Probabilistic Model of Bilateral Lymphatic Spread in Head and Neck Cancer

Roman Ludwig[1,2*], Yoel Perez Haas[1,2], Sergi Benavente[3], Panagiotis Balermpas[2], Jan Unkelbach[1,2]

[1]Physics, University of Zurich.

[2]Radiation Oncology, University Hospital Zurich.

[3]Radiation Oncology, University Hospital Vall d'Hebron.

*Corresponding author(s). E-mail(s): roman.ludwig@usz.ch;

Contributing authors: yoel.perezhaas@usz.ch; sergi.benavente@vallhebron.cat; panagiotis.balermpas@usz.ch; jan.unkelbach@usz.ch;

Abstract

Purpose: Current guidelines for elective nodal irradiation in oropharyngeal squamous cell carcinoma (OPSCC) recommend including large portions of the contralateral lymphatic system in the clinical target volume (CTV-N), even for lateralized tumors with no clinical lymph node involvement in the contralateral neck. This study introduces a probabilistic model of bilateral lymphatic tumor progression in OPSCC to estimate personalized risks of occult disease in specific lymph node levels (LNLs) based on clinical lymph node involvement, T-stage, and tumor lateralization.

Methods: Building on a previously developed hidden Markov model for ipsilateral lymphatic spread, we extend the approach to contralateral neck involvement. The model represents LNLs I, II, III, IV, V, and VII on both sides of the neck as binary hidden variables (healthy or involved), connected via arcs representing spread probabilities. These probabilities are learned using Markov chain Monte Carlo (MCMC) sampling from a dataset of 833 OPSCC patients, enabling the model to reflect the underlying lymphatic progression dynamics.

Results: The model accurately and precisely describes observed patterns of lymph node involvement with a compact set of interpretable parameters. Midline extension of the primary tumor is identified as the primary risk factor for contralateral involvement, with advanced T-stage and extensive ipsilateral involvement further increasing risk. The model highlights key hierarchical dependencies: occult disease in contralateral LNL III is highly unlikely if upstream LNL II is clinically negative, and in contralateral LNL IV, occult disease is exceedingly rare without LNL III involvement.

Conclusions: This model offers an interpretable, probabilistic framework to inform personalized elective CTV-N volume reduction. For lateralized tumors that do not cross the midline, it suggests the contralateral neck may safely be excluded from elective irradiation. For tumors extending across the midline but with a clinically negative contralateral neck, elective irradiation could be limited to LNL II, reducing unnecessary exposure while maintaining oncological safety.

Introduction

In treating head and neck squamous cell carcinomas (HNSCC) with radiotherapy or surgery, both the primary tumor and clinically detected lymph node metastases are targeted. Current guidelines also include large portions of the neck in the elective clinical target volume (CTV-N) [1–8] to mitigate the risk of regional recurrences from untreated microscopic disease undetectable by in-vivo imaging modalities such as computed tomography (CT), magnetic resonance imaging (MRI), or positron emission tomography (PET). However, this approach must balance minimizing the risk of occult disease in the lymphatic drainage region against the toxicity of unnecessarily irradiating healthy tissue.

These CTV-N guidelines rely on anatomically defined lymph node levels (LNLs) [2] and the overall prevalence of lymph node metastases within these levels. They often recommend extensive irradiation of both sides of the neck. However, the general prevalence of metastasis in a given LNL does not correspond to an individual patient's risk of occult disease in that region, which depends on their specific state of tumor progression. For example, a patient with no clinically detectable nodal disease (cN0) who has a small, clearly lateralized T1 tumor would receive the same contralateral CTV-N as a patient with significant ipsilateral nodal involvement and an advanced tumor crossing the mid-sagittal plane: Both patients receive elective irradiation of the contralateral LNLs II, III, and IVa [5].

To better quantify individualized risk of occult disease, we previously developed an intuitive probabilistic hidden Markov model (HMM) [9, 10], originally based on a conceptually similar a Bayesian network model [11]. However, these models have been

limited to predicting ipsilateral nodal involvement. This work extends the model to include contralateral risk predictions, enabling more personalized radiation volume recommendations. By identifying patients with low contralateral risk, the model could guide reductions in contralateral CTV-N, thereby decreasing radiation-induced toxicity and improving quality of life.

The main contributions of this paper are as follows:

1. Section 3 presents a multi-centric dataset on lymph node involvement in 833 OPSCC patients, identifying key risk factors for contralateral lymph node involvement and outlining requirements for a bilateral model extension (section 3).
2. Section 3 introduces a bilateral HMM that incorporates primary tumor lateralization, T-category, and clinical involvement as risk factors for contralateral involvement. Model training and computational experiments are described in section 3.
3. Section 3 demonstrates the model's ability to replicate observed contralateral lymph node involvement patterns and estimates occult disease risk for typical patients. Implications for volume-deescalated radiotherapy are discussed in section 3.

Data on Lymphatic Progression Patterns

To develop models for lymphatic tumor progression for all relevant LNLs, including contralateral regions, we compiled a detailed dataset of 833 patients with newly diagnosed oropharyngeal squamous cell carcinomas (OPSCC) [12, 13]. The dataset includes lymph node involvement per LNL for each patient in tabular form, along with primary tumor and patient characteristics such as T-category, subsite, primary tumor lateralization, and HPV p16 status. Patient records were collected from four institutions, and an overview of patient characteristics is provided in table 1.

Data from Inselspital Bern (ISB) and Centre Léon Bérard (CLB) consist exclusively of patients who underwent neck dissections. In contrast, the majority of patients from the University Hospital Zürich (USZ) and the Hospital Vall d'Hebron (HVH) were treated with definitive radiotherapy. Since surgical treatment is more common for early T-category patients, ISB and CLB datasets include a higher proportion of these cases compared to USZ and HVH. For 83 patients in the CLB dataset, the primary tumor's lateralization was not reported.

Source: Article Notebook

Table 1: Overview over the five datasets from four different institutions used to train and evaluate our model. Here, we briefly characterize the total number of OPSCC patients from the respective institution, their median age, what proportion received some form of neck dissection, the N0 portion of patients, what percentage presented with early T-category, and the prevalence of primary tumor midline extension. For a much more detailed look at the data, visit lyprox.org.

Table 1

Institution	Total	Age (median)	Neck Dissection	N0	Early T-Cat.	Mid.
Centre Léon Bérard	325	60	100%	19%	69%	18%
Inselspital Bern	74	61	100%	18%	66%	14%
University Hospital Zurich	287	66	26%	18%	52%	31%
Vall d'Hebron Barcelona Hospital	147	58	5%	21%	34%	34%

Source: Article Notebook

Consensus on Involvement Status

Pathological involvement is available only for surgically treated patients and for the levels that were dissected. For non-surgical patients, involvement status is determined clinically, i.e. using imaging. To ensure uniformity in analysis, diagnostic information was synthesized into a consensus decision for each patient and LNL. This consensus reflects the most likely state of involvement and accounts for the sensitivity and specificity of various diagnostic modalities, as reported in the literature [14, 15].

The consensus process is detailed in section 3. Briefly, pathological findings from neck dissections are treated as the gold standard, overriding any conflicting clinical diagnoses. For levels not dissected, PET-CT is typically the primary source for determining the most likely state of involvement.

Data Availability

The complete dataset, including additional patients with tumors in primary sites other than the oropharynx, is publicly accessible. It can be downloaded from LyProX, where it is also available for interactive exploration, from GitHub, or from Zenodo. Additional details about the datasets and data format are provided in designated Data-in-Brief publications [12, 13]. These publications do not include the latest data from HVH, which will be described in a future publication.

Patterns of Contralateral Involvement

The datasets enable analysis of correlations between LNL involvement and key risk factors. In figure 1, we illustrate the prevalence of contralateral LNL involvement, stratified by T-category, the number of ipsilaterally involved LNLs, and whether the tumor extends across the mid-sagittal plane.
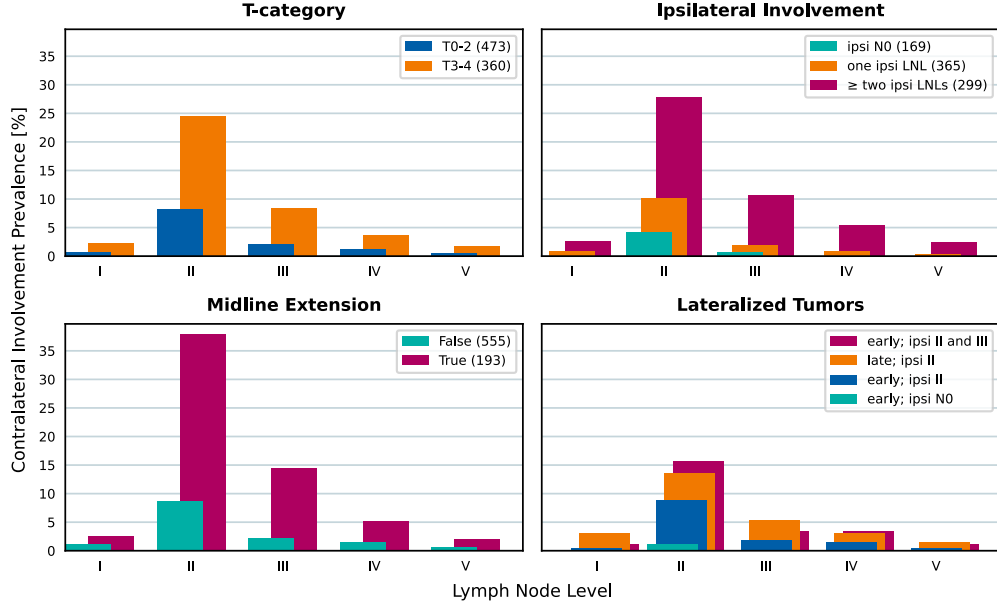
Figure 1: Contralateral involvement stratified by T-category (top left panel), the number of metastatic LNLs ipsilaterally (top right panel), and whether the primary tumor extended over the mid-sagittal line or was clearly lateralized (bottom left panel). Additionally, in the bottom right, we show that even for lateralized tumors, T-category and ipsilateral involvement are correlated with contralateral metastasis by comparing the contralateral involvement prevalence for selected scenarios that vary in their T-category and ipsilateral involvement extent.

Midline Extension

The bottom left panel of figure 1 shows that tumors crossing the mid-sagittal plane have a substantially higher prevalence of contralateral involvement compared to clearly lateralized tumors. This aligns with the anatomy of the head and neck lymphatic system, which is symmetric, with no major lymph vessels crossing the midline. Interstitial fluids from the primary tumor, presumed to carry malignant cells, can merely diffuse to contralateral lymphatic vessels over short distances. Thus, contralateral spread is more likely when the tumor approaches or crosses the midline.

