CLINICAL REVIEW

WILEY

Prognostic imaging variables for recurrent laryngeal and hypopharyngeal carcinoma treated with primary chemoradiotherapy: A systematic review and meta-analysis

Hilde J. G. Smits MSc[1] | **Sanam Assili MSc[1]** | **Frans Kauw MD[1]** |
Marielle E. P. Philippens MD, PhD[1] | **Remco de Bree MD, PhD[2]** |
Jan W. Dankbaar MD, PhD[1]

[1]Department of Radiology, University Medical Center Utrecht, University of Utrecht, Utrecht, the Netherlands

[2]Department of Head and Neck Surgical Oncology, University Medical Center Utrecht, University of Utrecht, Utrecht, the Netherlands

Correspondence
Hilde J. G. Smits, Department of Radiology, University Medical Center Utrecht, University of Utrecht, Heidelberglaan 100, 3584 CX, Utrecht, the Netherlands.
Email: h.j.g.smits-6@umcutrecht.nl

Abstract

Background: In this systematic review, we aim to identify prognostic imaging variables of recurrent laryngeal or hypopharyngeal carcinoma after chemoradiotherapy.

Methods: A systematic search was performed in PubMed and EMBASE (1990–2020). The crude data and effect estimates were extracted for each imaging variable. The level of evidence of each variable was assessed and pooled risk ratios (RRs) were calculated.

Results: Twenty-two articles were included in this review, 17 on computed tomography (CT) and 5 on magnetic resonance imaging (MRI) variables. We found strong evidence for the prognostic value of tumor volume at various cut-off points (pooled RRs ranging from 2.09 to 3.03). Anterior commissure involvement (pooled RR 2.19), posterior commissure involvement (pooled RR 2.44), subglottic extension (pooled RR 2.25), and arytenoid cartilage extension (pooled RR 2.10) were also strong prognostic factors.

Conclusion: Pretreatment tumor volume and involvement of several subsites are prognostic factors for recurrent laryngeal or hypopharyngeal carcinoma after chemoradiotherapy.

KEYWORDS
computed tomography, laryngeal carcinoma, magnetic resonance imaging, prognosis, recurrence

1 | INTRODUCTION

Total laryngectomy (TLE) has for many years been a valid curative treatment of laryngeal and hypopharyngeal carcinoma. However, a TLE can have a substantial impact on patients' lives, due to the crucial role the larynx has in speech and communication.[1,2] That is why besides oncological outcome, the preservation of laryngeal functions has progressively gained an important role in treatment decision-making in patients with laryngeal and hypopharyngeal carcinoma.[3,4] Larynx-preserving alternatives to TLE are radiotherapy and/or chemotherapy, and different types of organ-preserving surgeries,

[Corrections added after online publication, 09 April, 2021: The forest plot column in tables 2 and 4 have been updated.]

like transoral laser microsurgery and external partial laryngectomy.[5,6]

However, not all cases are suited for larynx-preserving treatment. Mainly patients with a higher T classification are at risk for undertreatment, and thus recurrent disease, when not opting for primary TLE. Patients with recurrent disease have lower survival rates and salvage TLE after radiotherapy comes with more complications than primary TLE.[7-9] Inadequate patient selection for larynx-preserving treatment may therefore have contributed to the fact that survival rates of patients with laryngeal carcinoma have barely improved in the last 20 years.[3,4,10,11]

Identification of imaging variables, which are associated with recurrent disease, may help in treatment decision-making. Various imaging modalities, such as magnetic resonance imaging (MRI), computed tomography (CT), and fluorinedeoxyglucose positron emission tomography (FDG-PET),[18] may be used to identify patients who are unlikely to benefit from radiotherapy. Quantitative prognostic factors on pretreatment imaging could help improve patient selection and may increase the overall survival rate of patients with laryngeal and hypopharyngeal carcinoma.

Several possible prognostic imaging variables have been described in the literature. In this paper, we conducted a systematic review to identify pretreatment, quantitative imaging variables that are associated with recurrent disease in patients with laryngeal or hypopharyngeal carcinoma treated with chemoradiotherapy. A meta-analysis of the imaging variables was performed.

2 | METHODS

This study is compliant with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) criteria.[12]

2.1 | Search strategy and study selection

A systematic search was performed in PubMed/Medline and Embase (1990–July 2020) with synonyms of "larynx or hypopharynx carcinoma" AND ("imaging modalities" OR "imaging variables") AND "recurrences" AND "radiotherapy" AND "prognosis" (see Table S1, Supporting Information for full search strategy). The title and abstract of all articles were first screened by one author (Hilde J. G. Smits) to identify relevant articles. The following inclusion criteria were applied: (1) study population of patients with laryngeal or hypopharyngeal carcinoma; (2) treatment with chemoradiotherapy; (3) tumor imaging before the start of the treatment; (4) recurrences or local control was the studied outcome; and (5) ≥10 recurrences were reported within the study population. Case reports, reviews, conference abstracts, non-English articles, and animal studies were excluded. Full text analysis determined whether articles met the inclusion criteria. Finally, additional eligible articles were selected by cross-reference check of the included articles.

2.2 | Quality assessment

The quality of the selected studies was assessed with a scoring system evaluating 10 predefined criteria (Table S2). The scoring system was based on the Newcastle Ottawa Quality Assessment Scale[13] and adapted to fit the topic of this review.[14-16] A study that scored <8 out of a maximum of 13 points for the 10 criteria was excluded from further analysis. The quality assessment was performed independently by two authors (Hilde J. G. Smits and Jan W. Dankbaar). Disagreements between authors on the scores were solved by consensus.

2.3 | Data extraction

Study characteristics were extracted from all high quality articles. If two or more articles were found to have an overlapping patient population, we included the results of both articles if unique variables were studied. If the same variable was reported, only the result from the most recent article was included in this review.

For each imaging variable, effect estimates (risk ratio [RR], odds ratio, or hazard ratio) with 95% confidence interval (CI) were collected. If no relevant effect estimates were reported, crude data were used to calculate a univariate RR if possible. The consistency and clinical relevance of the effect estimates were assessed for each variable and they were classified as either a prognostic factor, a nonprognostic factor or a factor with inconsistent evidence (Figure 1).[15,16] The level of evidence of each prognostic and nonprognostic factor was then determined based on the number of studies that reported a significant or neutral effect estimate (Figure 1).

2.4 | Statistical analysis

Variables were sorted by imaging modality and if possible, a formal meta-analysis was performed. Inverse variance random effect models were used to pool RRs in Review Manager (RevMan, Version 5.3. Copenhagen:
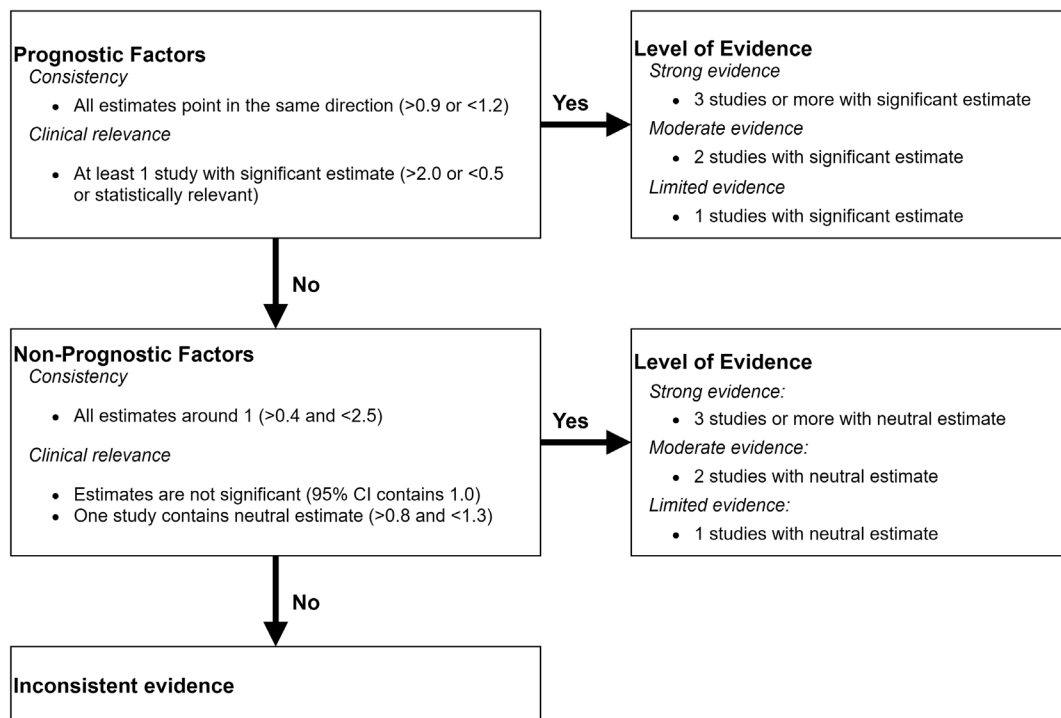
Prognostic Factors
Consistency
- All estimates point in the same direction (>0.9 or <1.2)

Clinical relevance
- At least 1 study with significant estimate (>2.0 or <0.5 or statistically relevant)

Yes →

Level of Evidence
Strong evidence
- 3 studies or more with significant estimate

Moderate evidence
- 2 studies with significant estimate

Limited evidence
- 1 studies with significant estimate

↓ **No**

Non-Prognostic Factors
Consistency
- All estimates around 1 (>0.4 and <2.5)

Clinical relevance
- Estimates are not significant (95% CI contains 1.0)
- One study contains neutral estimate (>0.8 and <1.3)

Yes →

Level of Evidence
Strong evidence:
- 3 studies or more with neutral estimate

Moderate evidence:
- 2 studies with neutral estimate

Limited evidence:
- 1 studies with neutral estimate

↓ **No**

Inconsistent evidence

FIGURE 1 Criteria for level of evidence of imaging variables

The Nordic Cochrane Centre, The Cochrane Collaboration, 2014). The weight for each study was calculated based on the variance, and pooled RRs and 95% confidence intervals were reported.

