

Chapter 1

Detailed patient-individual reporting of lymph node involvement in oropharyngeal squamous cell carcinoma with an online interface

One critical aspect of our effort to model and predict the lymphatic tumor progression is the data we use to train the model. As previously explained, our model essentially consumes tables with rows of patients and columns involvement by lymph node level (LNL). Data in this relatively simple format has been extracted in the past to create studies like [4] or [25]. However, the authors then used the data to compute statistics of it – e.g. the prevalence of involvement – but stopped short of publishing that data in its raw format. From these statistics it is – with one exception [24] – usually not possible to reconstruct the correlations between the involvement of LNLs.

With almost no usable data, of course, our methodology for modelling lymphatic progression cannot be tested or applied. So, we decided to start at the University Hospital Zurich (USZ) to extract all patterns of lymphatic progression in patients with newly diagnosed oropharyngeal squamous cell carcinoma (OP-SCC) between 2013 and 2019. We then not only used that data for inference on it, but also published it freely, hoping that other researchers might find it useful and that it may even motivate them to share their data in a similar fashion in the future.

In the following, I will include large parts of the publication [18], in which we detailed the extraction of the dataset, its characteristics and how we made it available. It is important to note that the first authorship is shared in this publication: **Jean-Marc Hoffmann**, a radiation oncologist at the USZ, extracted most of the data from digital patient and imaging records. **Bertrand Pouymayou**, a medical physicist and postdoctoral researcher at the USZ built up a complex template for easier extraction and storage of the patient information. He also created the initial interface for viewing the data. My contribution to this work was the processing of the data, creating figures and tables for the publication, host the cohort in the form of a comma separated values (CSV) table in online repositories and, lastly,

develop and deploy an online interface akin to what Bertrand Pouymayou had implemented earlier (see ?? for more details).

1.1 Abstract

Purpose/Objective

Whereas the prevalence of LNL involvement in head and neck squamous cell carcinoma (HNSCC) has been reported, the details of lymphatic progression patterns are insufficiently quantified. In this study, we investigate how the risk of metastases in each LNL depends on the involvement of upstream LNLs, T-category, Human Papillomavirus (HPV) status and other risk factors.

Results

We retrospectively analyzed patients with newly diagnosed OPSCC treated at a single institution, resulting in a dataset of 287 patients. For all patients, involvement of LNLs I-VII was recorded individually based on available diagnostic modalities (positron emission tomography (PET), magnetic resonance imaging (MRI), computed tomography (CT), fine needle aspiration (FNA)) together with clinicopathological factors. To analyze the dataset, a web-based graphical user interface (GUI) was developed, which allows querying the number of patients with a certain combination of co-involved LNLs and tumor characteristics.

Results

The full dataset and GUI is part of the publication. Selected findings are: Ipsilateral level IV was involved in 27% of patients with level II and III involvement, but only in 2% of patients with level II but not III involvement. Prevalence of involvement of ipsilateral levels II, III, IV, V was 79%, 34%, 7%, 3% for early T-category patients (T1/T2) and 85%, 50%, 17%, 9% for late T-category (T3/T4), quantifying increasing involvement with T-category. Contralateral levels II, III, IV were involved in 41%, 19%, 4% and 12%, 3%, 2% for tumors with and without midline extension, respectively. T-stage dependence of LNL involvement was more pronounced in HPV negative than positive tumors, but overall involvement was similar. Ipsilateral level VII was involved in 14% and 6% of patients with primary tumors in the tonsil and the base of tongue, respectively.

Conclusions

Detailed quantification of LNL involvement in HNSCC depending on involvement of upstream LNLs and clinicopathological factors may allow for further personalization of elective clinical target volume (CTV-N) definition in the future.

1.2 Introduction

HNSCC spread through the lymphatic system of the neck and form metastases in regional lymph nodes. Therefore, the target volume in radiotherapy of HNSCC

patients includes, in addition to the primary tumor, parts of the lymph drainage volume [2], [12]. The nodal gross tumor volume nodal gross tumor volume (GTV-N) contains detectable macroscopic lymph node metastases, while the elective clinical target volume CTV-N contains parts of lymph drainage volume that is at risk of harboring microscopic tumor, i.e. occult metastases that are not yet visible with current imaging techniques.