T-Category

The top left panel shows a correlation between T-category and contralateral involvement, reflecting T-category's role as a surrogate for the time elapsed between disease onset and diagnosis. Advanced T-category tumors (e.g., T4) generally represent a longer disease progression timeline, providing more opportunity for metastatic spread compared to smaller tumors (e.g., T1).

Ipsilateral Involvement

The top right panel reveals a positive correlation between ipsilateral and contralateral metastases. Extensive ipsilateral involvement likely indicates a longer or faster disease progression. Additionally, it has been hypothesized that bulky ipsilateral nodal disease may reroute lymphatic drainage toward the contralateral side, potentially facilitating contralateral metastasis.

Interplay of Risk Factors

Midline extension, T-category, and ipsilateral involvement are interrelated risk factors for contralateral metastasis. For instance, 45.56% of advanced T-category tumors exhibit midline extension compared to 6.13% of early T-category tumors. While the higher fraction of midline extensions in advanced T-category patients partially accounts for the higher contralateral metastasis rates, T-category itself and ipsilateral involvement also play an additional role.

The bottom right panel of figure 1 isolates patients with lateralized tumors that do not cross the midline, further demonstrating these relationships. Among early T-category patients with no ipsilateral nodal involvement (levels I-V), only 1.85% (2 of 108 patients) show involvement in contralateral level II. This proportion increases to 6.73% (14 of 208) if ipsilateral level II is involved, to 14.29% (10 of 70) if ipsilateral levels II and III are involved, and further to 39.68% (25 of 63) for advanced T-category tumors with ipsilateral levels II and III involved.

Requirements for a Bilateral Model

Based on the observations in section 3 above, any model predicting the risk of contralateral nodal involvement should account for the following:

1. **Midline Extension:** Tumors extending across the mid-sagittal plane should result in a significantly higher probability of contralateral metastases.
2. **T-Category and Disease Progression:** Advanced T-category should correspond to an increased risk of nodal disease. This can be modeled using the expected time of diagnosis, as demonstrated in a prior hidden Markov model [9].
3. **Ipsilateral Involvement as a Progression Marker:** The extent of ipsilateral involvement should provide additional information on the time elapsed between disease onset and diagnosis, supplementing the insights gained from T-category alone.

Unilateral Model for Lymphatic Progression

This paper builds on the previously developed unilateral model for ipsilateral lymph node involvement presented in [16]. Below, we briefly recap the unilateral model to introduce the notation required for extending the framework to a bilateral model, described in section 3. For further details on the ipsilateral model, refer to earlier publications [9, 16].

We represent a patient's state of involvement at an abstract time-step t as a vector of hidden binary random variables, where each component corresponds to a lymph node level (LNL):

$$\mathbf{X}[t] = (X_v[t]) \qquad v \in \{1, 2, ..., V\} \tag{1}$$

Here, V is the number of LNLs the model considers. The values a LNL's may take on are $X_v[t] = 0$ (`False`), meaning the LNL v is healthy or free of metastatic disease, or $X_v[t] = 1$ (`True`), corresponding to some form of tumor presence (i.e., occult or macroscopic disease). In total, there are 2^V distinct possible lymphatic involvement patterns, which we enumerate from $\xi_0 = (0\ 0\ \cdots\ 0)$ to $\xi_{2^V} = (1\ 1\ \cdots\ 1)$. Each LNL's state is observed via another binary random variable Z_v that describes the clinical involvement of a LNL based on imaging: $Z_v = 0$ (`False`) indicates that the LNL v is healthy based on clinical diagnosis, and $Z_v = 1$ (`True`) indicates that LNL v was classified as involved. X_v and Z_v are connected through the sensitivity and specificity of the diagnositc modality.

Based on this, our HMM is fully described by defining the following three quantities:

1. A starting state $\mathbf{X}[t=0]$ at time $t=0$ just before the patient's tumor formed. In our case, this is always the state ξ_0 where all LNLs are still healthy.
2. The *transition matrix*

$$\mathbf{A} = (A_{ij}) = (P\left(\mathbf{X}[t+1] = \xi_j \mid \mathbf{X}[t] = \xi_i\right)) \tag{2}$$

   where the value at row i and column j represents the probability to transition from state ξ_i to ξ_j during the time-step from t to $t+1$. Note that we prohibit self-healing, meaning that during a transition, no LNL may change their state from $X_v[t] = 1$ to $X_v[t+1] = 0$. Consequently, many elements of the transition matrix are zero.
3. Lastly, the *observation matrix*

$$\mathbf{B} = (B_{ij}) = (P\left(\mathbf{Z} = \zeta_j \mid \mathbf{X}[t_D] = \xi_i\right)) \tag{3}$$

   where in row i and at column j we find the probability to *observe* a lymphatic involvement pattern $\mathbf{Z} = \zeta_j$, given that the true (but hidden) state of involvement at the time of diagnosis t_D is $\mathbf{X}[t_D] = \xi_i$.

The transition matrix $\mathbf{A}$ is parameterized using a directed acyclic graph (DAG) that represents the underlying lymphatic network. Edges from the primary tumor to an LNL are associated with a probability b_v for direct spread to LNL v during one time step. Arcs from a LNL v to a LNL r are parameterized with the probability rate t_{vr} for inter-nodal progression. In this paper, we build on the DAG shown in figure 2 which was obtained by maximizing the model evidence as described in [16].

We can evolve the distribution of all possible hidden states from $\mathbf{X}[t=0] = \xi_0$ step by step, by successively multiplying this vector with the transition matrix $\mathbf{A}$:
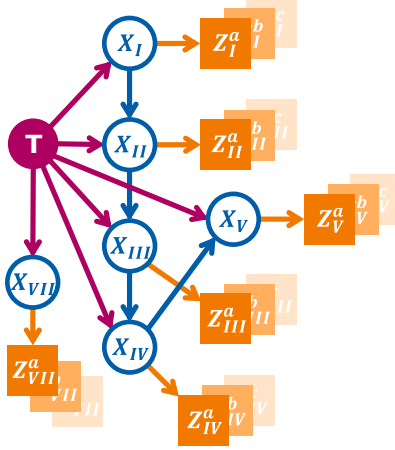
Figure 2: Directed acyclic graph (DAG) representing the abstract lymphatic network in the head and neck region. Blue nodes are the LNLs' hidden random variables, the red node represents the tumor, and the orange square nodes depict the binary observed variables. Red and blue arcs symbolize the probability of lymphatic spread along that edge during one time-step. The orange arcs represent the sensitivity and specificity of the observational modality (e.g. CT, MRI, pathology, …).

$\mathbf{X}[t+1] = \mathbf{X}[t] \cdot \mathbf{A}$. For later use, we define at this point a matrix Λ that collects these distributions for all considered time-steps:

$$\Lambda = P\left(\mathbf{X} \mid \mathbf{t}\right) = \begin{pmatrix} \pi^\top \cdot \mathbf{A}^0 \\ \pi^\top \cdot \mathbf{A}^1 \\ \vdots \\ \pi^\top \cdot \mathbf{A}^{t_{\max}} \end{pmatrix} \tag{4}$$

Where the k-th row in this matrix corresponds to the probability distribution over hidden states after $t = k - 1$ time-steps.

At the time of diagnosis, $0 \leq t_D \leq t_{\max}$, we multiply the evolved state distribution with the observation matrix $\mathbf{B}$ to obtain the distribution over all possible diagnoses. However, the exact time of diagnosis, t_D, is unknown; that is, we do not know the number of time-steps over which the HMM should be evolved. To address this, we marginalize over all possible diagnosis times, allowing the diagnosis to occur at any time-step, albeit with different weights. These weights are defined by a prior distribution over t_D, which can vary depending on the patient's T-category. For example, the time-prior for early T-category patients, $P(t_D \mid \text{early})$, may put more weight on earlier time-steps, reflecting – on average – earlier detection, compared to the prior for advanced T-category patients, $P(t_D \mid \text{advanced})$.

$$P\left(\mathbf{X} \mid \mathrm{T}x = \text{early}\right) = \sum_{t=0}^{t_{\max}} P\left(\mathbf{X} \mid t\right) \cdot P(t \mid \text{early})$$

In this work, we use binomial distributions $\mathfrak{B}\left(t_D, p_{\mathrm{T}x}\right)$ as time-priors which have one free parameter $p_{\mathrm{T}x}$ for each group of patients we differentiate based on T-category. Also, we fix $t_{\max} = 10$, which means that the expected number of time-steps from the onset of a patient's disease to their diagnosis is $\mathbb{E}\left[t_D\right] = 10 \cdot p_{\mathrm{T}x}$.

Likelihood Function of the Unilateral Model

The likelihood function for a patient to present with a diagnosis ζ_i a T-category $\mathrm{T}x$ tumor can now be written as:

$$\ell = P\left(\mathbf{Z} = \zeta_i \mid \mathrm{T}x\right) = \sum_{t=0}^{t_{\max}} \left[\xi_0 \cdot \mathbf{A}^t \cdot \mathbf{B}\right]_i \cdot P\left(t \mid \mathrm{T}x\right) \tag{5}$$

With $[...]_i$ we denote the i-th component of the vector in the square brackets. Note that it is also possible to account for missing involvement information: If a diagnosis (like fine needle aspiration (FNA)) is only available for a subset of all LNLs, we can sum over all those possible complete observed states ζ_j that match the provided diagnosis.

The single-patient likelihood ℓ in equation 5 depends on the spread parameters shown in figure 2 via the transition matrix $\mathbf{A}$ and on the binomial parameters $p_{\mathrm{T}x}$ via time-priors. In this work, we will only differentiate between "early" (T1 & T2) and "advanced" (T3 & T4) T-categories. Therefore, the parameter space of the unilateral model is:

$$\theta = \left(\{b_v\}, \{t_{vr}\}, p_{\text{early}}, p_{\text{adv.}}\right) \quad \text{with} \quad \substack{v \leq V \\ r \in \text{pa}(v)} \tag{6}$$

And it is our goal to infer optimal parameter values of from a given dataset $\mathcal{D}$ (consisting of diagnoses and T-categories) of OPSCC patients. The likelihood to observe this cohort of N patients, given a set of parameters θ is simply the product of their individual likelihoods as defined in equation 5. For numerical reasons, we typically compute the data likelihood in log space:

$$\log \mathcal{L}\left(\mathcal{D} \mid \theta\right) = \sum_{i=1}^{N} \log \ell_i \tag{7}$$

The methodology we use to infer the model's parameters is detailed in section 3.

Extension to a Bilateral Model

A straightforward approach to modeling contralateral lymphatic spread would be to use two independent unilateral models, as described in section 3, with shared parameters such as the distribution of diagnosis times or spread between LNLs (t_{vr}). However, this method would fail to capture the correlation between ipsilateral and contralateral involvement discussed in section 3, particularly the observed increase in contralateral involvement with greater severity of ipsilateral spread.