3 | RESULTS

The literature search yielded 2115 results after the duplicates were removed (Figure 2). Two hundred and eighty-five articles remained after the title and abstract screening, of which 37 articles met all inclusion criteria upon full text analysis. Of those, five articles reported neither the crude data nor ratios with confidence interval and could not be further analyzed, three articles had overlapping patient populations, and two articles scored <8 on the quality assessment. One additional publication was found through cross-reference. In the 28 remaining articles, the following imaging modalities were used: CT (17 articles), MRI (5), FDG-PET (3), dual-energy CT (1), ^{11}C-Tyrosine-PET (1), and sonography (1). The articles about the latter three modalities were excluded from the review, as were the articles about FDG-PET, since the variables were too heterogeneous to analyze. The remaining 22 articles about CT and MRI were included in this review, and their study characteristics are shown in Table 1.[17-38] The included articles cover 28 imaging variables on CT, of which 15 are studied in more than one article, and 18 articles

on MR, of which 6 are studied in more than one article.

Most articles analyzed tumor volume (TV) as a dichotomized variable, using various cut-off points ranging from 0.28 to 19 cm^3. For the purpose of the meta-analysis, the results of articles that used similar cut-off points ($\pm$0.5 cm^3) were grouped, for example, the variable TV >2 cm^3 on CT contains the range of cut-off points 1.6–2.5 cm^3). The level of evidence of all variables studied in multiple articles is shown in Tables 2–4. Results of variables or TV cut-off points that were only studied in one article can be found in Tables S3 and S4. No articles published univariate RRs, so all univariate RRs were calculated by us based on crude data. This data was either the number of recurrences within the study period, or the local control rates. Only a few studies provided other estimates or multivariable analyses. These results were included in Tables 2–4, but the pooled RRs are only based on the univariate RRs.

3.1 | Prognostic factors

Seven imaging variables were found to be prognostic factors with a strong level of evidence, all of them CT-variables (Table 2). TV >1 cm^3 (pooled RR = 3.03 [95% CI = 1.98–4.66]), TV >2 cm^3 (pooled RR = 2.48 [1.71–3.61]), TV >8 cm^3 (pooled RR = 2.09 [1.28–3.41]), anterior commissure involvement (pooled RR = 2.19 [1.45–

FIGURE 2 Flowchart diagram of included articles
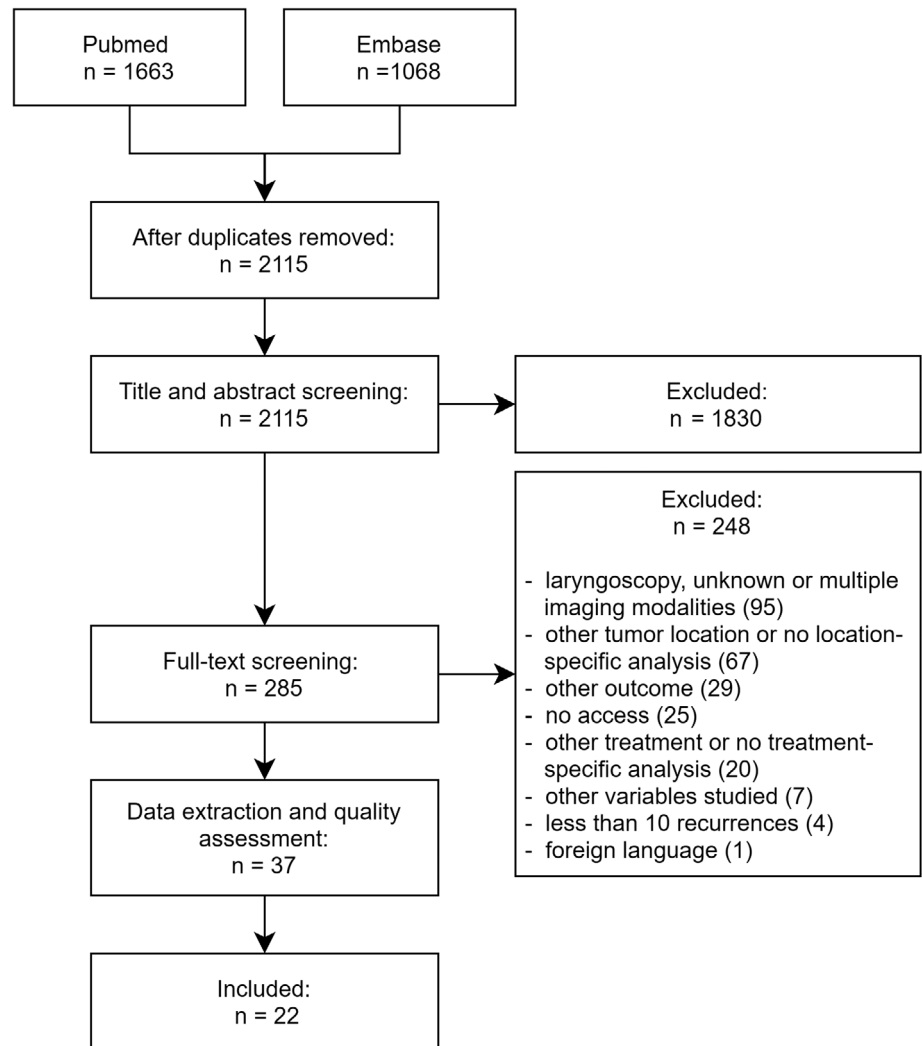
3.32]), posterior commissure involvement (pooled RR = 2.44 [1.56–3.81]), subglottic extension (pooled RR = 2.25 [1.59–3.19]), and arytenoid cartilage adjacent or invaded (pooled RR = 2.10 [1.39–3.19]) were all found to increase the risk of recurrence significantly in three or more studies. Arytenoid cartilage adjacent or invaded means the tumor has grown adjacent to the arytenoid or has invaded the arytenoid cartilage.

A moderate level of evidence was found for TV >6 cm^3 (pooled RR = 2.46 [1.39–4.35]), TV >16 cm^3 (pooled RR = 2.94 [1.98–4.37]), pre-epiglottic space involvement (pooled RR = 2.12 [1.40–3.22]), paraglottic space involvement (pooled RR = 1.72 [1.18–2.50]), and cricoid cartilage involvement (pooled RR = 2.17 [1.27–3.71]) as prognostic factors on CT (Table 2). On MRI, TV >3 cm^3 (pooled RR = 2.44 [1.54–3.88]), and pre-epiglottic space involvement (pooled RR = 2.62 [1.33–5.14]) were found to have a moderate level of evidence for being a prognostic factor of recurrent disease (Table 3).

An additional three variables on CT had a limited level of evidence: laryngeal ventricle involvement, any cartilage

involvement and vocal cord involvement (Table 2). For MRI, four variables were found to have a limited level of evidence: anterior commissure involvement, subglottic extension, hypopharyngeal extension and extralaryngeal spread beyond anterior commissure (Table 3).

3.2 | Nonprognostic factors

Only two variables were found to be nonprognostic factors with limited level of evidence: extralaryngeal spread (pooled RR = 1.35 [0.85–2.15]) and the involvement of more than one cartilage site (pooled RR = 1.48 [0.78–2.82]) (Table S5). Both are CT variables.