GTV-N definition is primarily performed through imaging techniques (PET-CT/MRI, MRI or CT) as well as FNA. Imaging criteria for lymph node metastases include size, round rather than oval shape, central necrosis, and FDG uptake as summarized by Biau et al [2]. Goel et al. gives an overview over clinical practice in PET/CT for the management of HNSCC [8]. However, all imaging techniques have finite sensitivity and specificity [21], [15], [23], i.e. they fail to detect small metastases or may incorrectly identify suspicious lymph nodes as tumor.

For standardized reporting of the location of lymph node metastases as well as delineation of the CTV-N, the lymph drainage system of the neck is divided into anatomically defined regions called LNL [10]. CTV-N definition amounts to the decision which LNLs to include into the elective CTV-N and is based on international consensus guidelines. Such guidelines were first published by Grégoire et al in 2000 and have been updated in 2006, 2014 and 2019 [2], [10], [9], [11]. Current recommendations for the selection of lymph node levels in OPSCC can be found in Table 2 of the guidelines published in 2019 by Biau et al. [2]. Current guidelines are primarily based on the prevalence of LNL involvement for a given primary tumor location, i.e. the percentage of patients diagnosed with metastases in each level. It is recommended that the elective CTV-N includes all LNLs that are involved in 10–15% of patients or more. Patients are primarily stratified by primary tumor location. For example, tumors of the soft palate, the posterior pharyngeal wall and the base of tongue show lymph node metastases on both sides via crossing lymph vessels. For this reason, even for lateralized tumors of these localizations, bilateral neck treatment is recommended. However, the lymphatic drainage of the tonsil is mainly unilateral, therefore an ipsilateral irradiation is recommended for lateralized low T-category (T1/T2) tumors (at least up to lymph node stage N2a). Volume-reduced elective nodal irradiation has been or is being investigated in several trials [3], [19].

While the general patterns of lymph drainage in the neck is understood and prevalence of LNL involvement has been reported in the literature [9], [4], [14], [1], the details of progression patterns in OPSCC are poorly quantified. How much does the risk of level IV involvement increase depending on whether levels II and III harbors macroscopic metastases? How much does the risk of involvement increase for late versus early T-category? Are progression patterns different for HPV positive versus HPV negative tumors? Answering these questions quantitatively may allow for further personalizing CTV-N definition based on an individual patient's clinical presentation at the time of diagnosis.

The basis for better quantification of LNL involvement are detailed datasets of HNSCC patients for whom involvement is reported per individual LNL together with tumor and patient characteristics. For example, to answer the question of how much the risk in level IV increases depending on the involvement of upstream levels II and III, it is insufficient to only report the prevalence of LNL involvement in levels II, III, and IV. Instead, the observed frequency of certain involvement

combinations must be known, e.g. how often levels II, III and IV are involved simultaneously, versus how often only the levels II and III are involved without level IV. The contributions of this work are:

- We provide a dataset of lymphatic progression patterns in 287 OPSCC patients treated at our institution in whom involvement of LNLs together with tumor characteristics are reported on a patient-individual basis.
- To visualize and explore the complex dataset, a graphical user interface is provided that allows the user to query the number of patients who were diagnosed with a specific combination of simultaneously involved LNLs and tumor characteristics.

We hope that this work provides the basis for collecting large multicenter datasets of lymphatic progression patterns, which can then inform future guidelines on further personalized CTV-N definition.

1.3 Material & methods

Data curation

We included patients diagnosed with OPSCC (primary diagnosis) between 2013 and 2019 and treated at the department of radiation oncology and/or head and neck surgery of the USZ. Patients with prior radiotherapy or surgery to the neck were excluded, resulting in a dataset of 287 patients. Specific subsites of oropharyngeal cancer included the base of tongue, the tonsils as well as the oropharyngeal side of the vallecula and the posterior or lateral wall of the oropharynx. Patient information consisted of the date of birth, gender, the date of the 1st histological confirmation of the tumor, the performed treatment (surgery with neck dissection prior to RT/RCHT vs. surgery only vs. definitive radio(chemo)therapy), risk factors such as nicotine abuse and HPV-status (p16 pos/neg), the TNM-classification (UICC 7th edition until 2017, 8th edition since 2017), the position of the primary tumor (left/right neck) as well as positive vs. negative mid-sagittal plane extension. Further details are described in the accompanying data-in-brief article [16].