Thus, we extend the formalism in section 3 in such a way that the model's ipsi- and contralateral side evolve synchronously over time. To achieve that, we start by writing down the posterior distribution of involvement in analogy to equation 15, which is now a joint probability of an involvement $\mathbf{X}^i$ ipsilaterally *and* an involvement $\mathbf{X}^c$ contralaterally, given a diagnosis of the ipsilateral LNLs $\mathbf{Z}^i$ and of the contralateral ones $\mathbf{Z}^c$:

$$P\left(\mathbf{X}^i, \mathbf{X}^c \mid \mathbf{Z}^i, \mathbf{Z}^c\right) = \frac{P\left(\mathbf{Z}^i, \mathbf{Z}^c \mid \mathbf{X}^i, \mathbf{X}^c\right) P\left(\mathbf{X}^i, \mathbf{X}^c\right)}{P\left(\mathbf{Z}^i, \mathbf{Z}^c\right)} \tag{8}$$

For the sake of brevity, we omit the dependency on the parameters and the T-category here.

The probability of the diagnoses given a hidden state factorises: $P\left(\mathbf{Z}^i, \mathbf{Z}^c \mid \mathbf{X}^i, \mathbf{X}^c\right) = P\left(\mathbf{Z}^i \mid \mathbf{X}^i\right) \cdot P\left(\mathbf{Z}^c \mid \mathbf{X}^c\right)$, and the two factors are described through observation matrices $\mathbf{B}^i$ and $\mathbf{B}^c$.

The term representing the model's prior probability of hidden involvement does not factorize. However, we assume no direct lymphatic drainage from ipsilateral to contralateral LNLs, as major lymph vessels do not cross the mid-sagittal plane. In the graphical model, this translates to the absence of directed arcs between ipsilateral and contralateral LNLs, implying that contralateral tumor spread occurs solely via the primary tumor. We can thus write the joint probability $P\left(\mathbf{X}^i, \mathbf{X}^c\right)$ as a factorising sum:

$$
\begin{aligned}
P\left(\mathbf{X}^i, \mathbf{X}^c\right) &= \sum_{t=0}^{t_{\max}} P(t) \cdot P\left(\mathbf{X}^i, \mathbf{X}^c \mid t\right) \\
&= \sum_{t=0}^{t_{\max}} P(t) \cdot P\left(\mathbf{X}^i \mid t\right) \cdot P\left(\mathbf{X}^c \mid t\right)
\end{aligned}
\tag{9}
$$

This assumption is intuitive: since no major lymph vessels cross the midline, the ipsilateral and contralateral sides of the lymphatic network evolve independently over time. However, they are indirectly coupled through time. For example, a joint state with severe contralateral involvement and limited ipsilateral involvement is improbable: Severe contralateral involvement typically occurs at later time steps, when limited ipsilateral involvement is unlikely.

10

Using equation 9 along with equation 4, we can write the above distribution algebraically as a product:

$$P\left(\mathbf{X}^{\mathrm{i}} = \xi_n, \mathbf{X}^{\mathrm{c}} = \xi_m\right) = \left[\Lambda_{\mathrm{i}}^{\top} \cdot \operatorname{diag} P(\mathbf{t}) \cdot \Lambda_{\mathrm{c}}\right]_{n,m} \tag{10}$$

Parameter Symmetries

The matrices $\Lambda\mathrm{i}$ and $\Lambda\mathrm{c}$ could, in principle, be parameterized with entirely separate parameters, allowing ipsilateral and contralateral spread rates to differ completely. However, we simplify the parameter space by sharing parameters between the two sides, based on the following three assumptions:

1. **Shared Graph Structure**: Both ipsilateral and contralateral spread are described by the same graph shown in figure 2.
2. **Symmetric Spread Among LNLs**: The spread among LNLs is assumed to be the same on both sides, reflecting the symmetric structure of the lymphatic system. Consequently, the spread rates between nodes should also be symmetric. This is formalized as:

$$\begin{aligned} b_v^{\mathrm{c}} &\neq b_v^{\mathrm{i}} \\ t_{rv}^{\mathrm{c}} &= t_{rv}^{\mathrm{i}} \end{aligned} \tag{11}$$

   for all $v \leq V$ and $r \in \operatorname{pa}(v)$.
3. **Asymmetric Spread from Tumor**: Tumor spread to the contralateral side varies depending on whether the tumor crosses the mid-sagittal plane. When the tumor extends over the midline, its spread rate to the contralateral LNLs should be larger than if the tumor were lateralized, but still smaller than the tumor's spread to the ipsilateral side. To model this, we introduce a mixing parameter $\alpha \in [0, 1]$, representing the "degree of asymmetry" effectively mixing these two extremes:

$$b_v^{\mathrm{c}, \epsilon=\texttt{True}} = \alpha \cdot b_v^{\mathrm{i}} + (1 - \alpha) \cdot b_v^{\mathrm{c}, \epsilon=\texttt{False}} \tag{12}$$

   This parameterization results in three sets of rates for tumor spread to the LNLs: (1) the spread to ipsilateral LNLs b_v^{i}, (2) the spread to contralateral LNLs as long as the tumor is lateralized $b_v^{\mathrm{c}, \epsilon=\texttt{False}}$, (3) the spread to contralateral LNLs when the tumor crosses the midline $b_v^{\mathrm{c}, \epsilon=\texttt{True}}$. Despite this additional complexity, these rates collectively account for only $2 \cdot 2^V + 1$ parameters due to their coupling via α.

The parameter space now becomes:

$$\theta = \left(\{b_v^{\mathrm{i}}\}, \{b_v^{\mathrm{c}}\}, \alpha, \{t_{vr}\}, p_{\mathrm{early}}, p_{\mathrm{adv.}}\right) \quad \text{with} \quad \begin{smallmatrix} v \leq V \\ r \in \mathrm{pa}(v) \end{smallmatrix} \tag{13}$$

This results in less than a doubling of parameters compared to the unilateral model. From these parameters, we construct three transition matrices: the unchanged $\mathbf{A}_{\mathrm{i}}$ for the ipsilateral side, $\mathbf{A}_{\mathrm{c}}^{\epsilon=\texttt{False}}$ for contralateral progression while the tumor is lateralized, and $\mathbf{A}_{\mathrm{c}}^{\epsilon=\texttt{True}}$ for cases where the tumor crosses the mid-sagittal plane.

Modelling Midline Extension

Most tumors crossing the midline at the time of diagnosis likely began as lateralized tumors that grew over the midline at a later stage. As a result, the transition matrix $\mathbf{A}_c^{\epsilon=\texttt{True}}$ applies only to a subset of time-steps.

To account for this, we model the tumor's extension over the mid-sagittal line as a binary random variable. A tumor starts as lateralized, with a finite probability p_ϵ at each time step of crossing the midline. The overall probabilities of a patient having a clearly lateralized tumor or one extending over the mid-sagittal line after t time steps are then given by

$$P(\epsilon = \texttt{False} \mid t) = (1 - p_\epsilon)^t$$
$$P(\epsilon = \texttt{True} \mid t) = 1 - P(\epsilon = \texttt{False} \mid t)$$

Using this, it is straightforward to write down the matrix of state distributions for all time-steps, as in equation 4 covering the contralateral hidden state evolution:

$$\Lambda_c^{\epsilon=\texttt{False}} = \begin{pmatrix} \pi^\top \cdot (\mathbf{A}_c^{\epsilon=\texttt{False}})^0 \\ (1 - p_\epsilon) \cdot \pi^\top \cdot (\mathbf{A}_c^{\epsilon=\texttt{False}})^1 \\ \vdots \\ (1 - p_\epsilon)^{t_{\max}} \cdot \pi^\top \cdot (\mathbf{A}_c^{\epsilon=\texttt{False}})^{t_{\max}} \end{pmatrix}$$

where we used the transition matrix $\mathbf{A}_c^{\epsilon=\texttt{False}}$ that depends on the base spread parameters $b_v^{c,\epsilon=\texttt{False}}$.

The case of midline extension is more complex: we already marginalize over the exact time step when the tumor grows over the mid-sagittal line. However, at the point of crossing, the contralateral transition matrix must switch to the increased spread rates, $b_v^{c,\epsilon=\texttt{True}}$, as defined by the linear mixing in equation 12. To correctly perform this marginalization, we iteratively construct the joint distribution $P\left(\mathbf{X}^c, \epsilon = \texttt{True} \mid t\right)$.

We begin at $t = 0$, where all contralateral LNLs are healthy (i.e., $\mathbf{X}_c = \xi_0$) and the tumor is lateralized ($\epsilon = \texttt{False}$):

$$P\left(\mathbf{X}^c = \xi_0, \epsilon = \texttt{False} \mid t = 0\right) = 1$$

while all other states have zero probability.

At some later time step $t = \tau + 1$, there are two scenarios to marginalize over:

1. **The tumor was lateralized at $t = \tau$ and grew over the midline at $t = \tau+1$:** In this case, the probability of midline extension at $t = \tau+1$ is p_ϵ. This probability weights the contralateral state distribution that had previously evolved without increased contralateral spread.

2. **The tumor had already crossed the midline before** $t = \tau$**:**
   Here, the tumor remains in the midline-crossed state with probability 1. To account for this scenario, we simply include the distribution $P\left(\mathbf{X}^{\mathrm{c}}, \epsilon = \mathtt{True} \mid \tau\right)$ from the previous time step.

Combining these scenarios leads to a recursive formulation:

$$P\left(\mathbf{X}^{\mathrm{c}}, \epsilon = \mathtt{True} \mid \tau + 1\right) = \left[p_\epsilon P\left(\mathbf{X}^{\mathrm{c}}, \epsilon = \mathtt{False} \mid \tau\right) + P\left(\mathbf{X}^{\mathrm{c}}, \epsilon = \mathtt{True} \mid \tau\right)\right]^\top \cdot \mathbf{A}_{\mathrm{c}}^{\epsilon = \mathtt{True}}$$

We can collect the iteratively computed distributions for the midline extension case to define the matrix over the states given all time-steps, in analogy to equation 4:

$$\Lambda_{\mathrm{c}}^{\epsilon = \mathtt{True}} = \begin{pmatrix} P\left(\mathbf{X}^{\mathrm{c}}, \epsilon = \mathtt{True} \mid 0\right) \\ P\left(\mathbf{X}^{\mathrm{c}}, \epsilon = \mathtt{True} \mid 1\right) \\ \vdots \\ P\left(\mathbf{X}^{\mathrm{c}}, \epsilon = \mathtt{True} \mid t_{\max}\right) \end{pmatrix}$$

Using this, we can again write the joint of ipsi- and contralateral involvement - now also for the case of mid-sagittal extension - algebraically as before in equation 10:

$$P\left(\mathbf{X}^{\mathrm{i}} = \xi_n, \mathbf{X}^{\mathrm{c}} = \xi_m, \epsilon\right) = \left[\Lambda_{\mathrm{i}}^\top \cdot \mathrm{diag}\, P(\mathbf{t}) \cdot \Lambda_{\mathrm{c}}^\epsilon\right]_{n,m}$$

With the above, we compute the likelihood of all patients with and without midline extension separately. And if for some patients the information of tumor lateralization is not available, we can simply marginalize over the unknown variable $\epsilon \in \{\mathtt{False}, \mathtt{True}\}$.