3.3 | Inconsistent evidence

On CT, TV >4 cm^3, arytenoid cartilage invasion and thyroid cartilage involvement were found to have inconsistent evidence (Table 4). TV >4 cm^3 was studied

TABLE 1 Study characteristics

Study	Design	Subsites	T classification	Sample size	Eligible sample size	#LF	LC calculation method	Follow-up (months)	Quality score	Variables studied
CT										
Agarwal et al.[17]	PC	L, H	1–4	60	60	18	2-year LC	Median 24	11	ENT*, MPP*
Bockel, van et al.[18]	RC	SPG, G	2–4	150	150	33	2-year LC	Mean 52	9	TV
Chen et al.[19]	RC	H	1–4	63	63	30	5-year LC	Median 38	10	LAI*
Dagan et al.[20]	RC	G	2	80	80	18	#LF	Median 85	8	PAGS, SGE, SPGE*
Dziegielewski et al.[21]	PC	G	3	107	60	22	#LF	NA	9	TV
Hermans et al. (1)[22]	RC	G	1–4	287	119	28	#LF	>36	9	TV, ACI, PCI, SGE, PEGS, PAGS, ArCI, ArCAI, CCI, TCI, CI, VCI, ELS, #CSI
Hermans et al. (2)[23]	RC	SPG	1–4	147	103	34	#LF	>24	10	TV, ACI, PCI, SGE, PEGS, PAGS, ARI, ARAI, CCI, TCI, CI, VCI, ELS, #CSI
Janssens et al.[24]	PC	SPG, G	2–4	270	270	64	5-year LC	Median 44	11	TV
Kraas et al.[25]	RC	SPG	1–4	28	28	11	#LF	Mean 34	9	TV
Mancuso et al.[26]	RC	SPG	1–4	63	63	16	#LF	>24	10	TV, PEGS
Murakami et al.[27]	PC	G	1, 2	68	68	16	2-year LC	Mean 27	11	TV, ACI, SGE, ArCI, ArCAI, TCI, VEN, TDM*, TD*, TCAI*
Nur et al.[28]	RC	G	is, 1, 2	114	112	15	5-year LC	Median 65	9	ACI
Pameijer et al.[29]	RC	G	3	42	42	12	#LF	>24	8	TV, ACI, PCI, SGE, PAGS, ArCAI, VEN, PAGS-FV*, #SSI*, CTVC*
Rutkowski et al.[30]	RC	G	2	115	115	33	#LF	>36	9	TV
Scherl et al.[31]	RC	L, H	3, 4	463	236	42	5-year LC	Mean 47.6	8	CI
Tsou et al.[32]	RC	H	2–4	51	51	29	#LF	Mean 24.6	8	TV, CN*
Zouhair et al.[33]	RC	G	1, 2	122	122	19	#LF	Median 85	8	ACI
MRI										
Castelijns et al. (1)[34]	RC	L	1–4	80	80	30	#LF	>24	12	TV, VCI*
Castelijns et al. (2)[35]	RC	L	1–4	80	80	30	#LF	>24	11	CAI*, CI*, #CSI*
Ljumanovic et al. (1)[36]	RC	SPG	1–4	84	84	28	5-year LC	>24	10	TV, ACI, HE, SGE, PEGS, ELA, ELTC*, TCI*, CCI*, GE*
Ljumanovic et al. (2)[37]	RC	G	1–4	118	118	39	2-year LC	>24	12	TV, ACI, HE, SGE, PEGS, ELA, TC/CAI*
Ljumanovic et al. (3)[38]	RC	SPG, G	1–4	64	64	28	2-year LC	>24	11	TV, SGE, PAGS*, DCE*

Note: Asterisks indicate variable that is only included in one article for that modality.

Abbreviations: ACI, anterior commissure involvement; ArCI, arytenoid cartilage invasion; ArCAI, arytenoid cartilage adjacent or invaded; CAI, cricoarytenoid involvement; CCI, cricoid cartilage involvement; CI, any cartilage involvement; CN, central necrosis; #CSI, number of cartilage sites involved; CT, computed tomography; CTVC, contra-lateral true vocal cord; DCE, dynamic contrast enhancement; ELA, extralaryngeal spread beyond anterior commissure; ELS, extralaryngeal spread; ELTC, extralaryngeal spread beyond thyroid or cricoid cartilage; ENT, entropy; G, glottis; GE, glottic extension; H, hypopharynx; HE, hypopharyngeal extension; L, larynx; LAI, larynx involvement; LC, local control; #LF, number of local failures; MPP, mean positive pixels; MRI, magnetic resonance imaging; NA, not available; PAGS, paraglottic space involvement; PAGS-FV, paraglottic space involvement at the false vocal cord; PC, prospective cohort; PCI, posterior commissure involvement; PEGS, pre-epiglottic space involvement; RC, retrospective cohort; SGE, subglottic extension; SPG, supraglottic extension; #SSI, number of subsites involved; TCAI, thyroid cartilage adjacent or invaded; TCAI, thyroid and/or cricoayrtenoid cartilage involvement; TCI, thyroid cartilage involvement; TD, tumor detectability; TDM, tumor dimension; TV, tumor volume; VCI, vocal cord involvement; VEN, laryngeal ventricle involvement.

in five articles, four of which reported this cut-off point to be significantly prognostic of tumor recurrence, but one study reported a neutral RR of 0.87 [0.47–1.61]. Arytenoid cartilage invasion was found to be significant in two out of three studies and thyroid cartilage involvement in only one out of three studies. The other studies for these variables all showed neutral effect estimates.

3.4 | Tumor volume cut-off point per subsite

In an effort to determine the optimal TV cut-off point, we stratified the studies according to the tumor subsite of their populations. This could only be done for supraglottic and glottic tumors, with five studies analyzing TV in supraglottic tumors and seven studies in glottic

tumors. The sensitivity and specificity of all TV cut-off points found in these studies were plotted for both glottic and supraglottic and glottic tumors (Figure 3). This graph contains results of both CT and MRI studies.

In general, lower cut-off points result in a higher sensitivity and lower specificity, while the reverse is true for higher cut-off points. Glottic tumors seem to have a lower optimal cut-off point (between 1 and 4 cm^3) than supraglottic tumors (between 4 and 10 cm^3). However, there is a high variability in sensitivity and specificity between studies, so no optimal cut-off point could be determined.

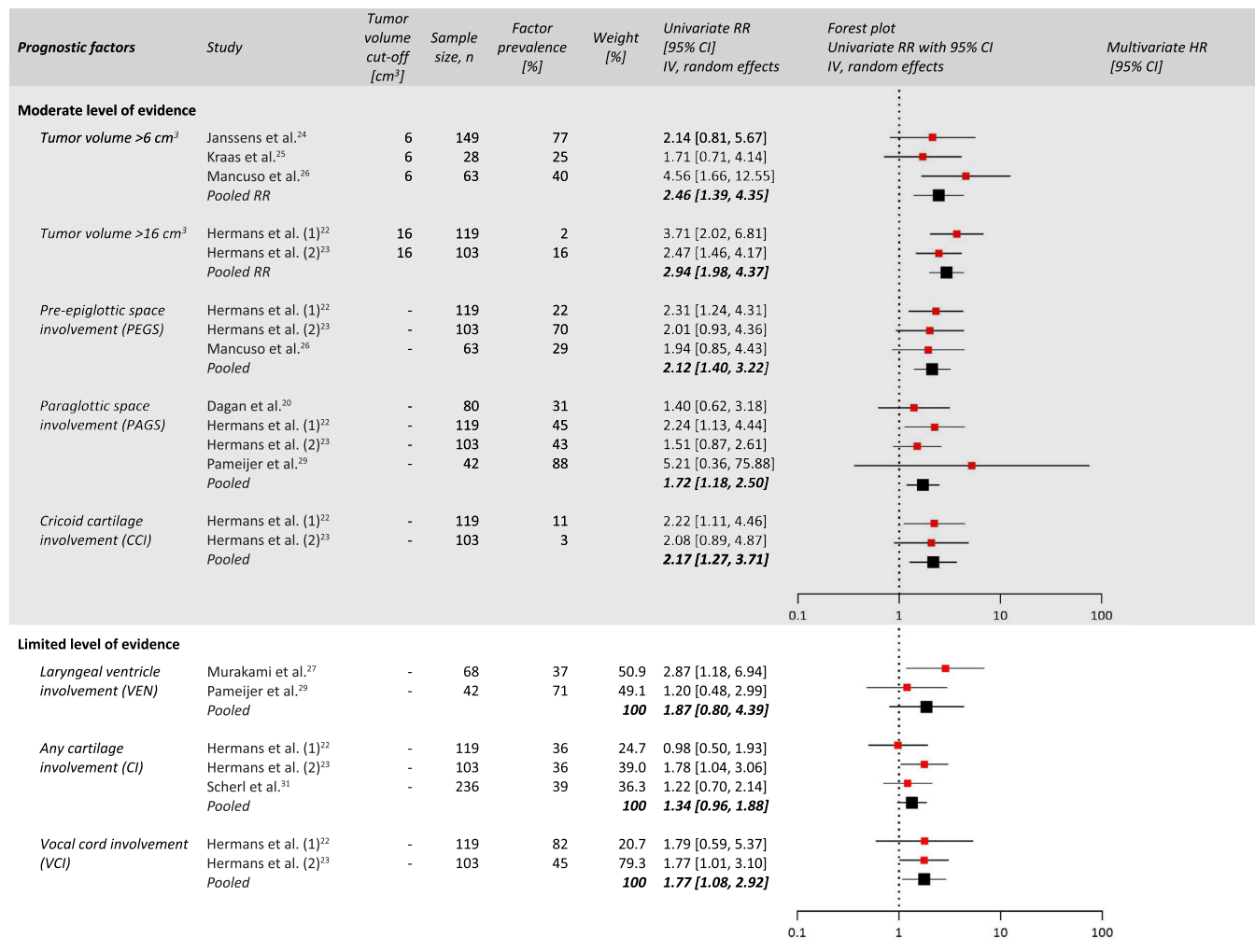## 3.5 | Anatomical subsites sensitivity and specificity

We also compared the sensitivity and specificity of anatomical subsite parameters for all studies on both CT and

TABLE 2 Prognostic factors on CT [Color figure can be viewed at wileyonlinelibrary.com]