The analysis of the lymphatic spread included levels Ia, Ib, IIa, IIb, III, IV, V, VII and was performed separately for the diagnostic imaging modalities available for a patient (FDG PET-CT, FDG PET-MRI, MRI, CT) as well as FNA and radiotherapy planning CT if available. This was performed by 2 experienced radiation oncologists by reviewing radiology and pathology reports together with the diagnostic images. Criteria for considering a lymph node as malignant followed the description in Biau et al [2] and are described in detail in the data-in-brief article [16].

Data base

The full dataset is available as a CSV-file via the data-in-brief article linked to this publication [16] and on GitHub at <https://github.com/rmnldwg/lydata> in a folder named 2021-usz-oropharynx.

In addition, the dataset has been archived and given a persistent identifier: <https://doi.org/10.5281/zenodo.6024778>.

Graphical user interface

We developed an online GUI based on the Django framework [7] and provide it to explore the dataset. It allows the user to conveniently determine the number of patients that show a particular combination of co-involved LNLs and tumor characteristics. The GUI is available at <https://2021-oropharynx.lyprox.org>; its source code under MIT license is available on GitHub at <https://github.com/rmnlwdwg/lyprox>. Documentation is provided within the GUI; a video demonstrating the use of the GUI is available online.

1.4 Results

In this section, we summarize selected findings. To that end, LNL involvement based on CT, MRI, PET and FNA was converted into a consensus decision via a logical OR, i.e. a LNL is considered involved if it was positive for one of these 4 modalities. However, only a small fraction of the information contained in the full dataset can be summarized here in tables. Thus, the interested reader is encouraged to explore the full dataset directly.

Graphical user interface

Figure 1.1 shows the GUI for analyzing the dataset. The main patient characteristics (smoking status, HPV status, and primary treatment) are shown on the top left. In the bottom-left panel, the user can specify characteristics of the primary tumor. In the example, the user considers all subsites combined, but restricts the patient selection to late T-category tumors (T3/T4) that extend over the midsagittal plane. On the top-right, the user selected the diagnostic modalities CT, MRI, PET and FNA, which are connected via a logical OR, i.e. a lymph node level is considered as involved for a patient if it was considered positive for at least one of the modalities available for that patient. The main panel shows the involvement of LNLs. In the example, the user restricts the selection to patients with positive ipsilateral level III while all other levels are unspecified. In total, 38 out of 287 patients meet these selection criteria. The main panel now displays the number of patients with involvement of the other levels. E.g. 21 patients have contralateral level II involved, and 12 patients ipsilateral level IV.

Prevalence

Overall prevalence of lymph node level involvement is reported in table 1.1 and visualized in fig. 1.2 panel a and b.

Dependence on T-category

Table 1.1 and fig. 1.2 panels a and b compare the prevalence of LNL involvement for early (T1/T2) versus late (T3/T4) T-category-patients. Consistent with common intuition, higher involvement was observed for late T-categories. Involvement of ipsilateral level II was high also for T1/T2 & (79%) and therefore increased only moderately for T3/T4 & (83%). However, involvement of the downstream levels