The final parameter space of our extended model has now reached this size:

$$\theta = \left(\{b_v^{\mathrm{i}}\}, \{b_v^{\mathrm{c}}\}, \alpha, \{t_{vr}\}, p_{\mathrm{early}}, p_{\mathrm{adv.}}, p_\epsilon\right) \quad \text{with} \quad \substack{v \leq V \\ r \in \mathrm{pa}(v)} \tag{14}$$

Model Prediction in the Bayesian Context

Our stated goal is to compute the risk for a patient's true ipsi- and contralateral nodal involvement states $\mathbf{X}^{\mathrm{i}}$ and $\mathbf{X}^{\mathrm{c}}$, *given* their individual diagnosis $d = \left(\zeta_k^{\mathrm{i}}, \zeta_\ell^{\mathrm{c}}, \epsilon, \mathrm{T}x\right)$. Here, this diagnosis consists of the observed ipsi- and contralateral nodal involvements, the patient's midline extension ϵ, and their tumor's T-category $\mathrm{T}x$. Using Bayes' law, we can write this risk as:

$$P(\mathbf{X}^{\mathrm{i}}, \mathbf{X}^{\mathrm{c}} \mid d, \hat{\theta}) = \frac{P\left(\zeta_k^{\mathrm{i}} \mid \mathbf{X}^{\mathrm{i}}\right) P\left(\zeta_\ell^{\mathrm{c}} \mid \mathbf{X}^{\mathrm{c}}\right) P(\mathbf{X}^{\mathrm{i}}, \mathbf{X}^{\mathrm{c}}, \epsilon \mid \hat{\theta}, \mathrm{T}x)}{\sum_{i=0}^{2^V} \sum_{j=0}^{2^V} \mathcal{C}_{ij}} \tag{15}$$

with the normalization constants

13

$$\mathcal{C}_{ij} = P\left(\zeta_k^i \mid \mathbf{X}^i = \xi_i^i\right) P(\zeta_\ell^c \mid \mathbf{X}^c = \xi_j^c) P(\mathbf{X}^i = \xi_i^i, \mathbf{X}^c = \xi_j^c, \epsilon \mid \hat{\theta}, \mathrm{T}x)$$

The terms $P\left(\zeta_k^i \mid \mathbf{X}^i\right)$ and $P\left(\zeta_\ell^c \mid \mathbf{X}^c\right)$ are defined solely by sensitivity and specificity of the diagnostic modality. Terms like these already appeared in the definition of the observation matrx in equation 3. The *prior* $P(\mathbf{X}^i, \mathbf{X}^c, \epsilon \mid \hat{\theta}, \mathrm{T}x)$ in the above equation is the crucial term that is supplied by a trained model and its parameters $\hat{\theta}$.

It is possible to compute this *posterior* probability of true involvement not only for one fully defined state $(\mathbf{X}^i, \mathbf{X}^c)$, but also for e.g. individual LNLs: For example, the risk for involvement in the contralateral level IV would be a marginalization over all ipsilateral states ξ_i^i and all contralateral states ξ_j^c where $\xi_{j4}^c = 1$. Formally:

$$P(\mathrm{cIV} \mid \mathbf{Z}^i = \zeta_k^i, \mathbf{Z}^c = \zeta_\ell^c, \hat{\theta}, \mathrm{T}x) = \sum_k \sum_{\ell\,:\,\xi_{\ell4}=1} P(\mathbf{X}^i = \xi_k^i, \mathbf{X}^c = \xi_\ell^c \mid \zeta_k^i, \zeta_\ell^c, \epsilon, \hat{\theta}, \mathrm{T}x)$$

(16)

Computational methods

This section details the experimental setup. All figures, tables, and results are fully reproducible via the GitHub repository rmnldwg/bilateral-paper.

Involvement Data Consensus

Rather than providing the model with a patient's diagnoses from multiple modalities with different sensitivities and specificities (e.g. PET-CT *and* MRI), we combined them into a single "consensus" diagnosis before inference. This approach was necessary because literature-reported diagnostic imaging sensitivities and specificities did not plausibly match our observations. For instance, in the USZ cohort, 78% of OPSCC patients were diagnosed with ipsilateral LNL II involvement -— a finding difficult to reconcile with typical imaging sensitivities around 80% and specificities below 100% [14, 15].

MCMC Sampling

We used the Python package `emcee` [17] for parameter inference, implementing efficient MCMC sampling with parallel affine-invariant samplers. The sampling algorithms employed differential evolution moves [18, 19], with the likelihood implemented by our `lymph-model` Python package.

We initialized 12 parallel samplers ("walkers") with random values from the unit cube, effectively representing a uniform prior distribution over the model parameters. Convergence was determined by two criteria:

1. The change in autocorrelation time was less than 5.0e-2.

2. The autocorrelation estimate dropped below n / 50, where n is the chain length. Earlier autocorrelation estimates might not be trustworthy.

Samples from this *burn-in phase* before convergence were discarded. After that, we drew 10 additional samples, spaced 10 steps apart.

We verified sampling convergence in figure 3 by examining the MCMC chain's autocorrelation time and walker acceptance fractions.
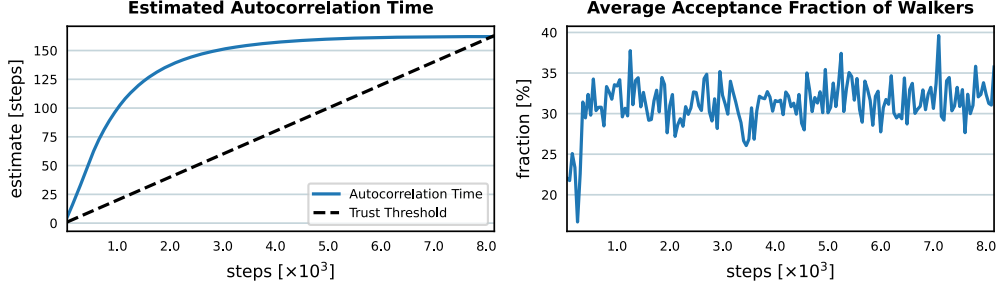

Figure 3: Burn-in phase monitoring of MCMC sampling. Left: Estimated autocorrelation time, indicating converging when stable and below the trust threshold. Right: Average acceptance fraction of parallel walkers, with ~30% indicating good mixing.

Computing the Observed and Predicted Prevalence of Involvement Patterns

We evaluate the model's ability to approximate the observed distribution of lymphatic involvement patterns. We compare the prevalence of selected involvement patterns in the data to the model's predicted prevalence, given patient scenarios.

A "scenario" includes the patient's T-category Tx and whether the tumor extended over the mid-sagittal line, i.e. $\epsilon = \texttt{True}$ or $\epsilon = \texttt{False}$. An involvement pattern specifies all ipsi- and contralateral LNLs' status as "healthy", "involved", or "masked" (ignored).

For example, we may assess contralateral LNL II involvement prevalence under early T-category (T0-T2) and no midline extension ($\epsilon = \texttt{False}$). In the data, 379 such patients were observed, with 27 exhibiting contralateral LNL II involvement – a prevalence of 7.12%.

To visualize the data prevalence, we plot a *beta posterior* over the true prevalence q. It is the result of a uniform beta prior over q, multiplied with the binomial likelihood for k out of n patients, given q. The resulting distribution has its maximum at $q = k/n$ and nicely captures the uncertainty in the observed cohort: if only 3 out of 10 patients are observed with a specific pattern, the distribution is much wider than if we observe 300 out of 1000.

The model's predicted prevalence to compare it with is computed as:

$$P\left(\mathrm{II^c} \mid \epsilon = \mathtt{False}, \mathrm{T}x = \mathrm{early}\right) = \frac{P\left(\mathrm{II^c}, \epsilon = \mathtt{False} \mid \mathrm{T}x = \mathrm{early}\right)}{P\left(\epsilon = \mathtt{False} \mid \mathrm{T}x = \mathrm{early}\right)}$$

In the enumator, we marginalize over all ipsi- and contralateral LNLs' involvements, except for LNL II contralaterally. This is similar to the marginalization in equation 16, although we are summing over different quantities. In the denominator, we can simply insert the joint distribution over midline extension and diagnose time $P\left(\epsilon, t\right)$ marginalized over t using the early T-category's time-prior.

We display model predictions as histograms, each value computed from one of the MCMC samples. Ideally, these approximate the location and width of the beta posteriors from the data showing an accurate and precise fit.

Note that we omit the y-axis in these figures, as their numerical value is not intuitively interpretable. We instead use the free space to label e.g. rows in an array of subplots.

Results: Model evaluation

In table 2, we tabulate the mean and standard deviation of the sampled parameters. The bilateral model mostly reproduces the ipsilateral spread parameter values reported in the earlier publication on the unilateral model [10]. Any discrepancies may be due to differences in the patient cohorts. Therefore, we omit the analysis of the ipsilateral spread patterns and focus instead on the analysis of contralateral involvement. Lastly, note the relatively small contralateral spread parameters b_v^c compared to the ipsilateral parameters b_v^i. This accurately reflects the low prevalence of contralateral lymph node involvement for lateralized tumors.

Source: Article Notebook

Table 2: Mean sampled parameter estimates of the midline model and the respective standard deviation. The parameters set to fixed values are the maximum number of time steps $t_{max} = 10$ and the time prior parameter for early T-category patients $p_{early} = 0.3$.

Table 2

Parameter	Mean	Std. Dev.
Mid. ext. probability	8.16%	± 0.48%
ipsi: T I	2.80%	± 0.26%
ipsi: T II	34.89%	± 1.40%
ipsi: T III	5.45%	± 0.66%
ipsi: T IV	0.94%	± 0.18%
ipsi: T V	1.83%	± 0.22%
ipsi: T VII	2.32%	± 0.26%
contra: T I	0.29%	± 0.09%

Parameter	Mean	Std. Dev.
contra: T II	2.46%	± 0.29%
contra: T III	0.14%	± 0.07%
contra: T IV	0.19%	± 0.08%
contra: T V	0.05%	± 0.04%
contra: T VII	0.50%	± 0.17%
Mixing	33.87%	± 4.32%
I II	62.50%	± 16.69%
II III	14.23%	± 1.64%
III IV	15.86%	± 1.93%
IV V	14.58%	± 3.76%
late T-cat. binom. prob.	44.98%	± 1.99%

Source: Article Notebook

Illustration of the model

In this subsection, we illustrate key aspects of the mathematical framework introduced earlier. The top panel of figure 4 shows the prior distribution over diagnosis times, $P(t)$. Based on the parameterization, early T-category tumors are on average diagnosed after 3 time steps, while advanced T-category tumors are diagnosed later, averaging 4.5 time steps. This is due to the learned value of $p_{\text{adv.}}$, which is 44.98% for advanced T-category tumors. The tumor's average probability per time step of growing over the midline, p_ϵ, was found to be 8.16%. Using this value, the conditional probability of midline extension, $P(\epsilon \mid t)$, can be computed for a given time step t (red line in the top panel of figure 4). The bottom panel visualizes the joint probability $P(\epsilon, t)$, showing the likelihood of diagnosis at time t with specific states of midline extension and T-category.