Prognostic factors	Study	Tumor volume cut-off [cm^3]	Sample size, n	Factor prevalence [%]	Weight [%]	Univariate RR [95% CI] IV, random effects	Forest plot Univariate RR with 95% CI IV, random effects	Multivariate HR [95% CI]
Strong level of evidence								
Tumor volume >1 cm^3	Hermans et al. (1)[22]	1	119	52	33.5	2.30 [1.10, 4.80]		
	Murakami et al.[27]	1	68	12	37.9	4.50 [2.25, 9.00]		
	Rutkowski et al.[30]	0.7	115	64	28.6	2.49 [1.12, 5.54]		
	Pooled RR				100	3.03 [1.98, 4.66]		
Tumor volume >2 cm^3	Dziegielewski et al.[21]	2.5	60	58	19.2	2.43 [1.03, 5.71]		5.0 [2.0, 10.0]
	Hermans et al. (1)[22]	2	119	29	35.7	2.17 [1.16, 4.06]		
	Hermans et al. (2)[23]	2	103	91	4.0	2.97 [0.46, 19.26]		
	Rutkowski et al.[30]	1.76	115	36	41.1	2.78 [1.55, 4.98]		
	Rutkowski et al.[30]	1.6				HR 3.93 [1.06-14.54]		3.21 [1.07-9.63]
	Pooled RR				100	2.48 [1.71, 3.61]		
Tumor volume >8 cm^3	Hermans et al. (1)[22]	8	119	7	21.0	2.31 [1.06, 5.04]		
	Hermans et al. (2)[23]	8	103	48	27.6	1.10 [0.62, 1.96]		
	Kraas et al.[25]	8	28	18	21.6	2.63 [1.23, 5.61]		
	Rutkowski et al.[30]	8.2	115	7	29.8	2.97 [1.77, 4.98]		
	Pooled RR				100	2.09 [1.28, 3.41]		
Anterior commissure involvement (ACI)	Hermans et al. (1)[22]	-	119	44	25.1	1.99 [1.02, 3.88]		
	Hermans et al. (2)[23]	-	103	9	14.6	1.01 [0.38, 2.66]		
	Murakami et al.[27]	-	68	32	17.1	3.48 [1.45, 8.36]		
	Nur et al.[28]	-	112	38	17.3	2.52 [1.06, 6.01]		
	Pameijer et al.[29]	-	42	67	15.5	1.50 [0.59, 3.81]		
	Zouhair et al.[33]	-	122	50	10.5	5.33 [1.64, 17.37]		
	Pooled RR				100	2.19 [1.45, 3.32]		
Posterior commissure involvement (PCI)	Hermans et al. (1)[22]	-	119	3	18.6	2.21 [0.78, 6.24]		
	Hermans et al. (2)[23]	-	103	3	27.6	2.08 [0.89, 4.87]		
	Pameijer et al.[29]	-	42	14	53.8	2.73 [1.48, 5.01]		
	Pooled RR				100	2.44 [1.56, 3.81]		
Subglottic extension (SGE)	Dagan et al.[20]	-	80	48	14.0	2.87 [1.13, 7.31]		
	Hermans et al. (1)[22]	-	119	21	28.1	1.78 [0.92, 3.44]		
	Hermans et al. (2)[23]	-	102	7	39.2	2.34 [1.34, 4.09]		
	Murakami et al.[27]	-	68	4	14.3	3.10 [1.23, 7.81]		
	Pameijer et al.[29]	-	42	93	4.5	1.15 [0.22, 6.00]		
	Pooled RR				100	2.25 [1.59, 3.19]		
Arytenoid cartilage adjacent or invaded (ArCAI)	Hermans et al. (1)[22]	-	119	57	29.0	2.25 [1.04, 4.88]		
	Hermans et al. (2)[23]	-	103	50	50.3	1.87 [1.04, 3.36]		
	Murakami et al.[27]	-	68	4	6.3	1.44 [0.27, 7.60]		
	Pameijer et al.[29]	-	42	57	14.4	3.25 [1.09, 9.73]		
	Pooled RR				100	2.10 [1.39, 3.19]		

0.1   1   10   100

(Continues)

TABLE 2 (Continued)

Prognostic factors	Study	Tumor volume cut-off [cm³]	Sample size, n	Factor prevalence [%]	Weight [%]	Univariate RR [95% CI] IV, random effects	Forest plot Univariate RR with 95% CI IV, random effects	Multivariate HR [95% CI]
Moderate level of evidence								
Tumor volume >6 cm³	Janssens et al.[24]	6	149	77		2.14 [0.81, 5.67]		
	Kraas et al.[25]	6	28	25		1.71 [0.71, 4.14]		
	Mancuso et al.[26]	6	63	40		4.56 [1.66, 12.55]		
	Pooled RR					**2.46 [1.39, 4.35]**		
Tumor volume >16 cm³	Hermans et al. (1)[22]	16	119	2		3.71 [2.02, 6.81]		
	Hermans et al. (2)[23]	16	103	16		2.47 [1.46, 4.17]		
	Pooled RR					**2.94 [1.98, 4.37]**		
Pre-epiglottic space involvement (PEGS)	Hermans et al. (1)[22]	-	119	22		2.31 [1.24, 4.31]		
	Hermans et al. (2)[23]	-	103	70		2.01 [0.93, 4.36]		
	Mancuso et al.[26]	-	63	29		1.94 [0.85, 4.43]		
	Pooled					**2.12 [1.40, 3.22]**		
Paraglottic space involvement (PAGS)	Dagan et al.[20]	-	80	31		1.40 [0.62, 3.18]		
	Hermans et al. (1)[22]	-	119	45		2.24 [1.13, 4.44]		
	Hermans et al. (2)[23]	-	103	43		1.51 [0.87, 2.61]		
	Pameijer et al.[29]	-	42	88		5.21 [0.36, 75.88]		
	Pooled					**1.72 [1.18, 2.50]**		
Cricoid cartilage involvement (CCI)	Hermans et al. (1)[22]	-	119	11		2.22 [1.11, 4.46]		
	Hermans et al. (2)[23]	-	103	3		2.08 [0.89, 4.87]		
	Pooled					**2.17 [1.27, 3.71]**		
Limited level of evidence								
Laryngeal ventricle involvement (VEN)	Murakami et al.[27]	-	68	37	50.9	2.87 [1.18, 6.94]		
	Pameijer et al.[29]	-	42	71	49.1	1.20 [0.48, 2.99]		
	Pooled				100	**1.87 [0.80, 4.39]**		
Any cartilage involvement (CI)	Hermans et al. (1)[22]	-	119	36	24.7	0.98 [0.50, 1.93]		
	Hermans et al. (2)[23]	-	103	36	39.0	1.78 [1.04, 3.06]		
	Scherl et al.[31]	-	236	39	36.3	1.22 [0.70, 2.14]		
	Pooled				100	**1.34 [0.96, 1.88]**		
Vocal cord involvement (VCI)	Hermans et al. (1)[22]	-	119	82	20.7	1.79 [0.59, 5.37]		
	Hermans et al. (2)[23]	-	103	45	79.3	1.77 [1.01, 3.10]		
	Pooled				100	**1.77 [1.08, 2.92]**		

Abbreviations: CI, confidence interval; CT, computed tomography; HR, hazard ratio; IV, inverse variance; RR, risk ratio.

MRI (Figure 4). Most parameters show a great variability across studies, with a few exceptions. All studies that analyzed cricoid cartilage involvement, extralaryngeal spread beyond anterior commissure, and posterior commissure involvement found a high specificity (90%–100%) and a low sensitivity (0%–30%). Findings for any cartilage involvement were also fairly consistent, with a specificity between 60% and 70% and a sensitivity between 30% and 50%.

4 | DISCUSSION

In this systematic review and meta-analysis, the prognostic value of quantitative imaging variables was assessed for recurrent laryngeal or hypopharyngeal carcinoma after radiotherapy.

4.1 | Tumor volume

TV is one of the most studied prognostic imaging factors and this review supports that patients with higher TV have an increased risk of recurrence. All cut-off points were a prognostic factor, except for TV >4 cm³, which had inconsistent evidence. However, only one study that used 4 cm³ as a cut-off point reported a neutral effect size, while the other four found an increased risk of recurrence for TV >4 cm³. Therefore, it seems likely that the first study is an anomaly.

