Appendix A.

A.1. Data collection

A.1.1. Patient data

This study is a retrospective analysis of 132 HNSCC patients (mean age at the time of diagnosis: 60.3±8.3y) treated with definitive radiotherapy or radiochemotherapy at our institution between 2002 and 2013. Pre-therapeutic standard diagnostic procedures like FDG-PET/CT, panendoscopy, fine needle aspiration (FNA) were already standard in 2002 in our academic cancer center. Planning CT-protocols and radiology reports did also not undergo crucial changes over this time-period. For the purpose of characterizing common HNSCC lymphatic progression patterns, we included only patients with newly diagnosed primary tumors located in the oral cavity, oropharynx, hypopharynx and larynx. Patients with a past medical history that may affect lymphatic progression patterns, such as prior neck surgery or cancer in the head & neck, were excluded. Data collection was based on radiology, pathology, and radiation oncology reports stored in the electronic medical record system. In case of ambiguous wording or imprecise specification of the location of lymph node metastases, the original medical images were reviewed by a senior radiation oncologist. Overall, 132 planning CT, 122 FDG-PET examinations, 57 diagnostic CTs, 27 MRI and 66 pathology reports were used to describe the lymphatic progression of 132 patients.

A.1.2. Primary tumor characteristic

To characterize the PT, its location and TNM stage (including the edition used) were extracted from the radiation oncology report. The PT location was assigned to one of the 4 categories: oral cavity, oropharynx, hypopharynx, and larynx. For tumors overlapping multiple regions of the pharynx, we used the GTV-T contours on the planning CT to determine which PT category concentrated most of the tumor mass. The position was considered lateralized if the major part of the tumor mass was located on one side (left or right). The extension over the mid-sagittal plane was also reported. The tumor was considered central when it visually extended equally to both sides. Additional information such as the HPV status, history of nicotine/alcohol abuse and whether the patient received induction chemotherapy before definitive (chemo)radiotherapy were recorded if available.

A.1.3. Multimodality LNL involvement assessment

LNL involvement was recorded in a binary fashion as positive or negative (or unknown if a diagnostic modality was not available). This reflects the clinical practice of radiotherapy where a LN is either part of the GTV-N or not. However, such binary labels face the problems that inherently arise when reducing continuous features such as size or FDG uptake to a binary label. For the different modalities, the following criteria were used:

- Radiotherapy planning CT: Whether a LNL was considered positive on the radiotherapy planning CT, was decided based on what was delineated as GTV-N in the clinically delivered treatment. GTV-N contours sometimes overlaid multiple LNL. In such cases, those LNL that contained at least one separate malignant LN were reported as positive. If instead a single enlarged LN overlaid two neighboring LNL, only the LNL containing the main part of the LN was labeled

