PET-CT-guided surveillance after (chemo) radiotherapy for HNSCC resulted in comparable OS, compared to patients underwent planned neck dissection. Furthermore, PET-CT-guided surveillance resulted in significantly less neck dissections than did planned dissection surgery (54 vs. 221) while the rates of surgical complications were similar in the two groups.

With regard to the use FDG-PET for the response evaluation of the primary site, a meta-analysis of 2.335 patients with HNSCC showed a NPV of 95% and PPV of 59% for FDG-PET/CT [12]. Kim et al. [11] described a series of 78 patients with locally advanced HNSCC treated with chemoradiotherapy, none of the 41 patients with a CMR had a RD opposed to 10 out of 37 with a PMR. Very comparable results are shown in our study in the 50 patients who received FDG-PET/CT for response evaluation. The sensitivity and NPV was 100% and the PPV was 45%. When this analysis was repeated for the small subgroups, 17 out of 18 patients in the intermediate-risk had MCR and none of these patients has RD.

The limitations of our study include the possible bias of the retrospective gathering of our data. Our definition of RD (local recurrence within 6 months after start of treatment) is not based on an universally accepted definition. However, if we extended this definition to 12 months the predictive values of Ojiri-score, CS and the combination remained significant and our conclusions did not alter (Supplementary Table 2). We also included long-term LFFS of Ojiri-score and CS to prove our results are not dependent on the RD definition and truly identify patient groups with different LF/RD risks.

Concluding, the combination of CS and Ojiri-score improved the predictive accuracy by reducing false positives and allowed to stratify patients into 3 risk groups with regard to RD and the need for further investigations. The role of FDG-PET in the intermediate- and possibly high-risk group seems promising to further increase the accuracy of the detection of RD, thereby reducing the number of unnecessary EUA's and biopsies with subsequent reduction of the associated morbidity of these investigations. Internal and external validations studies are planned in our institution and in collaboration with other head and neck cancer centers.

Conflict of interest

None to declare.

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

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ctro.2017.07.002.

References

[1] Ferlay J, Soerjomataram II, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2014. http://dx.doi.org/10.1002/ijc.29210.
[2] van Monsjou HS, Wreesmann VB, van den Brekel MWM, Balm AJM. Head and neck squamous cell carcinoma in young patients. Oral Oncol 2013;49(12):1097–102. http://dx.doi.org/10.1016/j.oraloncology.2013.09.001.
[3] Braakhuis BJM, Visser O, Leemans CR. Oral and oropharyngeal cancer in The Netherlands between 1989 and 2006: increasing incidence, but not in young adults. Oral Oncol 2009;45(9):e85–9. http://dx.doi.org/10.1016/j.oraloncology.2009.03.010.
[4] Pignon JP, le Maître A, Maillard E, Bourhis J. Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): an update on 93 randomised trials and 17,346 patients. Radiother Oncol 2009;92(1):4–14. http://dx.doi.org/10.1016/j.radonc.2009.04.014.
[5] 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(6):567–78. http://dx.doi.org/10.1056/NEJMoa053422.
[6] Ridder M, Gouw ZAR, Sonke JJ, et al. Recurrent oropharyngeal cancer after organ preserving treatment: pattern of failure and survival. Eur Arch Oto-Rhino-Laryngol 2017;274(3):1691–700. http://dx.doi.org/10.1007/s00405-016-4413-7.
[7] Ojiri H, Mendenhall WM, Mancuso AA. CT findings at the primary site of oropharyngeal squamous cell carcinoma within 6–8 weeks after definitive radiotherapy as predictors of primary site control. Int J Radiat Oncol Biol Phys 2002;52(3):748–54. http://dx.doi.org/10.1016/S0360-3016(01)02671-2.
[8] van den Broek GB, Rasch CRN, Pameijer FA, Peter E, van den Brekel MWM, Balm AJM. Response measurement after intraarterial chemoradiation in advanced head and neck carcinoma: magnetic resonance imaging and evaluation under general anesthesia? Cancer 2006;106(8):1722–9. http://dx.doi.org/10.1002/cncr.21786.
[9] Bøje CR, Dalton SO, Grønborg TK, et al. The impact of comorbidity on outcome in 12 623 Danish head and neck cancer patients: a population based study from the DAHANCA database. Acta Oncol 2013;52(2):285–93. http://dx.doi.org/10.3109/0284186X.2012.742964.
[10] Tribius S, Raguse M, Voigt C, et al. Residual deficits in quality of life one year after intensity-modulated radiotherapy for patients with locally advanced head and neck cancer: results of a prospective study. Strahlenther Onkol 2015;191(6):501–10. http://dx.doi.org/10.1007/s00066-015-0824-4.
[11] Kim R, Ock C-Y, Keam B, et al. Predictive and prognostic value of PET/CT imaging post-chemoradiotherapy and clinical decision-making consequences in locally advanced head & neck squamous cell carcinoma: a retrospective study. BMC Cancer 2016;16(1):1–9. http://dx.doi.org/10.1186/s12885-016-2147-y.
[12] Gupta T, Master Z, Kannan S, et al. Diagnostic performance of post-treatment FDG PET or FDG PET/CT imaging in head and neck cancer: a systematic review and meta-analysis. Eur J Nucl Med Mol Imaging 2011;38(11):2083–95. http://dx.doi.org/10.1007/s00259-011-1893-y.
[13] Culie D, Benezery K, Chamorey E, et al. Salvage surgery for recurrent oropharyngeal cancer: post-operative oncologic and functional outcomes. Acta Otolaryngol 2015;6489:1–7. http://dx.doi.org/10.3109/00016489.2015.1066935.
[14] Schouten CS, Hoekstra OS, Rene C, De Bree R. Response evaluation after chemoradiotherapy for advanced staged oropharyngeal squamous cell carcinoma: a nationwide survey in the Netherlands. 2015:3507–13. http://dx.doi.org/10.1007/s00405-014-3389-4.
[15] King AD, Keung CK, Yu KH, Mo FK, Bhatia KS, Yeung DK, et al. T2-weighted MR imaging early after chemoradiotherapy to evaluate treatment response in head and neck. AJNR Am J Neuroradiol 2013;34:1237–41.
[16] Abdel Razek AAK, Kandeel AY, Soliman N, et al. Role of diffusion-weighted Echo-Planar MR imaging in differentiation of residual or recurrent head and neck tumors and posttreatment changes. Am J Neuroradiol 2007;28(6):1146–52. http://dx.doi.org/10.3174/ajnr.A0491.
[17] Vandecaveye V, Dirix P, De Keyzer F, et al. Diffusion-weighted magnetic resonance imaging early after chemoradiotherapy to monitor treatment response in head-and-neck squamous cell carcinoma. Int J Radiat Oncol Biol Phys 2012;82(3):1098–107. http://dx.doi.org/10.1016/j.ijrobp.2011.02.044.
[18] Mehanna H, Wong W-L, McConkey CC, et al. PET-CT surveillance versus neck dissection in advanced head and neck cancer. N Engl J Med 2016:1–11. http://dx.doi.org/10.1056/NEJMoa1514493.