We also attempted to find an optimal TV cut-off point that can be used to predict recurrences. This cut-off point seems to be lower for glottic tumors compared to supraglottic tumors. However, there is a lot of variability in sensitivity and specificity of studies that used the same cut-off points. This can at least partially be explained by

TABLE 3   Prognostic factors on MRI [Color figure can be viewed at wileyonlinelibrary.com]

Prognostic factors	Study	Tumor volume cut-off [cm³]	Sample size, n	Factor prevalence [%]	Weight [%]	Univariate RR [95% CI] IV, random effects	Forest plot Univariate RR with 95% CI IV, random effects	Multivariate HR [95% CI]
Moderate level of evidence								
Tumor volume >3 cm³	Ljumanovic et al. (1)[36]	3	84	80	12.1	3.55 [0.94, 13.46]		
	Ljumanovic et al. (2)[37]	3	118	28	87.9	2.32 [1.41, 3.80]		
	Pooled				100	**2.44 [1.54, 3.88]**		
Pre-epiglottic space involvement (PEGS)	Ljumanovic et al. (1)[36]	-	84	62	35.7	4.15 [1.60, 10.78]		4.5 [1.0, 6.9]
	Ljumanovic et al. (2)[37]	-	118	8	64.3	2.02 [1.13, 3.63]		
	Pooled				100	**2.62 [1.33, 5.14]**		
Limited level of evidence								
Anterior commissure involvement (ACI)	Ljumanovic et al. (1)[36]	-	84	18	45.2	1.40 [0.74, 2.64]		2.5 [1.7, 9.8]
	Ljumanovic et al. (2)[37]	-	118	39	54.8	2.50 [1.48, 4.25]		
	Pooled				100	**1.93 [1.09, 3.40]**		
Subglottic extension (SGE)	Ljumanovic et al. (1)[36]	-	84	31	51.2	1.41 [0.81, 2.45]		
	Ljumanovic et al. (2)[37]	-	118	47	48.8	2.48 [1.39, 4.44]		
	Ljumanovic et al. (3)[38]	-	64	39	-			3.5 [1.6-7.9]
	Pooled				100	**1.86 [1.07, 3.23]**		
Hypopharyngeal extension (HE)	Ljumanovic et al. (1)[36]	-	84	31	45.7	1.49 [0.85, 2.62]		
	Ljumanovic et al. (2)[37]	-	118	6	54.3	2.88 [1.90, 4.37]		2.7 [1.0-7.4]
	Pooled				100	**2.13 [1.12, 4.07]**		
Extralaryngeal spread beyond anterior commissure (ELA)	Ljumanovic et al. (1)[36]	-	84	8	38.8	1.18 [0.48, 2.91]		
	Ljumanovic et al. (2)[37]	-	118	8	61.2	2.73 [1.73, 4.33]		
	Pooled				100	**1.97 [0.88, 4.41]**		

Abbreviations: CI, confidence interval; IV, inverse variance; HR, hazard ratio; MRI, magnetic resonance imaging; RR, risk ratio.

heterogeneity in study populations. Sensitivity and specificity are influenced by the prevalence of a prognostic factor within a study population. If a population is more likely to contain larger tumors (above the cut-off point), the sensitivity will be higher and the specificity will be lower. Therefore heterogeneity in for example T classification distribution can cause a lot of variability. In order to gain more insight into the differences between patient groups, studies should use more homogenous patient cohorts or stratify their results by tumor location and T classification.

Besides heterogeneity in study populations, a lack of uniformity in TV delineation might also complicate the results. Differences in image quality, measurement methods, slice thickness, and delineation guidelines can lead to different conclusions on TV. Moreover, interobserver variability of TV delineation in the head and neck region on both CT and MRI is still a much debated topic. With some studies finding no significant variability between observers,[39,40] while others do.[41-43]

Moreover, Ligtenberg et al.[44] found that delineations on CT, MRI and FDG-PET were all larger than the reference TV delineation on pathology. Between modalities, significant differences were found between CT and MRI and between CT and FDG-PET. This indicates that CT measurements cannot be compared to MRI measurements when it comes to TV.

In the current TNM classification of laryngeal carcinoma, TV is not taken into account.[45] For hypopharyngeal carcinoma, tumor dimension is considered in the form of the one-dimensional greatest tumor diameter.[45] For head and neck sites in general, previous reviews have plead for the addition of TV to the TNM classification system,[46,47] arguing that TV assessment should at the very least supplement TNM classification in clinical decision-making. Although our analysis does indicate a significant impact of TV on recurrences after radiotherapy, only few studies included a multivariable analysis with both the T classification and TV. Therefore, it is difficult to determine the added prognostic value of TV.

4.2 | Anatomical subsite involvement

The present study found strong evidence that involvement of the anterior commissure (ACI) is a prognostic factor of local recurrence after radiotherapy. This is in line with two recent reviews studying ACI in early glottic carcinoma by Tulli et al.[48] and Eskiizmir et al.[49] A third review by Hendriksma et al.[50] found conflicting evidence, arguing that ACI should not be judged as a binary

TABLE 4 Factors with inconsistent evidence on CT

Inconsistent evidence	Study	Tumor volume cut-off [cm³]	Sample size, n	Factor prevalence [%]	Weight [%]	Univariate RR [95% CI] IV, random effects	Forest plot Univariate RR with 95% CI IV, random effects	Multivariate HR [95% CI]
Tumor volume >4 cm³	Hermans et al. (1)[22]	4	119	12	20.7	2.05 [1.01, 4.16]		
	Hermans et al. (2)[23]	4	103	72	16.1	2.74 [1.05, 7.13]		
	Janssens et al.[24]	3.5	114	42	22.6	0.87 [0.47, 1.61]		
	Pameijer et al.[29]	3.5	42	38	16.3	4.88 [1.90, 12.54]		
	Rutkowski et al.[30]	3.6	115	17	24.3	2.38 [1.38, 4.07]		
	Pooled				*100*	*2.11 [1.23, 3.63]*		
Arytenoid cartilage invasion (ArCI)	Hermans et al. (1)[22]	-	119	31	35.4	0.89 [0.43, 1.83]		
	Hermans et al. (2)[23]	-	103	34	54.0	1.73 [1.01, 2.95]		
	Murakami et al.[27]	-	68	3	10.6	2.20 [0.51, 9.43]		
	Pooled				*100*	*1.40 [0.86, 2.28]*		
Thyroid cartilage involvement (TCI)	Hermans et al. (1)[22]	-	119	8	33.0	0.94 [0.26, 3.34]		
	Hermans et al. (2)[23]	-	103	4	25.7	0.75 [0.13, 4.19]		
	Murakami et al.[27]	-	68	22	41.3	4.54 [2.03, 10.15]		
	Pooled				*100*	*1.70 [0.49, 5.96]*		

Abbreviations: CI, confidence interval; CT, computed tomography; HR, hazard ratio; IV, inverse variance; RR, risk ratio.
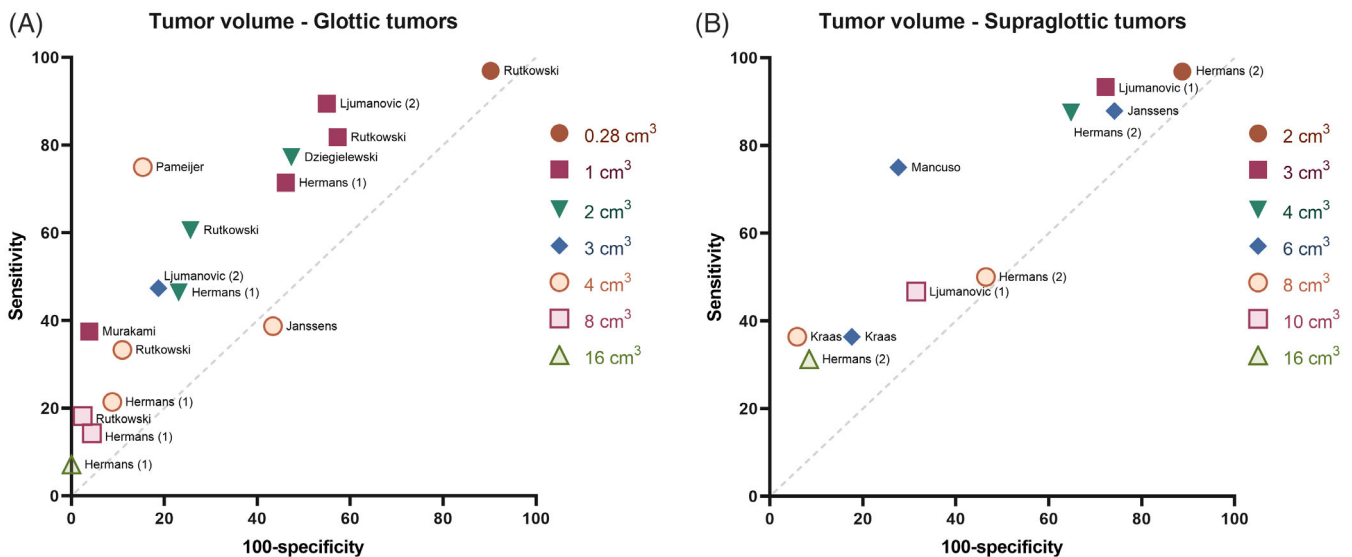
FIGURE 3 Sensitivity and specificity of different tumor volume cut-off points for predicting tumor recurrence. Each data point represents the results of one study with (A) glottic tumors and (B) supraglottic tumors [Color figure can be viewed at wileyonlinelibrary.com]

variable, but rather on a classification scale.[51] In the present review, all included studies treated ACI as a binary variable.

We found strong evidence of subglottic extension as a prognostic factor. It must be noted however, that the definition of subglottic extension is ambiguous. The boundaries of the subglottis itself are controversial and most publications use different definitions.[52] No study included in this review provided any definition for subglottic extension.