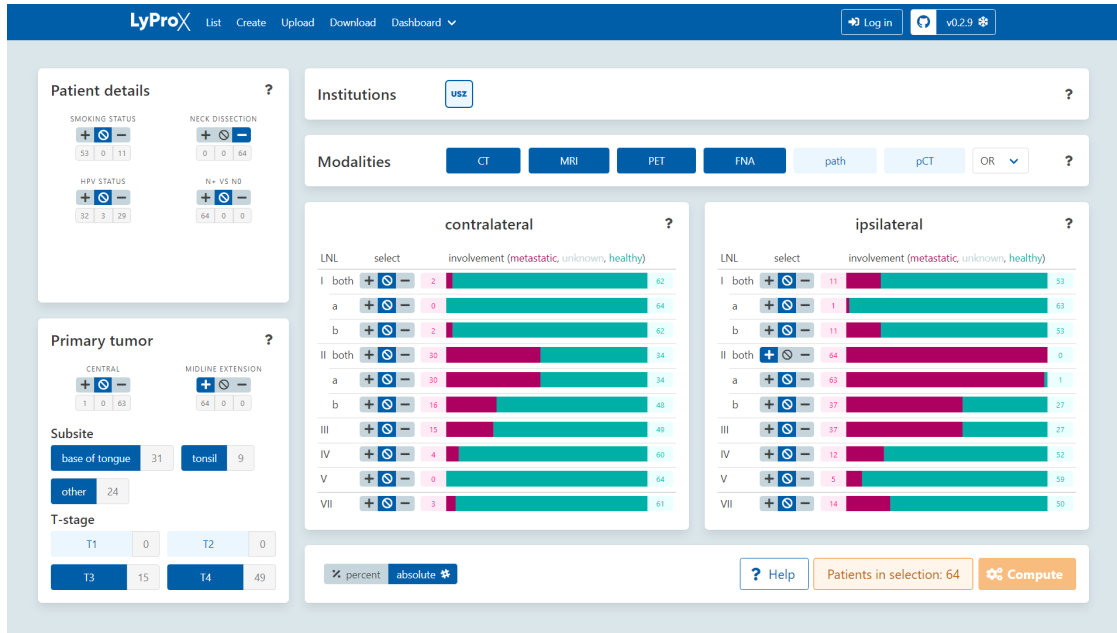


Figure 1.1: Graphical user interface to analyze lymphatic metastatic progression patterns.

Table 1.1: Prevalence of LNL involvement for the whole patient cohort (all) and stratified according to early (T1/T2) versus late (T3/T4) T-category and HPV positive (HPV +) versus HPV negative (HPV-) tumors. For each LNL, the first column indicates the number of patients showing involvement in the level, the second column the percentage of positive patients in the respective group.

	n	I		II		III		IV		V		VII		
ipsi	all	287	30	10%	232	81%	113	39%	31	11%	16	6%	30	10%
	T1/T2	150	10	7%	118	79%	48	32%	10	7%	4	3%	12	8%
	T3/T4	137	20	15%	114	83%	65	47%	21	15%	12	9%	18	13%
	HPV+	181	20	11%	155	86%	73	40%	20	11%	12	7%	18	10%
	HPV−	96	8	8%	69	72%	37	39%	11	11%	4	4%	12	13%
	HPV+, T1/T2	100	7	7%	86	86%	35	35%	8	8%	3	3%	10	10%
	HPV+, T3/T4	81	13	16%	69	85%	38	47%	12	15%	9	11%	8	10%
	HPV−, T1/T2	43	2	5%	27	63%	11	26%	2	5%	1	2%	2	5%
	HPV−, T3/T4	53	6	11%	42	79%	26	49%	9	17%	3	6%	10	19%
contra	all	287	3	1%	51	18%	21	7%	7	2%	2	1%	6	2%
	T1/T2	150	0	0%	13	9%	4	3%	2	1%	1	1%	1	1%
	T3/T4	137	3	2%	38	28%	17	12%	5	4%	1	1%	5	4%
	HPV+	181	3	2%	26	14%	13	7%	6	3%	2	1%	3	2%
	HPV−	96	0	0%	25	26%	8	8%	1	1%	0	0%	3	3%

III, IV, V increased from 32%, 7%, 3% for early T-category patients to 47%, 15%, 9% for late T-category. On the contralateral side, involvement of levels II, III, IV, V increased from 9%, 3%, 1%, 1% for T1/T2 patients to 28%, 12%, 4%, 1% for T3/T4.

Dependence on upstream levels

Table 1.2 considers the frequency of involvement in downstream levels depending on the involvement in upstream levels. On the ipsilateral side, level III harbored