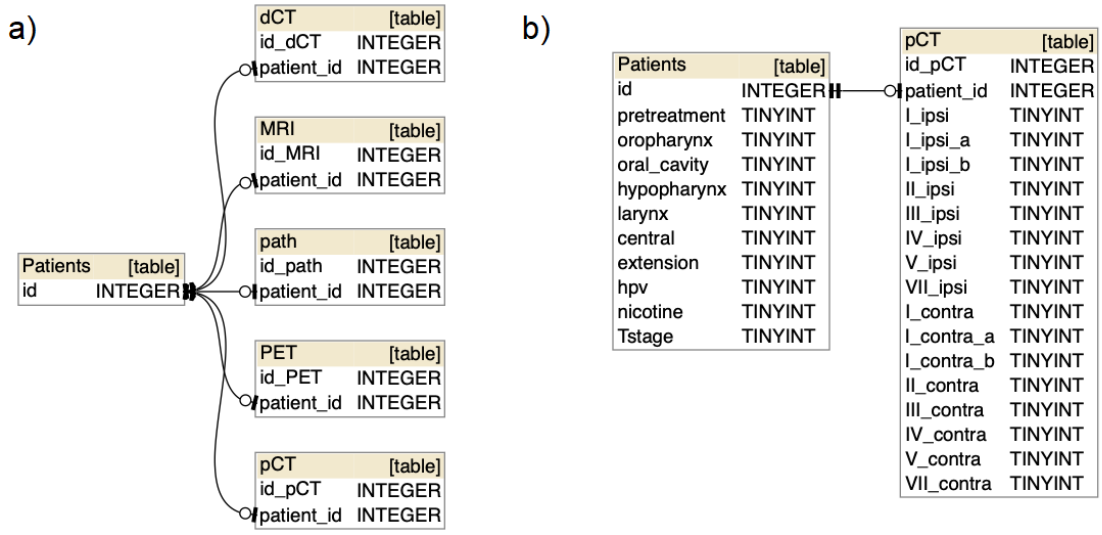

Figure A1. Entity-Relationship (ER) diagram representing our database. a) List of the tables contained in the database. The foreign key is represented via the arrow (a double-stop end represents an "one and only one" relationship, circle and bar represent an "one or none" relationship). b) List of variables stored in the respective tables.

positive.

- Diagnostic CT, MRI, and PET-CT: Most patients had PET-CT and an associate radiology report available for diagnosis. If a patient had an additional diagnostic CT or MRI image with an associated radiology report, LNL involvement was recorded separately for these modalities. Radiology reports are often characterized by the use of cautious formulations such as lymph nodes being described as "suspicious". If a LN was described as suspicious, the associated LNL was labeled as positive. The motivation is that, in order to consider a lymph node suspicious, it was abnormal in at least one criterion. Common criteria were SUV $\geq$ 3, diameter $\geq$ 1cm, round shape, inhomogenous contrast uptake, or central necrosis.
- USgFNA: Results of USgFNA were retrieved from pathology reports. LNL were classified as positive, negative, or unknown. If USgFNA did not show evidence of malignant tumor, the LNL was labeled negative. However, the interpretation of the 'negative' label must take into account that only selected LN are punctured and no information on other LN in the same level is provided. It was sometimes not possible to associate the punctured LN with a specific LNL unambiguously, in which case 'unknown' was recorded.

A.2. Database

The representation as a SQL database facilitates the selection and counting of patients presenting with a particular combination of LNL involvement and primary tumor characteristics. The Entity Relationship (E-R) diagram presented in figure A.1 summarizes the database structure.

Table A1. Number of patients with lateralized primary tumors presenting with a specific configuration of LNL involvement on PET examination. E.g. LI-II means involvement of levels I and II but not III and IV. Ipsi- and contralateral side are reported independently.
[c]counts the number of patients with the specific ipsilateral involvement configuration plus at least one contralateral LNL involved

PET					
	Oropharynx	Oral Cavity	Hypopharynx	Larynx	All
LNL involvement	Ipsi/Ipsi[c]/contra	Ipsi/Ipsi[c]/contra	Ipsi/Ipsi[c]/contra	Ipsi/Ipsi[c]/contra	Ipsi/Ipsi[c]/contra
None	10/1/48	2/0/7	6/0/20	5/0/7	23/1/82
L I	1/0/0	0/0/0	0/0/0	0/0/0	1/0/0
L II	35/14/17	1/0/2	5/1/3	2/1/1	43/16/23
L III	3/0/0	0/0/0	1/0/1	0/0/0	3/0/0
L IV	0/0/0	1/0/0	1/0/0	0/0/0	1/0/0
L I-II	2/1/1	4/1/0	0/0/0	0/0/0	6/2/1
L I-III	0/0/0	0/0/0	0/0/0	0/0/0	0/0/0
L I-IV	0/0/0	1/0/0	0/0/0	0/0/0	1/0/0
L II-III	17/6/4	1/1/0	9/3/1	1/0/0	27/9/5
L II-IV	0/0/0	1/1/0	0/0/0	0/0/0	1/1/0
L III-IV	0/0/0	0/0/0	1/0/0	0/0/0	1/0/0
L I-II-III	1/1/1	0/0/0	0/0/0	0/0/0	1/1/1
L I-II-IV	0/0/0	0/0/0	0/0/0	0/0/0	0/0/0
L I-III-IV	0/0/0	0/0/0	0/0/0	0/0/0	0/0/0
L II-III-IV	2/0/0	0/0/0	2/0/0	0/0/0	4/0/0
L I-II-III-IV	0/0/0	0/0/0	2/1/0	0/0/0	0/0/0
Total	71/23/71	9/2/9	24/4/24	8/1/8	112/30/112

A.3. Dataset description

We complete the results presented in table 1 with the corresponding PET findings (table A1). Involvement on the ipsilateral and contralateral side is reported separately. In addition, the second column labeled with 'Ipsi[c]' reports the number of patients with at least one contralateral LNL involved for the corresponding ipsilateral configuration.