A third anatomical prognostic factor for which we found strong evidence was arytenoid cartilage adjacent or invaded (ArCAI), while just arytenoid cartilage invasion (ArCI) gave inconsistent evidence. The difference between the two parameters is that both ArCAI and ArCI include tumors that have invaded the cartilage, but ArCAI also includes the tumors that have grown adjacent to it. These results suggest that the risk of recurrence is already higher when the tumor has reached the arytenoid cartilage and that actual invasion does not necessarily increase that risk further. It is possible that patients with invaded cartilage received a more rigorous treatment regime, but this was not mentioned, but this was not mentioned in the articles that made the distinction. More

FIGURE 4 Sensitivity and specificity of anatomical subsites variables on CT or MRI for predicting tumor recurrence. Each data point represents the results of one study. ACI, anterior commissure involvement; ArCAI, arytenoid cartilage adjacent or invaded; CI, any cartilage involvement; CCI, cricoid cartilage involvement; ELA, extralaryngeal spread beyond anterior commissure; HE, hypopharyngeal extension; PAGS, paraglottic space involvement; PCI, posterior commissure involvement; PEGS, pre-epiglottic space involvement; SGE, subglottic extension; VCI, vocal cord involvement; VEN, laryngeal ventricle involvement [Color figure can be viewed at wileyonlinelibrary.com]
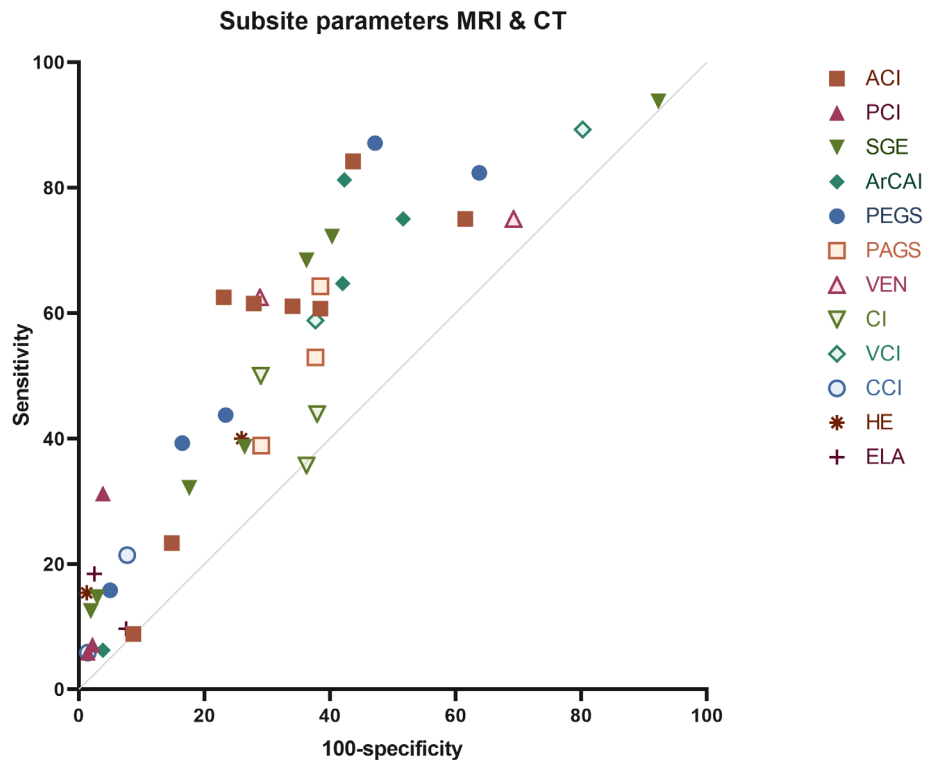
research into the prognostic value of cartilage adjacent tumors is needed in order to draw any conclusions.

The current TNM classification system[45] for all laryngeal subsites as well as for the hypopharynx takes into account infiltrated subsites, like pre-epiglottic and paraglottic space invasion, and thyroid and cricoid cartilage involvement. The inclusion of these variables is supported by this review, except for thyroid cartilage involvement, for which we found inconsistent evidence. However, the two studies by Hermans et al.[22,23] that reported a neutral effect estimate both had a very low prevalence of thyroid cartilage involvement within their study population (Table 4). It might therefore be an anomaly that these patients did relatively well.

As for the sensitivity and specificity of the anatomical variables as prognostic factors, most parameters showed great variability across studies. Three parameters had a high specificity and low sensitivity in all studies: posterior commissure involvement, cricoid cartilage involvement, and extralaryngeal spread beyond anterior commissure. In all study populations in which these variables were analyzed, the prevalence of these factors was relatively low, explaining the low sensitivity. However, because the specificity is so high, they might serve as a prognostic factor in those patients who do have these factors. Since posterior commissure was only analyzed in three studies and the other two factors in two studies, more research is necessary in order to draw definite conclusions.

4.3 | Other modalities/variables

Only 5 out of the 22 included articles studied MRI variables, all from the same research group. Due to overlap in patient pools, only the results of one or two articles could be compared for each variable. MR imaging in laryngeal carcinoma has its challenges due to relatively long acquisition times and significant organ movement. However, MRI techniques have improved a great deal over the last years and in the absence of motion artifacts, MRI is superior to CT in determining tumor extension. The imaging modality should not be overlooked when it comes to its prognostic value, especially more advanced MRI methods, like diffusion-weighted imaging (DWI) and dynamic contrast-enhanced (DCE) MRI.[53,54] On DWI-MRI, the apparent diffusion coefficient (ADC) has often been discussed as a prognostic factor in different head and neck sites, but mean or median ADC does often not reach significance.[54-56] More promising are ADC parameters that take into account tumor heterogeneity like ADC kurtosis.[54,57,58] On DCE-MRI, higher K^{trans}, a measure of vascular permeability, is associated with better treatment response and long-term outcome in various head and neck sites.[53,54,59,60]

The reason so few studies with FDG-PET variables were included was because most studies we found used other outcome measures than local control. FDG-PET is also not routinely used for pretreatment staging and is more commonly used in diagnostic studies than

prognostic ones. This does not mean that FDG-PET is not a promising modality for the prediction of radiotherapy outcome. High pretreatment metabolic tumor volume and total lesion glycolysis are both associated with worse patient outcomes in head and neck cancer.[61,62]

Furthermore, this review only looked at imaging variables, but clinical variables also play a role in patient prognosis. Factors like age, sex, tobacco use and WHO performance status have all been shown to affect patient outcome after radiotherapy in either laryngeal or hypopharyngeal carcinoma, or head and neck cancer in general.[49,63,64] These clinical factors should also be taken into account when making treatment-decisions.

4.4 | Limitations

For all variables included in the meta-analysis, the pooled RRs are based on few studies. Therefore, the uncertainty in the pooled RRs is large and they serve more as an indication than as conclusive RRs.

Another limitation is that only very few studies included a multivariable analysis; therefore, there is little control for possible confounders. For example, while TV was found to be a prognostic factor for recurrence, there is a correlation between TV and TNM classification. This review cannot control for this, due to the lack of multivariable analyses.

Publication bias might also play a role, as negative results are less likely to be published. Because of the limited number of studies per variable we could not test publication bias with a funnel plot.[65] There is also no formal way to assess selective reporting bias. If an article does not provide the crude data or effect estimates of non-significant results, their findings could not be used in this review.

There is heterogeneity between the tumor and treatment characteristics in the study populations, as well as in follow-up time and the outcome measures that were used (Table 1). The three outcome measures were the number of recurrences discovered during the follow-up time, the 2-year local control rate, and the 5-year local control rate. No differentiation was made based on the outcome, since most laryngeal and hypopharyngeal recurrences present within 2 years after treatment.[8,66] As for tumor characteristics, the prognostic value of different variables is likely to vary between different locations or tumor classifications, which might have affected the results. However, this review aimed to find prognostic imaging factors for all laryngeal and hypopharyngeal tumors and an inverse variance random effects model was applied to adjust for variation between studies. More research is needed to allow for further stratification.

We also noticed that only a few articles reported the time between the pretreatment imaging and the start of treatment (6 of the 23 studies articles). And when they did, this time could be quite long (one study reported an average of 43 days between imaging and start of treatment[38]). In this time period tumors will likely grow, causing discrepancies between the tumor on pretreatment imaging and the tumor at the start of treatment. Future studies assessing the prognostic value of imaging factors should report on the interval between imaging and treatment and aim to keep this time as short as possible.

5 | CONCLUSION

Pretreatment tumor volume determined on CT has clear prognostic value in laryngeal and hypopharyngeal carcinoma treated with radiotherapy. Due to heterogeneity in study populations and methodology, no conclusions can be drawn on optimal cut-off points.

On CT, there is strong evidence for the prognostic value of anterior and posterior commissure involvement, subglottic extension, and arytenoid cartilage extension. Moderate evidence was found for pre-epiglottic space involvement, paraglottic space involvement and cricoid cartilage involvement as prognostic factors on CT. On MRI, a moderate level of evidence was only found for tumor volume and pre-epiglottic space involvement. This is partly due to the low number of articles studying prognostic imaging variables on MRI. More research is needed in order to accurately assess the prognostic power of MR imaging variables.