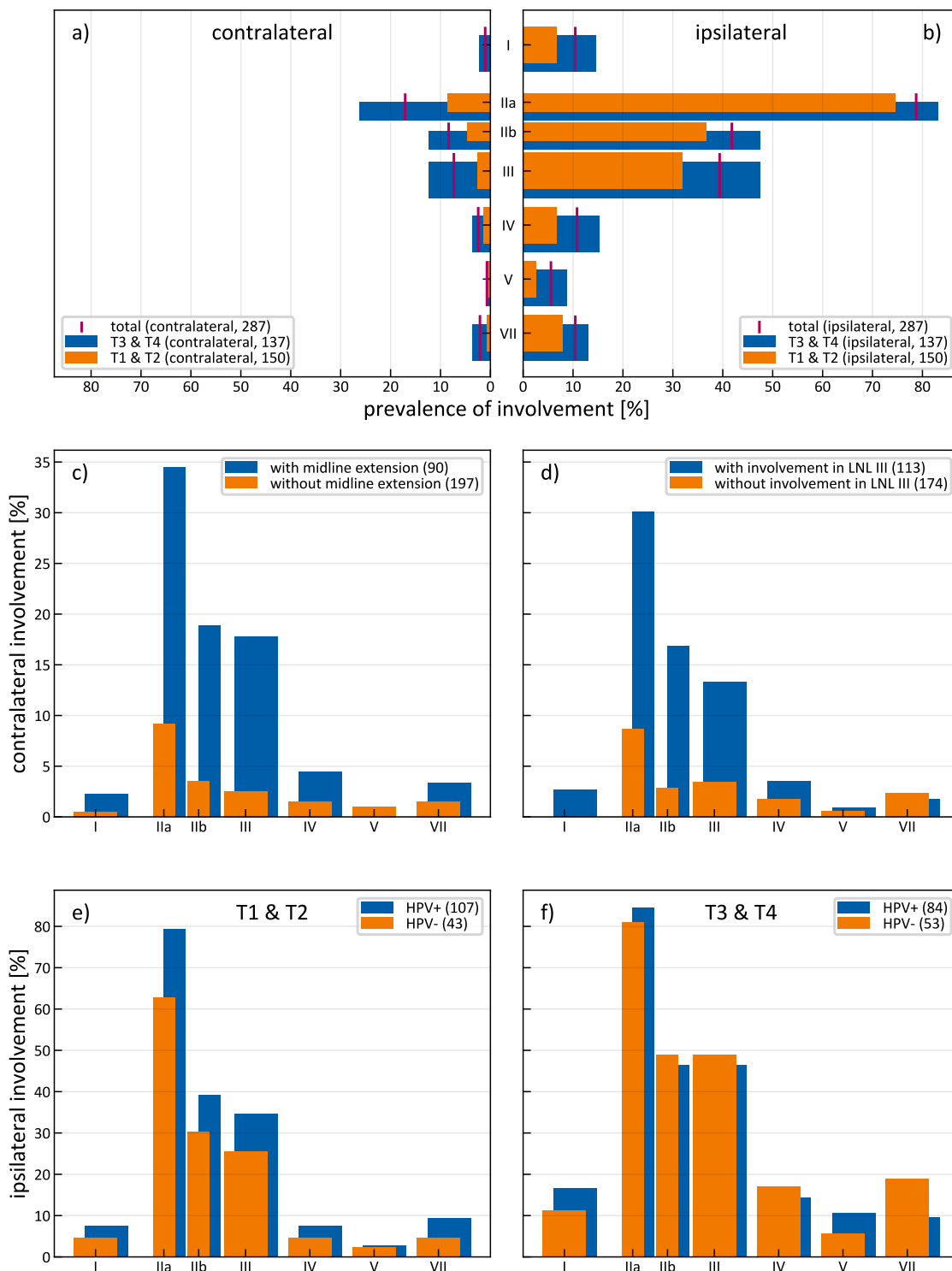


Figure 1.2: (a) Contralateral and (b) ipsilateral prevalence of LNL involvement for the whole patient cohort and stratified according to early (T1/T2) versus late (T3/T4) T-category. Contralateral LNL involvement stratified according to (c) midsagittal plan extension and (d) involvement of ipsilateral level III. Ipsilateral LNL involvement stratified according to HPV status for T1/T2 tumors (e) and T3/T4 tumors (f).

Table 1.2: Simultaneous involvement in levels II, III, and IV and frequency of skip metastases, for the whole patient cohort (all) and stratified according to early (T1/T2) versus late (T3/T4) T-category. Columns 1-3 define the 8 possible combinations of involvement; subsequent columns report the number of patients with the respective combination of co-involved levels.

II	III	IV	ipsi						contra					
			all		T1/T2		T3/T4		all		T1/T2		T3/T4	
+	+	+	28	10%	8	5%	20	15%	5	2%	1	1%	4	3%
+	+	-	80	28%	37	25%	43	31%	13	5%	2	1%	11	8%
+	-	+	2	1%	2	1%	0	0%	0	0%	0	0%	0	0%
+	-	-	122	43%	71	47%	51	37%	33	11%	10	7%	23	17%
-	+	+	0	0%	0	0%	0	0%	1	0%	1	1%	0	0%
-	+	-	5	2%	3	2%	2	1%	2	1%	0	0%	2	1%
-	-	+	1	0%	0	0%	1	1%	1	0%	0	0%	1	1%
-	-	-	49	17%	29	19%	20	15%	232	81%	136	91%	96	70%
			287		150		137		287		150		137	

metastases in 47% of patients (108 out of 232) when level II was positive, but in only 9% of patients (5 out of 55) when II was negative. Analogously, level IV harbored metastases in 25% of patients (28 out of 118) when level III was positive, but in only 2% of patients (3 out of 174) when III was negative (fig. 1.3). On the contralateral side, level III harbored metastases in 35% of patients (18 out of 51) when level II was positive, but in only 1% of patients (3 out of 236) when II was negative.