Overall the most frequent ipsilateral combinations on planning CT are: level II only (46), levels II and III (25) and none (15). Approximately half of the patients (54 over 122) present with contralateral involvement, mostly in level II alone (32), levels II and III (7) and levels I and II (7). Similar trends are observed for PET results.

The remaining levels were involved in few cases. Level VI was involved in 6 patients on the planning CT (6 ipsilateral including 1 contralateral) and in one additional patient (ipsilateral) on PET. Level VIII was involved in 2 oropharyngeal patients (once contralateral and once on both sides). In contrast level V was involved twice on the ipsilateral side for hypopharyngeal cases. Level VII was positive 12 times on the ipsilateral side and twice on the contralateral side.

A.4. Dataset limitations

Our dataset reflects the patient population in our radiation oncology department, which may not be representative for the whole population of head & neck cancer patients due to patterns of care. The decision whether patients are treated with surgery versus radiotherapy may introduce a bias, which may in part explain difference compared to literature values. This can be seen in the distribution of T-stages in our data set, which contains a majority of T2 (29%) and T4 (41%) cases and relatively few T1 (7%) and T3 (23%) cases. In addition, our data contain a low number of N0 cases. Table A2 compares the rate of involvement of levels I to IV in our dataset to the values

Table A2. Distribution of clinical LNL involvement reported in [1] compared to our dataset.

	N+	I	II	III	IV
Oral cavity[1]	36	43[a]/3.5[b]	79/8	18/3	1/0
USZ	78	86/14	86/43	29/0	14/0
Oropharynx[1]	64	13/2	81/24	23/5	9/2.5
USZ	92	23/10	96/51	36/11	3/0
Hypopharynx[1]	70	2/0	80/13	51/4	20/3
USZ	93	23/23	96/38	62/19	19/0
Supraglottic[1]	80	2/0	71/21	48/10	22/15
USZ	44	0/0	100/50	50/0	0/0

[a]/[b] percentage of ipsilateral/contralateral involvement in lateralized N+ patients

reported in the review by Grégoire et al.[1].

LNL involvement based on PET diagnostics was similar, but slightly lower, compared to the involvement based on planning CT. This may be explained by the assumption that the radiation oncologist's decision takes into account all diagnostic modalities, and suspicious lymph nodes are included in the GTV-N as soon as a lymph node appears abnormal on one of the modalities. There is a larger discrepancy between planning CT and PET for level I, which may be explained by the difficulty to separate level Ib and II.

Appendix B.

B.1. Bayesian network model for microscopic involvement

In this section we detail the BN model outlined in section 2.3. To that end,

- each LNL i is associated with a latent binary variable x_i for its microscopic state, which indicates if the LNL truly harbors tumor including occult metastases. This state is unobserved for radiotherapy patients and the goal is to estimate the probability that $x_i = 1$.
- each LNL i is further associated with a binary random variable z_{ij} for each diagnostic method j (PET, CT, MRI, FNA), which represent the macroscopic state and indicate if a LNL i harbors suspected metastases based on diagnostic modality j.