The framework models the joint probability distribution of midline extension and ipsi- and contralateral lymph node involvement, $P\left(\mathbf{X}^{\text{i}}, \mathbf{X}^{\text{c}}, \epsilon\right)$. This is visualized in figure 5, which represents the calculation defined in equation 10. To simplify interpretation, the example focuses only on LNLs II, III, and IV, reducing the state space to $2^3 = 16$ possible states per side, and $2 \times 8 \times 8 = 128$ total states. LNLs I, V, and VII are excluded, along with their spread parameters, while remaining parameters are set to their mean values from table 2.

The left matrix in figure 5 shows the time evolution of the probability distribution over the ipsilateral involvement states, starting from the healthy state $[0, 0, 0]$. The two right matrices show the contralateral state evolution, distinguishing between the cases of midline extension and no midline extension. At $t = 0$, the contralateral neck begins in the healthy state $[0, 0, 0]$ without midline extension. The central matrix shows the time prior for late T-category tumors. The matrix multiplication results in the joint distribution $P\left(\mathbf{X}^{\text{i}}, \mathbf{X}^{\text{c}}, \epsilon = \texttt{True}\right)$, visualized in figure 6.
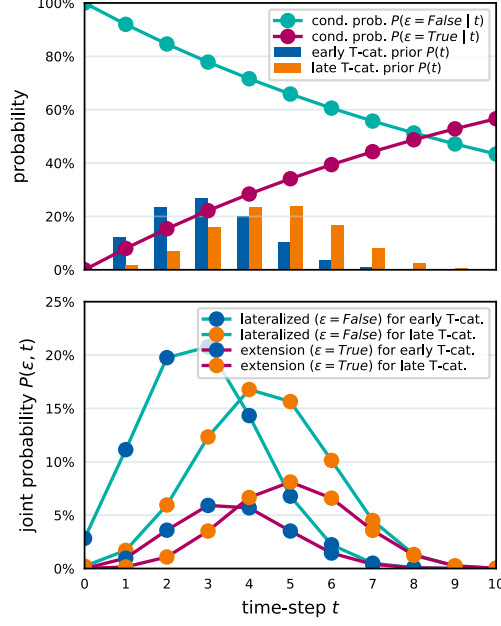
17

Figure 4: The top panel shows the prior probability to be diagnosed at time step t for early and late T-category tumors as bars. The conditional probability of midline extension ($\epsilon = \texttt{True}$) given time step t is shown as a line plot. The bottom panel illustrates the joint probability of being diagnosed at time t and having a tumor that crosses the midline.

This joint distribution is presented as two heatmaps, corresponding to the two states of midline extension. The most likely state involves a lateralized tumor with ipsilateral level II involvement and no contralateral involvement, having a probability of approximately 25%. The next most probable state involves ipsilateral levels II and III without contralateral involvement. The most likely state with contralateral involvement corresponds to tumors with midline extension, showing involvement of contralateral level II and ipsilateral levels II and III.

Prevalence predictions for contralateral involvement

The bilateral model was designed to meet the requirements outlined in section 3. Here, we evaluate the model's ability to quantitatively capture the observed patterns of lymph node involvement in the dataset. Specifically, we compare the model's predictions for contralateral involvement to the observed data across scenarios that vary by T-category, midline extension, and ipsilateral involvement.
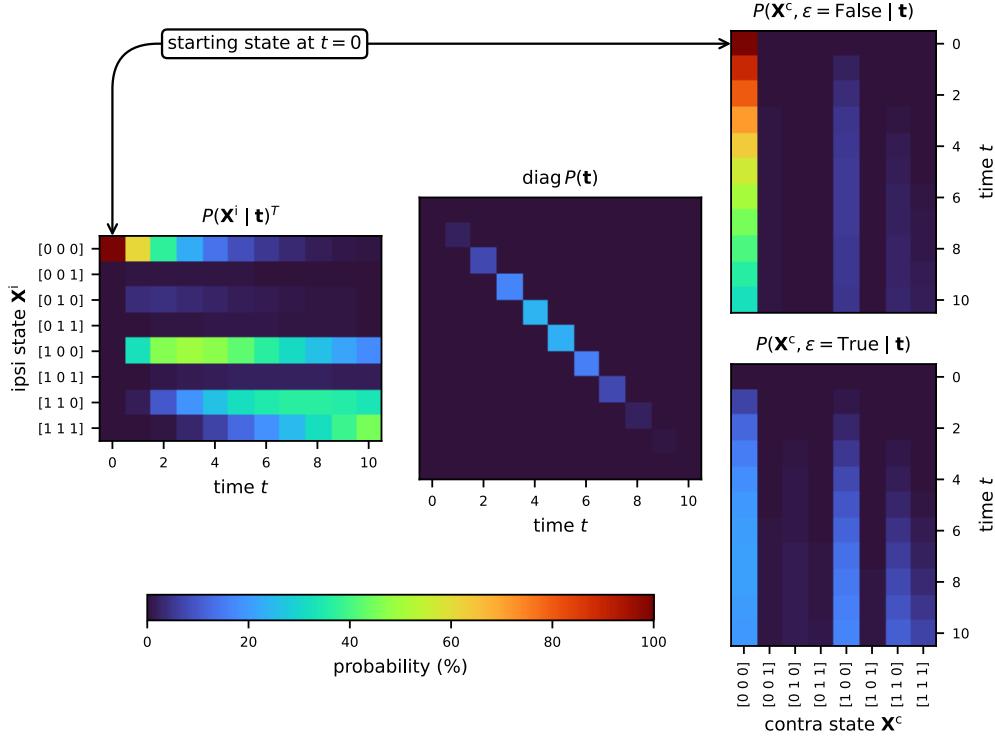
Figure 5: Visual representation of equation 10 for midline extension. The left and right matrices represent the time evolution of hidden states for the ipsi- and contralateral necks, respectively. The right matrices distinguish between the cases of no midline extension (top) and midline extension (bottom). The central diagonal matrix shows the time-prior for late T-category tumors. This computation yields the joint distribution $P(\mathbf{X}^i, \mathbf{X}^c, \epsilon)$, visualized in figure 6.

Dependence of Contralateral Involvement on T-Category and Midline Extension

In figure 7, we compare the prevalence of contralateral involvement for LNLs II, III, and IV across four scenarios: combinations of early and late T-category, and lateralized versus midline-extending tumors.

Figure 7 demonstrates the model's ability to account for key risk factors, such as T-category and midline extension. Consistent with the data, the model predicts that the prevalence of contralateral LNL II involvement increases from 7.1% for early T-category lateralized tumors to 39.17% for advanced T-category tumors that cross the midline. Similarly, contralateral LNL III involvement rises from around 1.55% for early T-category lateralized tumors to nearly 14.25% for advanced T-category midline-extending tumors.
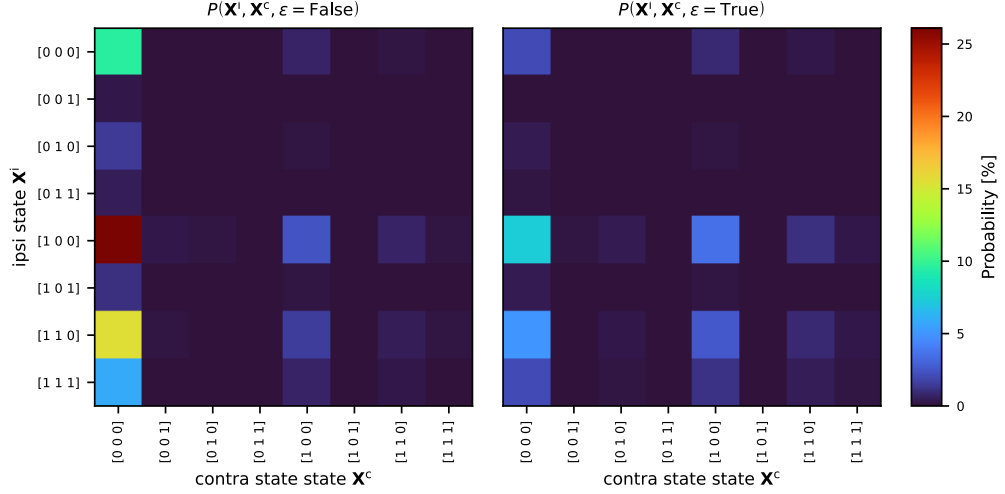
19

Figure 6: The full distribution $P\left(\mathbf{X}^{\mathrm{i}}, \mathbf{X}^{\mathrm{c}}, \epsilon = \text{False}\right)$ over ipsi- and contralateral states and midline extension for late T-category tumors. This three-dimensional distribution is shown as two separate heatmaps for the binary midline extension variable ϵ. These matrices are the result of equation 10, visualized in figure 5.
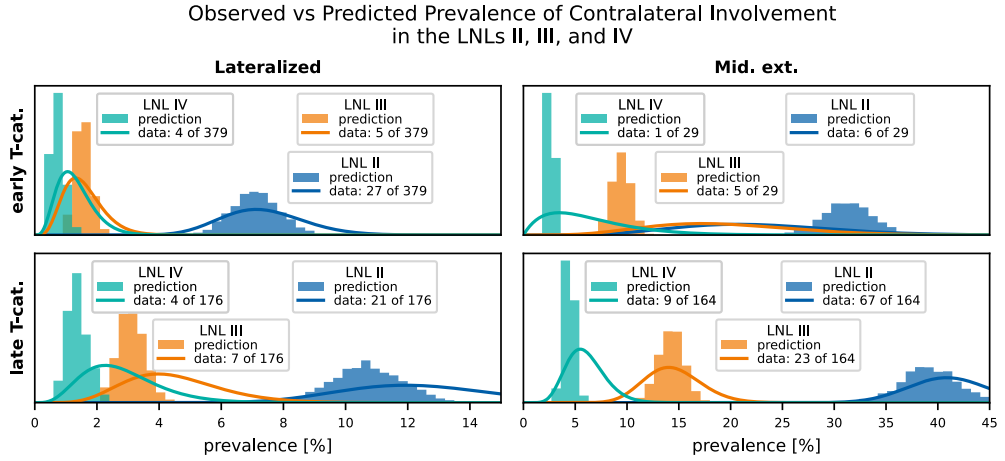