Contralateral involvement

Apart from late T-category (table 1.2), extension of the primary tumor across the midsagittal plane and higher ipsilateral involvement was associated with higher contralateral involvement. table 1.3 reports the prevalence of contralateral lymph node involvement depending on three factors: T-category, midsagittal extension, and whether ipsilateral level III was involved. For all 197 patients without midline extension, contralateral involvement in levels II, III, IV, V was 10%, 3%, 2%, 1% compared to 36%, 18%, 4%, 0% with midline extension (90 patients). In addition, out of 38 patients with late T-category, midsagittal extension, and positive ipsilateral level III, 21 & (55%) showed contralateral involvement in level II and 12 & (32%) in level III. Out of 39 late T-category patients with midsagittal extension but negative ipsilateral level III, contralateral involvement was lower (24% in level II, 7% in level III). In table 1.3, we consider ipsilateral level III rather than II, because level II is involved in 81% of all patients. However, when ipsilateral level II is not involved, contralateral involvement is unlikely (1 out of 55 patients showed metastases in contralateral level II). We further note that the three factors considered in table 1.3 are correlated. Out of 150 early T-category patients, 11 & (7%) showed midline extension, whereas 79 & (58%) out of 137 late T-category patients showed midline extension. As expected, contralateral involvement depended on primary tumor subsite. When considering only primary tumors strictly restricted to the tonsils (118 patients), contralateral involvement in levels II and III was 8% and 3%, respectively.

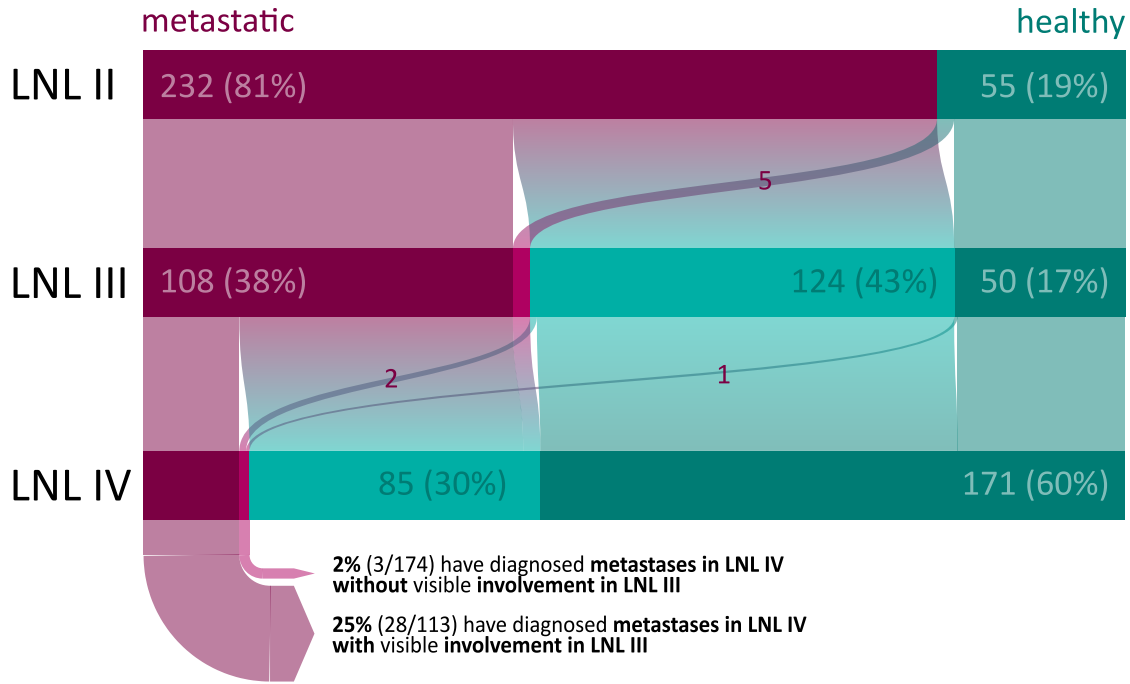


Figure 1.3: Ipsilateral involvement in levels III and IV depending on the involvement of upstream levels as flow plot.