A Bayesian network model specifies the probabilistic relationship between these random variables in terms of conditional probability distributions. Microscopic state x_i and macroscopic observation z_{ij} are related via sensitivity and specificity, i.e. the probability of a macroscopic observation z_{ij} given the ground truth x_i is given by the conditional probability distribution

$$\begin{pmatrix} P(z_{ij}=1|x_i=1) & P(z_{ij}=0|x_i=1) \\ P(z_{ij}=1|x_i=0) & P(z_{ij}=0|x_i=0) \end{pmatrix} = \begin{pmatrix} sens_j & (1-sens_j) \\ (1-spec_j) & spec_j \end{pmatrix} \tag{B1}$$

The probability distribution over the microscopic states is determined by the probability of the primary tumor to spread through the lymphatic network. Here, we illustrate this for the case where we only consider levels II and III and a given primary tumor location. A general and more detailed description of the methodology is pub-

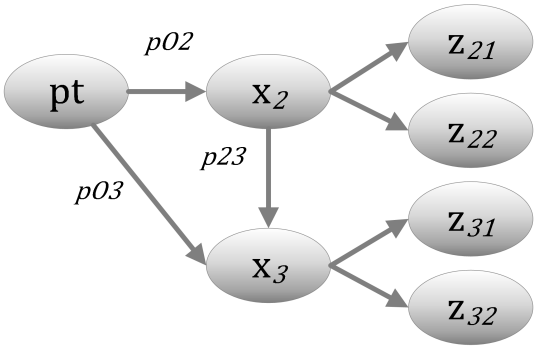

Figure B1. Bayesian network representation of the lymphatic spreads from the primary tumor to the ipsilateral levels II and III. Nodes represent random variables while the links represent conditional probabilities. In this example, each lymph node level is observed by two diagnostic modalities z_{i1} and z_{i2}. The tumor is represented as an additional node, which corresponds to a binary random variable that is always positive

lished elsewhere. Considering only levels II and III, the joint probability $P(x_2, x_3)$ can be parametrized via three independent parameters: p_{02} and p_{03}, describing the probability for the tumor to directly spread to levels II and III, respectively; and the probability p_{23} to spread from level II to level III. The latter is motivated by the fact that level III receives efferent lymphatics from level II. This yields

$$\begin{aligned} P(x_2, x_3) &= \begin{pmatrix} P(x_2 = 1, x_3 = 1) & P(x_2 = 0, x_3 = 1) \\ P(x_2 = 1, x_3 = 0) & P(x_2 = 0, x_3 = 0) \end{pmatrix} \\ &= \begin{pmatrix} p_{O2}(1 - (1 - p_{23})(1 - p_{O3})) & p_{O3}(1 - p_{O2}) \\ p_{O2}(1 - p_{23})(1 - p_{O3}) & (1 - p_{O2})(1 - p_{O3}) \end{pmatrix} \end{aligned} \tag{B2}$$

This formulation is also known as the independence of causal influence (ICI) model [2]. The graph of the corresponding Bayesian Network is shown in figure B.1 and the joint probability distribution associated with it is:

$$P(\{z_{3j}\}, x_3, \{z_{2j}\}, x_2) = P(x_2, x_3) \prod_j P(z_{3j}|x_3)(z_{2j}|x_2) \tag{B3}$$

We are now interested in the probability of microscopic involvement of level III, depending on whether or not level II harbors macroscopic metastases. For the sake of simplicity, we derive the estimation for only one diagnostic modality, such that the index j can be ignored. The probability of microscopic involvement given diagnostic findings z_2 and z_3 is given by the conditional probability $P(x_3 = 1|z_3, z_2)$, which is obtained by Bayesian inference:

$$P(x_3 = 1|z_3, z_2) = \frac{P(z_3, z_2, x_3 = 1, x_2 = 1) + P(z_3, z_2, x_3 = 1, x_2 = 0)}{P(z_3, z_2)} \tag{B4}$$

This formulation implicitly depends on the unknown model parameters (p_{O2}, p_{O3}, p_{23}). At this step, we include the knowledge gathered in our database $(\boldsymbol{D})$, which determines the posterior density function over the parameters $f(p_{O2}, p_{O3}, p_{23}|\boldsymbol{D})$. Equation (B4) becomes:

$$P(x_3 = 1|z_3, z_2) = \iiint_0^1 P(x_3 = 1|z_3, z_2; p_{O2}, p_{O3}, p_{23}) \dots \cdot f(p_{O2}, p_{O3}, p_{23}|\boldsymbol{D})\, dp_{O2}\, dp_{O3}\, dp_{23} \tag{B5}$$

Here, we assume uniform priors over the model parameters. Thus, the posterior density function $f(p_{O2}, p_{O3}, p_{23}|\boldsymbol{D})$ is directly proportional to the likelihood function L.