Figure 7: Comparison of predicted (histograms) vs observed (beta posteriors) prevalences, shown for the contralateral LNLs II (blue), III (orange), and IV (green). The top row shows scenarios with early T-category tumors, the bottom row for late T-category ones. The left column depicts scenarios where the primary tumor is clearly lateralized, the right column scenarios of tumors extending over the mid-sagittal line. This figure illustrates the model's ability to describe the prevalence of involvement for different combinations of the risk factors T-category and midline extension.
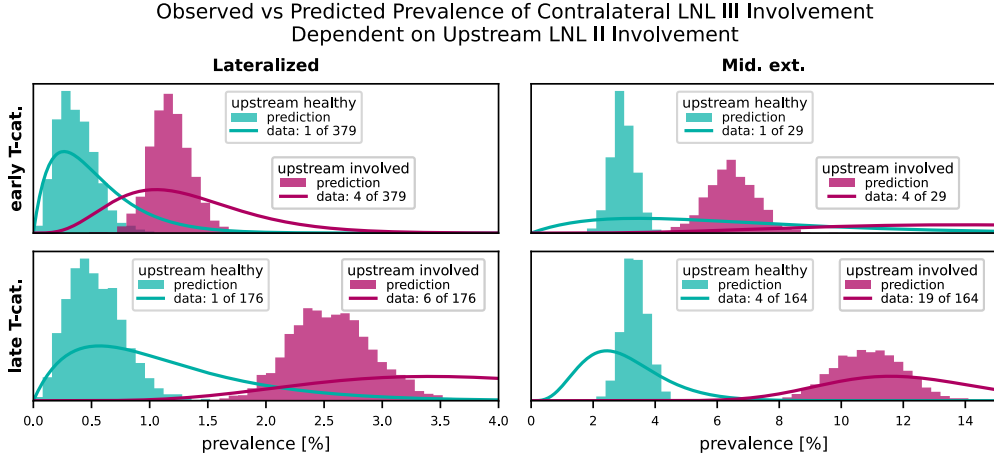
Figure 8: The influence of the upstream LNL II's involvement on the prevalence of contralateral level III for the four combinations of tumor lateralization (lateralized or extending over midline) and T-category (early or advanced). Our model predictions (histograms) are plotted against the observations in the data (beta posteriors).

Influence of Upstream Involvement on Contralateral Metastasis

Figure 8 highlights the influence of upstream LNL II involvement on contralateral LNL III metastasis. The contralateral LNL III rarely harbors metastases when its upstream LNL II is healthy, a correlation well-captured by the model. However, for early T-category tumors extending over the midline, the model slightly underestimates the observed prevalence, likely due to the relative rarity of such cases, resulting in broader uncertainty in the observed prevalence estimates.

Correlation between Ipsi- and Contralateral Involvement

In figure 9, the model's ability to capture the correlation between ipsi- and contralateral involvement is demonstrated. The marginals of the joint distribution highlight contralateral LNL II involvement alongside varying ipsilateral LNL involvement states. Despite having no direct connections between the two sides, the model successfully predicts these correlations, which arise purely through the shared diagnosis time.

For example, the model accurately predicts that contralateral LNL II involvement is rare when the ipsilateral neck is completely healthy (green histograms). However, if ipsilateral LNL II is involved, contralateral involvement becomes more likely. Notably, the model achieves this without specific parameters to quantify ipsi-contralateral correlations, relying instead on the inherent structure of the time-dependent dynamics.
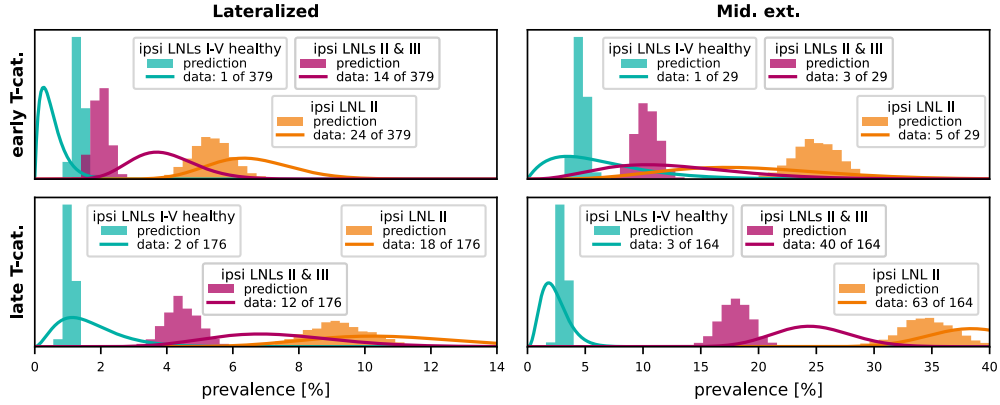
Figure 9: Comparison of the computed and observed prevalences for scenarios that illustrate the model's capability of accounting for the correlation between ipsi- and contralateral involvement. We show three scenarios where we consider the joint involvement of contralateral LNL II together with different ipsilateral involvements: 1) the ipsilateral neck shows no involvement in green (LNLs I to V are healthy, LNL VII is unspecified because data on it is missing for some patients), 2) where ipsilateral LNL II is involved in orange (LNLs I, III, IV, and V are healthy), and 3) where ipsilateral LNLs II and III are involved in red (LNLs I, IV, and V are healthy). These three scenarios are plotted for all combinations of T-category (early in top row, advanced in bottom row) and tumor lateralization (lateralized in left column, extending over mid-sagittal line in the right column).

Results: Prediction of Risk for Occult Disease

For clinical applications, the model needs to estimate the risk of occult metastases in clinically negative LNLs based on a patient's diagnosis. The diagnosis includes the T-category, tumor lateralization (whether it extends over the mid-sagittal plane), and the clinically detected involvement of LNLs based on imaging and possibly fine needle aspiration (FNA).

We assume imaging detects lymph node involvement with a sensitivity of 81% and a specificity of 76% [14], while FNA has a sensitivity of 80% and a specificity of 98%. This implies that FNA-confirmed involvement is highly reliable, with almost no false positives.

Contralateral LNL II

The predicted risk of occult disease in contralateral LNL II is shown in the left panel of figure 10. Tumor lateralization is the strongest determinant of risk. For example:

- A patient with a lateralized early T-category tumor and ipsilateral LNL II involvement has a predicted risk of 1.55% for occult contralateral LNL II disease (green
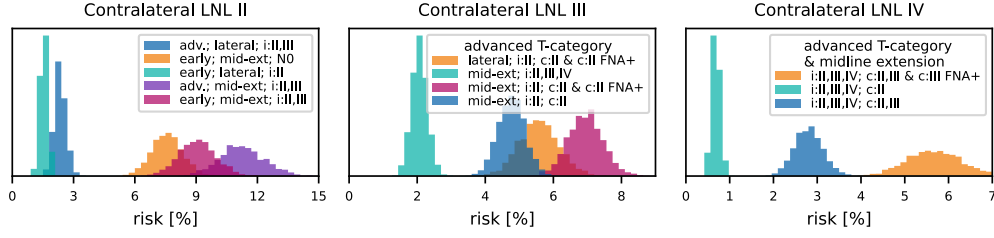
Figure 10: Histograms over the predicted risk of occult involvement in contralateral LNL II (left), III (middle), and IV (right), shown for various combinations of T-category, tumor lateralization, and clinical LNL diagnoses. All LNLs not explicitly mentioned in the legend, including the LNL for which the risk of occult disease was computed, were assumed to be clinically negative (specificity 76%, sensitivity 81%).

histogram).

- If the same tumor extends over the midline, the risk increases to 7.58% (orange histogram).

Advanced T-category further increases risk but has less impact than midline extension. For instance:

- An early T-category tumor crossing the midline with ipsilateral involvement of LNLs II and III has a 9.06% risk of contralateral LNL II disease (red histogram).

- For the same scenario with an advanced T-category tumor, the risk rises to 11.25% (purple histogram).

The degree of ipsilateral involvement also influences risk. For a midline-extending early T-category tumor, a clinically N0 ipsilateral neck predicts a risk of 7.58%, compared to 9.06% when LNLs II and III are involved.

In summary, midline extension is the primary risk factor for contralateral LNL II involvement, but advanced T-category and extensive ipsilateral involvement also contribute significantly.

Contralateral LNL III

As shown in the center panel of figure 10, the risk of occult contralateral LNL III involvement rarely exceeds 5% and depends strongly on upstream LNL II involvement:

- For an advanced T-category tumor extending over the midline and with extensive clinical involvement in the ipsilateral LNLs II, III, and IV, but a clinically negative contralateral neck, the risk is only 2.07% (green histogram).

23

- If contralateral LNL II is clinically involved, the risk for contralateral LNL III rises to 4.77% (blue histogram).

- When contralateral LNL II involvement is confirmed by FNA, the risk further increases to 6.94% (red histogram).

Even for lateralized tumors, FNA-confirmed involvement in contralateral LNL II predicts a 5.58% risk for LNL III involvement (orange histogram). This highlights the importance of upstream LNL II involvement in determining the risk for LNL III.

Contralateral LNL IV

For contralateral LNL IV, the predicted risk is below 1% in most scenarios, even for advanced T-category tumors extending over the midline with extensive ipsilateral and contralateral involvement (green histogram).

- If contralateral LNL III is also clinically involved, the risk increases to 2.81% (blue histogram).

- When contralateral LNL III involvement is confirmed by FNA, the risk rises significantly to 5.68% (orange histogram).

This higher risk occurs because FNA confirmation eliminates the possibility of false-positive diagnoses for contralateral LNL III, strongly increasing the likelihood of downstream LNL IV involvement.

Contralateral LNLs I, V, and VII

Contralateral LNLs I, V, and VII show very low predicted risks for occult disease. Even in advanced T-category tumors extending over the midline with extensive ipsi- and contralateral involvement, the risk for LNLs I and VII remains below 3%.

For contralateral LNL V, the risk is negligible unless there is severe contralateral involvement, including LNL IV, confirmed by FNA. In such extreme cases, the risk increases to 5.49%.

Discussion

Summary

This work introduces a formalism to model ipsi- and contralateral lymph node involvement in oropharyngeal SCC patients. Building on a previously developed ipsilateral model [9, 10], we extend it to the contralateral side while preserving the original model's structure. Our extension is both intuitive and interpretable, with parameters learned from a dataset of 833 patients across four institutions. The model uses clinically diagnosed LNL involvement, tumor T-category, and lateralization to provide personalized risk predictions for occult disease in any LNL of interest.