ORCID
Hilde J. G. Smits https://orcid.org/0000-0002-7212-2694
Frans Kauw https://orcid.org/0000-0001-5604-0659
Marielle E. P. Philippens https://orcid.org/0000-0003-0205-3459
Remco de Bree https://orcid.org/0000-0001-7128-5814

REFERENCES
1. Dropkin MJ. Body image and quality of life after head and neck cancer surgery. *Cancer Pract.* 1999;7(6):309-313. https://doi.org/10.1046/j.1523-5394.1999.76006.x.
2. Braz DSA, Ribas MM, Dedivitis RA, Nishimoto IN, Barros APB. Quality of life and depression in patients undergoing total and partial laryngectomy. *Clinics (Sao Paulo).* 2005;60(2):135-142. https://doi.org/10.1590/S1807-59322005000200010.

3. Forastiere AA, Ismaila N, Lewin JS, et al. Use of larynx-preservation strategies in the treatment of laryngeal cancer: American society of clinical oncology clinical practice guideline update. *J Clin Oncol.* 2018;36(11):1143-1169. https://doi.org/10.1200/JCO.2017.75.7385.

4. O'Neill CB, O'Neill JP, Atoria CL, et al. Treatment complications and survival in advanced laryngeal cancer: a population based analysis. *Laryngoscope.* 2014;124(12):2707-2713. https://doi.org/10.1002/lary.24658.

5. Mannelli G, Lazio MS, Luparello P, Gallo O. Conservative treatment for advanced T3–T4 laryngeal cancer: meta-analysis of key oncological outcomes. *Eur Arch Oto-Rhino-Laryngol.* 2018;275(1):27-38. https://doi.org/10.1007/s00405-017-4799-x.

6. Chawla S, Carney AS. Organ preservation surgery for laryngeal cancer. *Head Neck Oncol.* 2009;1:12. https://doi.org/10.1186/1758-3284-1-12.

7. Brenner B, Marshak G, Sulkes A, Rakowsky E. Prognosis of patients with recurrent laryngeal carcinoma. *Head Neck.* 2001;23(7):531-535. https://doi.org/10.1002/hed.1074.

8. Li P, Hu W, Zhu Y, Liu J. Treatment and predictive factors in patients with recurrent laryngeal carcinoma: a retrospective study. *Oncol Lett.* 2015;10(5):3145-3152. https://doi.org/10.3892/ol.2015.3640.

9. Ganly I, Patel S, Matsuo J, et al. Postoperative complications of salvage total laryngectomy. *Cancer.* 2005;103(10):2073-2081. https://doi.org/10.1002/cncr.20974.

10. Sessions DG, Lenox J, Spector GJ. Supraglottic laryngeal cancer: analysis of treatment results. *Laryngoscope.* 2005;115(8):1402-1410. https://doi.org/10.1097/01.MLG.0000166896.67924.B7.

11. Hoffman HT, Porter K, Karnell LH, et al. Laryngeal cancer in the United States: changes in demographics, patterns of care, and survival. *Laryngoscope.* 2006;116(9):1-13. https://doi.org/10.1097/01.mlg.0000236095.97947.26.

12. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 2009;6(7):e1000097. https://doi.org/10.1371/journal.pmed.1000097.

13. Wells GA, Shea B, O'Connell D, et al. Newcastle-Ottawa quality assessment scale. *Ottawa Hosp Res Inst.* 2014;3:2-4. https://doi.org/10.2307/632432.

14. Post B, Merkus MP, De Haan RJ, Speelman JD. Prognostic factors for the progression of Parkinson's disease: a systematic review. *Mov Disord.* 2007;22(13):1839-1851. https://doi.org/10.1002/mds.21537.

15. De Rooij NK, Rinkel GJE, Dankbaar JW, Frijns CJM. Delayed cerebral ischemia after subarachnoid hemorrhage: a systematic review of clinical, laboratory, and radiological predictors. *Stroke.* 2013;44(1):43-54. https://doi.org/10.1161/STROKEAHA.112.674291.

16. Kauw F, Takx RAP, De Jong HWAM, Velthuis BK, Kappelle LJ, Dankbaar JW. Clinical and imaging predictors of recurrent ischemic stroke: a systematic review and meta-analysis. *Cerebrovasc Dis.* 2018;45(5–6):279-287. https://doi.org/10.1159/000490422.

17. Agarwal JP, Sinha S, Goda JS, et al. Tumor radiomic features complement clinico-radiological factors in predicting long-term local control and laryngectomy free survival in locally advanced laryngo-pharyngeal cancers. *Br J Radiol.* 2020;93(1109):20190857. https://doi.org/10.1259/bjr.20190857.

18. Van Bockel LW, Monninkhof EM, Pameijer FA, Terhaard CHJ. Importance of tumor volume in supraglottic and glottic laryngeal carcinoma. *Strahlentherapie und Onkol.* 2013;189(12):1009-1014. https://doi.org/10.1007/s00066-013-0467-2.

19. Chen SW, Yang SN, Liang JA, Tsai MH, Shiau AC, Lin FJ. Value of computed tomography-based tumor volume as a predictor of outcomes in hypopharyngeal cancer after treatment with definitive radiotherapy. *Laryngoscope.* 2006;116(11):2012-2017. https://doi.org/10.1097/01.mlg.0000237804.38761.81.

20. Dagan R, Morris CG, Bennett JA, et al. Prognostic significance of paraglottic space invasion in T2N0 glottic carcinoma. *Am J Clin Oncol.* 2007;30(2):186-190. https://doi.org/10.1097/01.coc.0000251403.54180.df.

21. Dziegielewski PT, Reschly WJ, Morris CG, et al. Tumor volume as a predictor of survival in T3 glottic carcinoma: a novel approach to patient selection. *Oral Oncol.* 2018;79:47-54. https://doi.org/10.1016/j.oraloncology.2018.02.015.

22. Hermans R, Van den Bogaert W, Rijnders A, Doornaert P, Baert AL. Predicting the local outcome of glottic squamous cell carcinoma after definitive radiation therapy: value of computed tomography-determined tumour parameters. *Radiother Oncol.* 1999;50(1):39-46. https://doi.org/10.1016/S0167-8140(98)00114-5.

23. Hermans R, Van Den Bogaert W, Rijnders A, Baert AL. Value of computed tomography as outcome predictor of supraglottic squamous cell carcinoma treated by definitive radiation therapy. *Int J Radiat Oncol Biol Phys.* 1999;44(4):755-765. https://doi.org/10.1016/S0360-3016(99)00039-5.

24. Janssens GO, van Bockel LW, Doornaert PA, et al. Computed tomography-based tumour volume as a predictor of outcome in laryngeal cancer: results of the phase 3 ARCON trial. *Eur J Cancer.* 2014;50(6):1112-1119. https://doi.org/10.1016/j.ejca.2013.12.012.

25. Kraas JR, Underhill TE, D'Agostino RB, Williams DW, Cox JA, Greven KM. Quantitative analysis from CT is prognostic for local control of supraglottic carcinoma. *Head Neck.* 2001;23(12):1031-1036. https://doi.org/10.1002/hed.10030.

26. Mancuso BAA, Mukherji SK, Schmalfuss I, et al. Preradiotherapy computed tomography as a predictor of local control in supraglottic carcinoma. *J Clin Oncol.* 1999;17(2):631-637. https://doi.org/10.1200/JCO.1999.17.2.631.

27. Murakami R, Furusawa M, Baba Y, et al. Dynamic helical CT of T1 and T2 glottic carcinomas: predictive value for local control with radiation therapy. *AJNR Am J Neuroradiol.* 2000;21(7):1320-1326.

28. Nur DA, Oguz C, Kemal ET, et al. Prognostic factors in early glottic carcinoma implications for treatment. *Tumori.* 2005;91(2):182-187. https://doi.org/10.1177/030089160509100215.

29. Pameijer FA, Mancuso AA, Mendenhall WM, Parsons JT, Kubilis PS. Can pretreatment computed tomography predict local control in t3 squamous cell carcinoma of the glottic larynx treated with definitive radiotherapy? *Int J Radiat Oncol Biol Phys.* 1997;37(5):1011-1021. https://doi.org/10.1016/S0360-3016(96)00626-8.

30. Rutkowski T. Impact of initial tumor volume on radiotherapy outcome in patients with T2 glottic cancer. *Strahlentherapie und Onkol.* 2014;190(5):480-484. https://doi.org/10.1007/s00066-014-0603-7.

31. Scherl C, Mantsopoulos K, Semrau S, et al. Management of advanced hypopharyngeal and laryngeal cancer with and

without cartilage invasion. *Auris Nasus Larynx*. 2017;44(3):333-339. https://doi.org/10.1016/j.anl.2016.08.002.

32. Tsou YA, Hua JH, Lin MH, Tsai MH. Analysis of prognostic factors of chemoradiation therapy for advanced hypopharyngeal cancer—Does tumor volume correlate with central necrosis and tumor pathology? *ORL J Otorhinolaryngol Relat Spec*. 2006;68(4):206-212. https://doi.org/10.1159/000091803.

33. Zouhair A, Azria D, Coucke P, et al. Decreased local control following radiation therapy alone in early-stage glottic carcinoma with anterior commissure extension. *Strahlentherapie und Onkol*. 2004;180(2):84-90. https://doi.org/10.1007/s00066-004-1164-y.