Dependence on HPV status

The HPV status was positive for 181 & (63%), negative for 96 & (33%), and unknown for 10 patients & (4%). When considering LNL involvement for all T-categories and primary tumor subsites combined, the dataset provides no strong indication that LNL involvement is different for HPV positive versus HPV negative patients, neither regarding the patterns of spread to the LNLs nor in terms of prevalence of involvement. However, the data suggests that the association of higher lymph node involvement with more advanced T-category is more pronounced for HPV negative tumors than for HPV positive tumors, i.e. HPV-tumors tend to disseminate earlier to regional nodes (table 1.1). For example, for HPV

Table 1.3: Risk factors for contralateral involvement. Columns 1-3 define the 8 possible combinations of positive/negative mid-sagittal plan extension, late/early T-category, and positive/negative involvement of ipsilateral level III. Subsequent columns report the number of patients and percentages with involvement in the respective level for each combination of risk factors.

T-stage	midline	ipsi III	n	I	II	III	IV	V	VII
early	no	—	94	0 0%	4 4%	2 2%	1 1%	0 0%	1 1%
early	no	+	45	0 0%	8 18%	1 2%	1 2%	1 2%	0 0%
early	yes	—	8	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%
early	yes	+	3	0 0%	1 33%	1 33%	0 0%	0 0%	0 0%
late	no	—	31	0 0%	3 10%	1 3%	1 3%	1 3%	2 6%
late	no	+	27	1 4%	4 15%	1 4%	0 0%	0 0%	0 0%
late	yes	—	41	0 0%	10 24%	3 7%	1 2%	0 0%	1 2%
late	yes	+	38	2 5%	21 55%	12 32%	3 8%	0 0%	2 5%
			287	3	51	21	7	2	6

positive tumors, involvement of ipsilateral level III increased from 35% for early T-category to 47% for late T-category. Instead, for HPV negative tumors, involvement was 26% versus 49%.

Involvement of levels I, V and VII

Prevalence of ipsilateral involvement was 10% (30 out of 287 patients) in level I and 6% (16 out of 287 patients) in level V. No patient had metastases in ipsilateral level I or V without involvement of ipsilateral level II. Prevalence of ipsilateral level VII involvement was 10% (30 out of 287 patients) and was more frequent for tumors of the tonsil (14%, 17 out of 118 patients) than for tumors of the base-of-tongue (6%, 5 out of 83 patients). 4 patients had metastases in ipsilateral level VII without involvement of level II. Features of more advanced disease was associated with higher involvement. For example, involvement in levels I, V, VII was 6%, 1%, 6% in early T-category patient without metastases in ipsilateral level III (102 patients) and 22%, 14%, 17% for late T-category patient with metastases in ipsilateral level III (68 patients).

1.5 Discussion

The purpose of this study was to provide detailed per-level-quantifications of cervical lymph node involvement for oropharyngeal carcinoma on a patient-individual basis, depending on T-category, involvement of other nodal levels, and various clinicopathological factors such as smoking and HPV status. To the best of our knowledge, this is the only study providing such detailed quantitative information considering multimodal diagnostic modalities, which distinguishes this study from previous publications on the overall prevalence of LNL involvement for oropharyngeal cancer. Furthermore, an elaborate user-friendly GUI is provided to visualize and explore the dataset and study the dependence of LNL involvement as a function of the above parameters. While only selected information can be presented here in the form of tables and figures, the GUI can be used to access the full information contained in the dataset and study the influence of other factors such as primary tumor subsite or smoking status.

The main limitation of this dataset is that pathological involvement for the surgically treated patients was not available because neck dissection was performed en bloc. In addition, as a single-institution dataset, the number of patients is limited. However, dataset and GUI are made publicly available. The dataset can in the future be pooled with other datasets without loss of information, and the software platform and GUI developed may serve as the basis for collecting large multi-institutional datasets.

1.5.1 Comparison to previous works

Prevalence of LNL involvement

Overall patterns and prevalence of LNL involvement in our study (table 1.1) is consistent with previous studies [4, 9, 14]. Our study contained a relatively low number of N0 patients (16%) compared to previous reports, which may be explained by

differences in patient selection and diagnostic modalities used. Our study includes all patients treated at our institution between 2013 and 2019 irrespective of primary treatment. Hence, our patient cohort may be considered relatively unbiased. Studies reporting on surgically treated patients may introduce bias towards lower prevalence of LNL involvement when patients with advanced disease are referred to definitive chemoradiotherapy.