Despite its relatively few parameters, the model is highly interpretable, with each parameter having a clear, intuitive explanation. At the same time, it effectively describes the observed data on ipsilateral and contralateral nodal involvement. To our knowledge, this represents the most comprehensive model of lymphatic tumor progression in oropharyngeal SCC, surpassing prior efforts, which were conceptually different, limited in scope, or not trained on real patient data [20, 21]. The underlying dataset and code are publicly available, supporting reproducibility and further development.

Implications for Contralateral Elective Nodal Treatment

The model's predictions are already guiding a clinical trial on volume de-escalation at the University Hospital Zurich [22]. Based on a 5% acceptable risk threshold for occult disease in any LNL, the model recommends the following approach for contralateral elective irradiation, assuming the respective LNL is clinically healthy:

- **Lateralized tumors with no contralateral clinical involvement:**
  Unilateral radiotherapy is sufficient, regardless of T-category or ipsilateral involvement.

- **Tumors extending over the midline with no contralateral clinical involvement:**
  Elective irradiation is limited to LNL II.

- **Contralateral LNL III:**
  Irradiate only if LNL II is involved, regardless of tumor lateralization, T-category, or ipsilateral involvement. If contralateral LNL II is clinically negative, LNL III is not irradiated unless LNL IV is involved.

- **Contralateral LNL IV:**
  Irradiate only when LNL III involvement is confirmed.

- **Contralateral LNL V:**
  Elective irradiation is generally unnecessary. Only in extreme cases, such as advanced T-category tumors with midline extension and confirmed contralateral involvement in LNLs II to IV, is irradiation of LNL V considered.

- **Contralateral LNLs I and VII:**
  Elective irradiation is not recommended unless these nodes are clinically involved.

These recommendations align with the model results discussed in section 3, providing a basis for refined treatment guidelines. However, they should be interpreted in light of the limitations discussed in section 3 below.

Limitations and Future Work

T-Category Dependence

The model uses a single parameter, $p_{\text{adv.}}$, to account for differences in lymph node involvement patterns between early and advanced T-category tumors. Advanced T-category tumors are modeled as evolving over more time steps, while the probability of spread per time step, governed by the b_v parameters, remains constant. While this approach captures the overall differences between early and advanced T-category tumors well (as seen in figure 7), it is not perfect.

For example: - The observed differences between early and advanced T-category tumors may sometimes be greater or smaller than the model's predictions. This has been previously noted for ipsilateral LNL I involvement [10]. - For contralateral spread, the prevalence of midline extension is slightly overestimated for early T-category tumors and underestimated for advanced T-category tumors, as discussed in section 3.

Sensitivity and Specificity

As noted in section 3 and further detailed in section 3, we assumed that the consensus across diagnostic modalities reflects the true state X_v of lymph node involvement. While this assumption is reasonable for pathologically confirmed diagnoses, it is an approximation for clinically diagnosed involvement, which cannot detect occult disease by definition.

The model could, in principle, distinguish between pathologically confirmed and clinically diagnosed involvement by incorporating different sensitivity and specificity values for each diagnostic modality. However, this was not implemented in the current work for two reasons: 1. **Simplified evaluation:** This approximation allowed for direct comparison between the observed prevalence of involvement and the model's predictions, enabling an evaluation of the model's ability to describe the data with few interpretable parameters. 2. **Inconsistent literature values:** Reported sensitivity and specificity values for diagnostic modalities show inconsistencies with some observed data patterns, complicating their integration into the model.

Future efforts could focus on developing methods to rigorously differentiate between pathologically confirmed and clinically diagnosed involvement, potentially improving model accuracy for clinical settings.

Tumor Subsites

Oropharyngeal tumors occur in distinct subsites, such as the base of the tongue or the tonsils, which may exhibit slightly different lymphatic metastatic patterns. Incorporating subsite-specific information into the model could enhance its predictive accuracy. Preliminary investigations suggest that a mixture model may effectively capture these subsite-specific spread patterns [23, 24].

This approach could also facilitate extending the model to other tumor locations, such as the oral cavity, hypopharynx, and larynx. Including these additional tumor sites

would broaden the model's applicability and improve its clinical utility for head and neck cancer patients.

Acknowledgement

Consensus on most likely involvement

The consensus on the most likely involvement state of a LNL was formed as follows: Suppose the involvement status X_v of LNL v was assessed using different diagnostic modalities $\mathcal{O} = \{\text{MRI}, \text{CT}, \text{pathology}, ...\}$, each characterized by their own pair of sensitivity and specificity values s_N^o and s_P^o, with $o \in \mathcal{O}$. These values are tabulated in table 3. Then we have $|\mathcal{O}|$ observations $z_v^o \in [0, 1]$, where 0 stands for "healthy" and 1 for "involved". We can then compute the most likely true involvement X_v using the likelihood function

$$\ell\left(X_v \mid \{z_v^o\}_{o \in \mathcal{O}}\right) = \prod_{o \in \mathcal{O}} (1 - X_v) \cdot [z_v^o \cdot (1 - s_P^o) + (1 - z_v^o) \cdot s_P^o]$$
$$+ X_v \cdot [z_v^o \cdot s_N^o + (1 - z_v^o) \cdot (1 - s_N^o)]$$

We now assume the true state X_v to take on the value 1 if $\ell\left(X_v = 1 \mid ...\right) > \ell\left(X_v = 0 \mid ...\right)$ and 0 otherwise. For example, if we have $z_{\text{II}}^{\text{CT}} = 0$ and $z_{\text{II}}^{\text{MRI}} = 1$ we would compute the following likelihoods:

$$\ell\left(X_{\text{II}} = 1 \mid z_{\text{II}}^{\text{CT}} = 0, z_{\text{II}}^{\text{MRI}} = 1\right) = (1 - s_N^{\text{CT}}) \cdot s_N^{\text{MRI}} = 15.39\%$$
$$\ell\left(X_{\text{II}} = 0 \mid z_{\text{II}}^{\text{CT}} = 0, z_{\text{II}}^{\text{MRI}} = 1\right) = s_P^{\text{CT}} \cdot (1 - s_N^{\text{MRI}}) = 14.44\%$$

In this example, we would thus assume the true state to be involved ($X_{\text{II}} = 1$).

This method of computing a consensus also ensures that the pathology reports always override any conflicting clinical diagnosis, due to pathology's high sensitivity and specificity.

Table 3: Specificity and sensitivity values from the literature [14, 15].

Modality	Specificity	Sensitivity
CT	76%	81%
PET	86%	79%
MRI	63%	81%
FNA	98%	80%

27

Modality	Specificity	Sensitivity
pathology	100%	100%

Contralateral Prevalence of Involvement

Source: Article Notebook

Table 4: Contralateral involvement depending on whether the primary tumor extends over the mid-sagittal line, the T-category, and whether the ipsilateral LNL III was involved or healthy.

Table 4

T-cat.	ipsi	Mid. ext.	I		II		III		IV		total
			n	%	n	%	n	%	n	%	n
early	0	False	0	0.00	1	1.16	0	0.00	0	0.00	86
early	0	True	0	0.00	1	10.00	1	10.00	0	0.00	10
early	0	nan	0	0.00	0	0.00	0	0.00	0	0.00	12
early	1	False	1	0.53	11	5.82	2	1.06	1	0.53	189
early	1	True	1	11.11	2	22.22	0	0.00	0	0.00	9
early	1	nan	0	0.00	3	9.68	0	0.00	1	3.23	31
early	2	False	1	0.96	15	14.42	3	2.88	3	2.88	104
early	2	True	0	0.00	3	30.00	4	40.00	1	10.00	10
early	2	nan	0	0.00	3	13.64	0	0.00	0	0.00	22
advanced	0	False	0	0.00	2	5.88	0	0.00	0	0.00	34
advanced	0	True	0	0.00	3	12.50	0	0.00	0	0.00	24
advanced	0	nan	0	0.00	0	0.00	0	0.00	0	0.00	3
advanced	1	False	0	0.00	3	4.55	0	0.00	0	0.00	66
advanced	1	True	1	1.64	18	29.51	5	8.20	1	1.64	61
advanced	1	nan	0	0.00	0	0.00	0	0.00	0	0.00	9
advanced	2	False	4	5.26	16	21.05	7	9.21	4	5.26	76
advanced	2	True	3	3.80	46	58.23	18	22.78	8	10.13	79
advanced	2	nan	0	0.00	0	0.00	0	0.00	0	0.00	8

Source: Article Notebook

Prevalence of Midline Extension

In figure 11, we plot the prevalence of midline extension in the data versus our model's prediction. It is obvious the model cannot match the large spread between early and advanced T-category seen in the data. This is because to achieve that, it would need to increase the advanced T-category patient's prior distribution over diagnosis times and at the same time reduce the probability of the tumor to cross the midline during
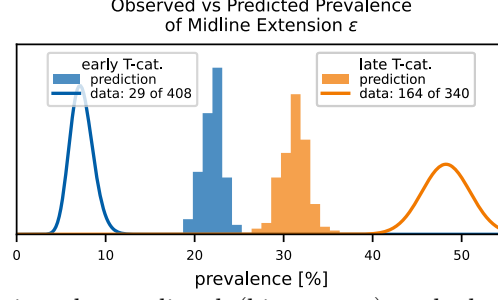
Figure 11: Comparing the predicted (histograms) and observed (lines depicting beta posteriors) prevalence of midline extension for early (blue) and late (orange) T-category. While the prevalence is predicted correctly when marginalizing over T-category, the model cannot capture the degree of separation observed in the data. Since the tumor's midline extension is virtually always part of the diagnosis and hence *given* when predicting a patient's risk, we do not consider this discrepancy a major issue.

a time-step. But since the time-priors parameter is also coupled with the spread probabilities among the LNLs, the model does not have that freedom.

However, we do not consider this discrepancy a major limitation of the model: We will not realistically be interested in the probability of midline extension, as it is always possible to assess it with high certainty. That is also the reason why we initially modelled the midline extension *not* as a random variable, but as a global risk factor that would have been turned on or off from the onset of a patient's disease evolution. This, however, lead to overly high risks for contralateral involvement in advanced T-category patients with midline extension, because then the model assumes an increased spread to the contralateral side from the onset of the disease. Which is probably not true in a majority of those cases. Thus, treating it as a random variable that only becomes true during a patient's disease evolution resulted in a better description of the data.