34. Castelijns JA, Van den Brekel MWM, Smit EMT, et al. Predictive value of MR imagining-dependent and non-MR imaging-dependent parameters for recurrence of laryngeal cancer after radiation therapy. *Radiology*. 1995;196(3):735-739. https://doi.org/10.1148/radiology.196.3.7644637.

35. Castelijns JA, van den Brekel MWM, Tobi H, et al. Laryngeal carcinoma after radiation therapy: correlation of abnormal MR imaging signal patterns in laryngeal cartilage with the risk of recurrence. *Radiology*. 1996;198(1):151-155. https://doi.org/10.1148/radiology.198.1.8539369.

36. Ljumanović R, Langendijk JA, Schenk B, et al. Supraglottic carcinoma treated with curative radiation therapy: identification of prognostic groups with MR imaging. *Radiology*. 2004;232(2):440-448. https://doi.org/10.1148/radiol.2322031001.

37. Ljumanovic R, Langendijk JA, Van Wattingen M, et al. MR imaging predictors of local control of glottic squamous cell carcinoma treated with radiation alone. *Radiology*. 2007;244(1):205-212. https://doi.org/10.1148/radiol.2441060593.

38. Ljumanovic R, Pouwels PJW, Langendijk JA, et al. Has the degree of contrast enhancement with MR imaging in laryngeal carcinoma added value to anatomic parameters regarding prediction of response to radiation therapy? *AJNR Am J Neuroradiol*. 2007;28(8):1540-1546. https://doi.org/10.3174/ajnr.A0599.

39. Geets X, Daisne JF, Arcangeli S, et al. Inter-observer variability in the delineation of pharyngo-laryngeal tumor, parotid glands and cervical spinal cord: comparison between CT-scan and MRI. *Radiother Oncol*. 2005;77(1):25-31. https://doi.org/10.1016/j.radonc.2005.04.010.

40. Mukherji SK, Toledano AY, Beldon C, et al. Interobserver reliability of computed tomography-derived primary tumor volume measurement in patients with supraglottic carcinoma. *Cancer*. 2005;103(12):2616-2622. https://doi.org/10.1002/cncr.21072.

41. Jager AA, Kasperts N, Caldas-Magalhaes J, et al. GTV delineation in supraglottic laryngeal carcinoma: Interobserver agreement of CT versus CT-MR delineation. *Radiat Oncol*. 2015;10(1):1-9. https://doi.org/10.1186/s13014-014-0321-4.

42. Hoorweg JJ, Kruijt RH, Heijboer RJJ, Eijkemans MJC, Kerrebijn JDF. Reliability of interpretation of CT examination of the larynx in patients with glottic laryngeal carcinoma. *Otolaryngol – Head Neck Surg*. 2006;135(1):129-134. https://doi.org/10.1016/j.otohns.2006.01.024.

43. Cooper JS, Mukherji SK, Toledano AY, et al. An evaluation of the variability of tumor-shape definition derived by experienced observers from CT images of supraglottic carcinomas

(ACRIN protocol 6658). *Int J Radiat Oncol Biol Phys*. 2007;67(4):972-975. https://doi.org/10.1016/j.ijrobp.2006.10.029.

44. Ligtenberg H, Jager EA, Caldas-Magalhaes J, et al. Modality-specific target definition for laryngeal and hypopharyngeal cancer on FDG-PET, CT and MRI. *Radiother Oncol*. 2017;123(1):63-70. https://doi.org/10.1016/j.radonc.2017.02.005.

45. Amin MB, Edge SB, Greene FL, et al. *AJCC Cancer Staging Manual*. 8th ed. New York: Springer; 2017.

46. Mendenhall WM, Mancuso AA, Strojan P, et al. Impact of primary tumor volume on local control after definitive radiotherapy for head and neck cancer. *Head Neck*. 2014;36(10):1391. https://doi.org/10.1002/hed.23454.

47. Rutkowski TW. The role of tumor volume in radiotherapy of patients with head and neck cancer. *Radiat Oncol*. 2014;9(1):1-9. https://doi.org/10.1186/1748-717X-9-23.

48. Tulli M, Re M, Bondi S, et al. The prognostic value of anterior commissure involvement in T1 glottic cancer: a systematic review and meta-analysis. *Laryngoscope*. 2020;130(8):1932-1940. https://doi.org/10.1002/lary.28395.

49. Eskiizmir G, Baskin Y, Yalcin F, Ellidokuz H, Ferris RL. Risk factors for radiation failure in early-stage glottic carcinoma: a systematic review and meta-analysis. *Oral Oncol*. 2016;62:90-100. https://doi.org/10.1016/j.oraloncology.2016.10.013.

50. Hendriksma M, Sjogren EV. Involvement of the anterior commissure in early glottic cancer (tis-T2): a review of the literature. *Cancers (Basel)*. 2019;11(9):1234. https://doi.org/10.3390/cancers11091234.

51. Rucci L, Gammarota L, Gallo O. Carcinoma of the anterior commissure of the larynx: II. Proposal of a new staging system. *Ann Otol Rhinol Laryngol*. 1996;105(5):391-396. https://doi.org/10.1177/000348949610500512.

52. Coskun H, Mendenhall WM, Rinaldo A, et al. Prognosis of subglottic carcinoma: Is it really worse? *Head Neck*. 2019;41(2):511-521. https://doi.org/10.1002/hed.25172.

53. King AD, Thoeny HC. Functional MRI for the prediction of treatment response in head and neck squamous cell carcinoma: potential and limitations. *Cancer Imaging*. 2016;16(1):1-8. https://doi.org/10.1186/s40644-016-0080-6.

54. Dai YL, King AD. State of the art MRI in head and neck cancer. *Clin Radiol*. 2018;73(1):45-59. https://doi.org/10.1016/j.crad.2017.05.020.

55. Peltenburg B, Driessen JP, Vasmel JE, et al. Pretreatment ADC is not a prognostic factor for local recurrences in head and neck squamous cell carcinoma when clinical T-stage is known. *Eur Radiol*. 2020;30(2):1228-1231. https://doi.org/10.1007/s00330-019-06426-y.

56. Chawla S, Kim S, Dougherty L, et al. Pretreatment diffusion-weighted and dynamic contrast-enhanced MRI for prediction of local treatment response in squamous cell carcinomas of the head and neck. *Am J Roentgenol*. 2013;200(1):35-43. https://doi.org/10.2214/AJR.12.9432.

57. Law BKH, King AD, Bhatia KS, et al. Diffusion-weighted imaging of nasopharyngeal carcinoma: can pretreatment DWI predict local failure based on long-term outcome? *Am J Neuroradiol*. 2016;37(9):1706-1712. https://doi.org/10.3174/ajnr.A4792.

58. Preda L, Conte G, Bonello L, et al. Combining standardized uptake value of FDG-PET and apparent diffusion coefficient of DW-MRI improves risk stratification in head and neck

squamous cell carcinoma. *Eur Radiol.* 2016;26(12):4432-4441. https://doi.org/10.1007/s00330-016-4284-8.

59. Chawla S, Kim S, Loevner LA, et al. Prediction of disease-free survival in patients with squamous cell carcinomas of the head and neck using dynamic contrast-enhanced MR imaging. *Am J Neuroradiol.* 2011;32(4):778-784. https://doi.org/10.3174/ajnr.A2376.

60. Shukla-Dave A, Lee NY, Jansen JFA, et al. Dynamic contrast-enhanced magnetic resonance imaging as a predictor of outcome in head-and-neck squamous cell carcinoma patients with nodal metastases. *Int J Radiat Oncol Biol Phys.* 2012;82(5):1837-1844. https://doi.org/10.1016/j.ijrobp.2011.03.006.

61. Pak K, Cheon GJ, Nam HY, et al. Prognostic value of metabolic tumor volume and total lesion glycolysis in head and neck cancer: a systematic review and meta-analysis. *J Nucl Med.* 2014;55 (6):884-890. https://doi.org/10.2967/jnumed.113.133801.

62. Bonomo P, Merlotti A, Olmetto E, et al. What is the prognostic impact of FDG PET in locally advanced head and neck squamous cell carcinoma treated with concomitant chemo-radiotherapy? A systematic review and meta-analysis. *Eur J Nucl Med Mol Imaging.* 2018;45(12):2122-2138. https://doi.org/10.1007/s00259-018-4065-5.

63. Bradford CR, Ferlito A, Devaney KO, Mäkitie AA, Rinaldo A. Prognostic factors in laryngeal squamous cell carcinoma. *Laryngoscope Investig Otolaryngol.* 2020;5(1):74-81. https://doi.org/10.1002/lio2.353.

64. Economopoulou P, de Bree R, Kotsantis I, Psyrri A. Diagnostic tumor markers in head and neck squamous cell carcinoma (HNSCC) in the clinical setting. *Front Oncol.* 2019;9:1-13. https://doi.org/10.3389/fonc.2019.00827.

65. Lau J, Ioannisid JPA, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. *Br Med J.* 2006;333(7568):597-600. https://doi.org/10.1136/bmj.38946.501215.68.

66. Kothari P, Trinidade A, Hewitt RJD, Singh A, O'Flynn P. The follow-up of patients with head and neck cancer: an analysis of 1,039 patients. *Eur Arch Oto-Rhino-Laryngol.* 2011;268(8):1191-1200. https://doi.org/10.1007/s00405-010-1461-2.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.