Dependence on upstream levels

The question on how the probability of metastases in a LNL depends on the involvement of upstream levels is poorly reported in the literature. To our knowledge, Sanguinetti et al [24] is the only publication reporting on this question for early T-category surgically treated OPSCC. For example, out of 42 patients with ipsilateral level III involved, 12 patients (29%) had also level IV involved ¹, which is similar to our findings (28 out of 113 patients, 25%, for all patients combined). In agreement with previous studies, our dataset confirms that skip metastases in levels III and IV occur but are rare (table 1.2). Furthermore, we observed no case of level I or V involvement without metastases in level II, which is also confirming previous publications. Further data collection and analysis in that direction could potentially lead to treatment-de-escalation strategies by not irradiating down-stream LNLs in the absence of metastases in up-stream levels, e.g. by identifying patients in whom level IV may be excluded from the CTV-N.

Contralateral spread

A prominent example of treatment de-escalation is the sparing of the elective contralateral irradiation or the pN0, negative, neck: Chronowski et al. [5] provided data of 102 patients with tonsillar carcinoma treated with unilateral radiotherapy, of which only 2% experienced contralateral recurrence. Very similar data, with contralateral recurrence rates of only 2-3.5% were reported from the Princess Margaret Hospital [13, 20]. Moreover, similarly to the results of the large meta-analysis of Al-Mamgani [19], we could demonstrate that the probability of contralateral involvement also strongly depends on T-category and midline extension. Concerning omission of radiotherapy to the pN0 neck, a recent prospective trial, with most of the patients included suffering from oropharyngeal cancer, could demonstrate excellent tumor control rates of 97% in the unirradiated neck [6]. These results show that CTV-N reduction is possible for selected patients. However, according to current guidelines, unilateral radiotherapy is recommended in specific circumstances. In our study, incidence of contralateral involvement was 20% and the data suggests that for many patients with favorable characteristics (early T-category, no midline extension, limited ipsilateral involvement), the risk of contralateral metastases is low (??). If supported by further multi-institutional data, this could identify additional patients in whom the contralateral neck may be safely excluded from the CTV-N, either completely or in part. E.g., radiotherapy could be limited to level II in some patients when level II still bears a relevant risk of occult metastases but the risk in levels III and IV is minimal.

¹These numbers are reconstructed from the data reported but are not directly contained in the publication.

HPV-status

Consistent with the findings of Bauwens et al [1] and the general clinical observation that HPV-positive tumors seem to metastasize early despite small primaries, our data suggests that the dependence of lymph node involvement on T-category is less pronounced in HPV-positive tumors (table 1.2). Beyond that, our data does not provide evidence that lymphatic progression patterns differ between HPV-positive versus negative tumors (consistent with Bauwens et al [1]).

1.5.2 From macroscopic progression patterns to microscopic involvement

In this work, we consider lymphatic progression patterns assessed through imaging. Hence, the distribution of macroscopic lymph node metastases is studied. For defining the elective clinical target volume (CTV) or the extent of surgical resection, we are instead interested in the conditional probability of microscopic involvement in a LNL given that no macroscopic metastases are seen in that level. This risk depends on two aspects: (1) The sensitivity and specificity of diagnostic imaging, i.e. the probability of not diagnosing lymph node metastases that are present; and (2) The probability that tumor cells have spread to a LNL, given the observed state of tumor progression for that patient. The latter can be obtained from datasets of metastatic progression patterns in a cohort of patients as presented in this paper. Statistical methods that combine both aspects to calculate risk of microscopic involvement have been developed for ipsilateral levels I-IV [22, 17]. Future work will extend these statistical models to contralateral spread and levels V and VII, informed by the data presented here. However, this is not part of the current paper.

1.5.3 Summary and prospect

Detailed datasets of lymphatic progression patterns, meaning reporting the combination of simultaneously involved LNLs together with tumor characteristics on a patient-individual basis, allows for better quantification of LNL involvement. This may in turn allow for further personalization of elective CTV-N definition based on the individual patient's state of tumor progression. In this paper we publish such a dataset together with a graphical user interface to explore the dataset. The software tools are made publicly available for others to study our dataset in detail and to contribute data for building larger multi-institutional datasets. Large datasets, possibly containing thousands of patients, together with the statistical methods for analysis, may eventually inform future clinical trials and guidelines on nodal CTV definition in head & neck cancer. Potential applications are to omit irradiation of ipsilateral level IV in selected patients, or to identify additional patients in whom contralateral neck irradiation can be omitted or limited to level II.