Formally, the wrong prediction of midline extension prevalence makes little difference, since it is always given: Instead of $P\left(\mathbf{X}^{\mathrm{i}}, \mathbf{X}^{\mathrm{c}}, \epsilon \mid \mathbf{Z}^{\mathrm{i}}, \mathbf{Z}^{\mathrm{c}}\right)$, we typically compute $P\left(\mathbf{X}^{\mathrm{i}}, \mathbf{X}^{\mathrm{c}} \mid \mathbf{Z}^{\mathrm{i}}, \mathbf{Z}^{\mathrm{c}}, \epsilon\right)$, which does not suffer from the wrong probability of midline extension, as the distribution over hidden states is renormalized:

$$P\left(\mathbf{X}^{\mathrm{i}}, \mathbf{X}^{\mathrm{c}} \mid \mathbf{Z}^{\mathrm{i}}, \mathbf{Z}^{\mathrm{c}}, \epsilon\right) = \frac{P\left(\mathbf{Z}^{\mathrm{i}}, \mathbf{Z}^{\mathrm{c}} \mid \mathbf{X}^{\mathrm{i}}, \mathbf{X}^{\mathrm{c}}, \epsilon\right) P\left(\mathbf{X}^{\mathrm{i}}, \mathbf{X}^{\mathrm{c}}, \epsilon\right)}{P\left(\mathbf{Z}^{\mathrm{i}}, \mathbf{Z}^{\mathrm{c}}, \epsilon\right)}$$

Note that a distribution over ϵ appears both in the enumerator and the denominator, which largely cancel each other, leaving only the midline extension's effect on the distribution over hidden states in the prediction.

Also, the discrepancy in midline extension prevalence between early and advanced T-category is particularly pronounced in oropharyngeal SCC patients. For example, in

29

oral cavity SCC, the midline extension only increases from 15.4% (20 out of 130) to 33.3% (13 out of 39).

Source: Article Notebook

References

[1] Grégoire, V., Levendag, P., Ang, K.K., Bernier, J., Braaksma, M., Budach, V., Chao, C., Coche, E., Cooper, J.S., Cosnard, G., Eisbruch, A., El-Sayed, S., Emami, B., Grau, C., Hamoir, M., Lee, N., Maingon, P., Muller, K., Reychler, H.: CT-based delineation of lymph node levels and related CTVs in the node-negative neck: DAHANCA, EORTC, GORTEC, NCIC,RTOG consensus guidelines. Radiotherapy and Oncology **69**(3), 227–236 (2003) https://doi.org/10.1016/j.radonc.2003.09.011

[2] Grégoire, V., Ang, K., Budach, W., Grau, C., Hamoir, M., Langendijk, J.A., Lee, A., Le, Q.-T., Maingon, P., Nutting, C., O'Sullivan, B., Porceddu, S.V., Lengele, B.: Delineation of the neck node levels for head and neck tumors: A 2013 update. DAHANCA, EORTC, HKNPCSG, NCIC CTG, NCRI, RTOG, TROG consensus guidelines. Radiotherapy and Oncology **110**(1), 172–181 (2014) https://doi.org/10.1016/j.radonc.2013.10.010

[3] Grégoire, V., Evans, M., Le, Q.-T., Bourhis, J., Budach, V., Chen, A., Eisbruch, A., Feng, M., Giralt, J., Gupta, T., Hamoir, M., Helito, J.K., Hu, C., Hunter, K., Johansen, J., Kaanders, J., Laskar, S.G., Lee, A., Maingon, P., Mäkitie, A., Micciche', F., Nicolai, P., O'Sullivan, B., Poitevin, A., Porceddu, S., Składowski, K., Tribius, S., Waldron, J., Wee, J., Yao, M., Yom, S.S., Zimmermann, F., Grau, C.: Delineation of the primary tumour clinical target volumes (CTV-P) in laryngeal, hypopharyngeal, oropharyngeal and oral cavity squamous cell carcinoma: AIRO, CACA, DAHANCA, EORTC, GEORCC, GORTEC, HKNPCSG, HNCIG, IAG-KHT, LPRHHT, NCIC CTG, NCRI, NRG Oncology, PHNS, SBRT, SOMERA, SRO, SSHNO, TROG consensus guidelines. Radiotherapy and Oncology **126**(1), 3–24 (2018) https://doi.org/10.1016/j.radonc.2017.10.016

[4] Eisbruch, A., Foote, R.L., O'Sullivan, B., Beitler, J.J., Vikram, B.: Intensity-modulated radiation therapy for head and neck cancer: Emphasis on the selection and delineation of the targets. Seminars in Radiation Oncology **12**(3), 238–249 (2002) https://doi.org/10.1053/srao.2002.32435

[5] Biau, J., Lapeyre, M., Troussier, I., Budach, W., Giralt, J., Grau, C., Kazmierska, J., Langendijk, J.A., Ozsahin, M., O'Sullivan, B., Bourhis, J., Grégoire, V.: Selection of lymph node target volumes for definitive head and neck radiation therapy: A 2019 Update. Radiotherapy and Oncology **134**, 1–9 (2019) https://doi.org/10.1016/j.radonc.2019.01.018

[6] Chao, K.S.C., Wippold, F.J., Ozyigit, G., Tran, B.N., Dempsey, J.F.: Determination and delineation of nodal target volumes for head-and-neck cancer based

on patterns of failure in patients receiving definitive and postoperative IMRT. International Journal of Radiation Oncology*Biology*Physics **53**(5), 1174–1184 (2002) https://doi.org/10.1016/S0360-3016(02)02881-X

[7] Vorwerk, H., Hess, C.F.: Guidelines for delineation of lymphatic clinical target volumes for high conformal radiotherapy: Head and neck region. Radiation Oncology **6**(1), 97 (2011) https://doi.org/10.1186/1748-717X-6-97

[8] Ferlito, A., Silver, C.E., Rinaldo, A.: Elective management of the neck in oral cavity squamous carcinoma: Current concepts supported by prospective studies. British Journal of Oral and Maxillofacial Surgery **47**(1), 5–9 (2009) https://doi.org/10.1016/j.bjoms.2008.06.001

[9] Ludwig, R., Pouymayou, B., Balermpas, P., Unkelbach, J.: A hidden Markov model for lymphatic tumor progression in the head and neck. Scientific Reports **11**(1), 12261 (2021) https://doi.org/10.1038/s41598-021-91544-1

[10] Ludwig, R., Schubert, A., Barbatei, D., Bauwens, L., Hoffmann, J.-M., Werlen, S., Elicin, O., Dettmer, M., Zrounba, P., Pouymayou, B., Balermpas, P., Grégoire, V., Giger, R., Unkelbach, J.: Modelling the Lymphatic Metastatic Progression Pathways of OPSCC from Multi-Institutional Datasets. arXiv (2023). https://doi.org/10.48550/arXiv.2312.11270

[11] Pouymayou, B., Balermpas, P., Riesterer, O., Guckenberger, M., Unkelbach, J.: A Bayesian network model of lymphatic tumor progression for personalized elective CTV definition in head and neck cancers. Physics in Medicine & Biology **64**(16), 165003 (2019) https://doi.org/10.1088/1361-6560/ab2a18

[12] Ludwig, R., Hoffmann, J.-M., Pouymayou, B., Morand, G., Däppen, M.B., Guckenberger, M., Grégoire, V., Balermpas, P., Unkelbach, J.: A dataset on patient-individual lymph node involvement in oropharyngeal squamous cell carcinoma. Data in Brief **43**, 108345 (2022) https://doi.org/10.1016/j.dib.2022.108345

[13] Ludwig, R., Schubert, A., Barbatei, D., Bauwens, L., Werlen, S., Elicin, O., Dettmer, M., Zrounba, P., Balermpas, P., Pouymayou, B., Grégoire, V., Giger, R., Unkelbach, J.: A multi-centric dataset on patient-individual pathological lymph node involvement in head and neck squamous cell carcinoma. Data in Brief **52**, 110020 (2024) https://doi.org/10.1016/j.dib.2023.110020

[14] De Bondt, R.B.J., Nelemans, P.J., Hofman, P.A.M., Casselman, J.W., Kremer, B., Van Engelshoven, J.M.A., Beets-Tan, R.G.H.: Detection of lymph node metastases in head and neck cancer: A meta-analysis comparing US, USgFNAC, CT and MR imaging. European Journal of Radiology **64**(2), 266–272 (2007) https://doi.org/10.1016/j.ejrad.2007.02.037

[15] Kyzas, P.A., Evangelou, E., Denaxa-Kyza, D., Ioannidis, J.P.A.: 18F-fluorodeoxyglucose positron emission tomography to evaluate cervical node metastases in patients with head and neck squamous cell carcinoma: A meta-analysis. JNCI: Journal of the National Cancer Institute **100**(10), 712–720 (2008) https://doi.org/10.1093/jnci/djn125

[16] Ludwig, R., Schubert, A.D., Barbatei, D., Bauwens, L., Hoffmann, J.-M., Werlen, S., Elicin, O., Dettmer, M., Zrounba, P., Pouymayou, B., Balermpas, P., Grégoire, V., Giger, R., Unkelbach, J.: Modelling the lymphatic metastatic progression pathways of OPSCC from multi-institutional datasets. Scientific Reports **14**(1), 15750 (2024) https://doi.org/10.1038/s41598-024-66012-1

[17] Foreman-Mackey, D., Hogg, D.W., Lang, D., Goodman, J.: Emcee: The MCMC Hammer. Publications of the Astronomical Society of the Pacific **125**(925), 306–312 (2013) https://doi.org/10.1086/670067

[18] Ter Braak, C.J.F., Vrugt, J.A.: Differential Evolution Markov Chain with snooker updater and fewer chains. Statistics and Computing **18**(4), 435–446 (2008) https://doi.org/10.1007/s11222-008-9104-9

[19] Nelson, B., Ford, E.B., Payne, M.J.: RUN DMC: AN EFFICIENT, PARALLEL CODE FOR ANALYZING RADIAL VELOCITY OBSERVATIONS USING N-BODY INTEGRATIONS AND DIFFERENTIAL EVOLUTION MARKOV CHAIN MONTE CARLO. The Astrophysical Journal Supplement Series **210**(1), 11 (2013) https://doi.org/10.1088/0067-0049/210/1/11

[20] Benson, N., Whipple, M., Kalet, I.J.: A Markov model approach to predicting regional tumor spread in the lymphatic system of the head and neck. AMIA ... Annual Symposium proceedings. AMIA Symposium, 31–35 (2006)

[21] Jung, H., Law, A., Grunblatt, E., Wang, L.L., Kusano, A., Mejino, J.L.V., Whipple, M.E.: Development of a Novel Markov Chain Model for the Prediction of Head and Neck Squamous Cell Carcinoma Dissemination. AMIA Annual Symposium Proceedings **2016**, 1832–1839 (2017)

[22] University of Zurich: Personalized Volume-deescalated Elective Nodal Irradiation in Oropharyngeal Head and Neck. Clinical Trial Registration NCT06563362, clinicaltrials.gov (August 2024)

[23] Sarrut, D., Rit, S.: Proceedings of the XX-th International Conference on the Use of Computers in Radiation Therapy (ICCR), Lyon, France (2024)

[24] Ludwig, R.: Modelling Lymphatic Metastatic Progression in Head and Neck Cancer. PhD thesis, University of Zurich, Zurich (2023). https://doi.org/10.5167/uzh-231470