$$f(p_{O2}, p_{O3}, p_{23}|\boldsymbol{D}) = \frac{L(\boldsymbol{D}|p_{O2}, p_{O3}, p_{23})}{\iiint L(\boldsymbol{D}|p_{O2}, p_{O3}, p_{23})\, dp_{O2}\, dp_{O3}\, dp_{23}} \tag{B6}$$

Since our model contains only three parameters, the likelihood function can be calculated numerically by discretizing the parameters space in 0.01 steps. We performed the BN numerical analysis using the Bayes Net Toolbox[3] for Matlab (R2016b,The Mathworks, USA). Assuming a dataset $\boldsymbol{D}$ of M patients with the observations (z_2^m, z_3^m), the likelihood function is given by:

$$L(\boldsymbol{D}|p_{O2}, p_{O3}, p_{23}) = \prod_{m=1}^{M} \left[\sum_{x_2 \in \{0,1\}} \sum_{x_3 \in \{0,1\}} P(x_2, x_3, z_2^m, z_3^m|p_{O2}, p_{O3}, p_{23}) \right] \tag{B7}$$

The computation proposed in (eq. B5) can therefore be carried out numerically to yield probabilities of microscopic involvement. The methodology of Bayesian networks is illustrated here only for the simplified case considering oropharyngeal tumors and levels II and III. The methodology can be extended to include all levels, however, a detailed presentation thereof will be presented elsewhere.

B.2. Microscopic risk estimation

The risk of microscopic involvement in level III as stated in section 3.3 is obtained from (eq. B5) by numerical integration of the likelihood function. We obtain, for the case of metastases in level II:

$$P(x_3 = 1|z_3 = 0, z_2 = 1) = 0.12 \tag{B8}$$

and for the case without metastases in level II:

$$P(x_3 = 1|z_3 = 0, z_2 = 0) = 0.09 \tag{B9}$$

For comparison, Sanguineti et al. [4] reports 16% risk of occult metastasis in level III for the case of positive level II assuming a sensitivity and specificity of 0.88 and 0.39, respectively. However, recalculating these values for their dataset but with sensitivity and specificity values of 0.79 and 0.86 yields 13.1% risk.

Most papers report the percentage of patients in which a particular LNL harbors metastases which is different from the probabilities stated in (eq. B8) and in (eq. B9).

In our dataset, ipsilateral level III is macroscopically involved in 23 out of 71 lateralized oropharynx patients, corresponding to 33%. In the statistical model described in section A.1, this corresponds to the joint probability distribution $P(z_3, x_3, z_2, x_2)$ specified via equations (eq. B1-B3) marginalized over all variables other than z_3, i.e.

$$P(z_3 = 1) = \iiint_0^1 \sum_{z_2} \sum_{x_3} \sum_{x_2} P(z_3 = 1|x_3) P(z_2|x_2) P(x_2, x_3) \ldots \cdot f(p_{O2}, p_{O3}, p_{23}|\boldsymbol{D})\, dp_{O2}\, dp_{O3}\, dp_{23} \tag{B10}$$

which evaluates to 0.36 (meaning that the statistical model adequately describes the dataset with regard to the cumulative incidence of level III involvement).

References

[1] Grégoire V, Coche E, Cosnard G, et al. Selection and delineation of lymph node target volumes in head and neck conformal radiotherapy. proposal for standardizing terminology and procedure based on the surgical experience. Radiotherapy and Oncology. 2000; 56(2):135–150.

[2] Dıez FJ, Druzdzel MJ. Canonical probabilistic models for knowledge engineering. Technical Report CISIAD-06-01, UNED, Madrid, Spain; 2006. Report.

[3] Murphy K. The bayes net toolbox for matlab. Computing science and statistics. 2001; 33(2):1024–1034.

[4] Sanguineti G, Califano J, Stafford E, et al. Defining the risk of involvement for each neck nodal level in patients with early t-stage node-positive oropharyngeal carcinoma. Int J Radiat Oncol Biol Phys. 2009;74(5):1356–64. Available from: https://www.ncbi.nlm.nih.gov/pubmed/19131181.