Modelling lymph node involvement in early stage breast cancer

Background

The treatment landscape for patients with early-stage breast cancer has evolved significantly over the last few decades, aiming for optimal outcomes while minimizing treatment-related harm [1]. Advances in targeted systemic therapies and technological innovations in surgery and radiotherapy have markedly improved patient outcomes [2]. Specifically, in several meta-analyses, radiotherapy to the locoregional lymph node levels after surgery has demonstrated a reduction in disease recurrences and breast cancer mortality, especially in early-stage breast cancer patients with positive lymph nodes [3].

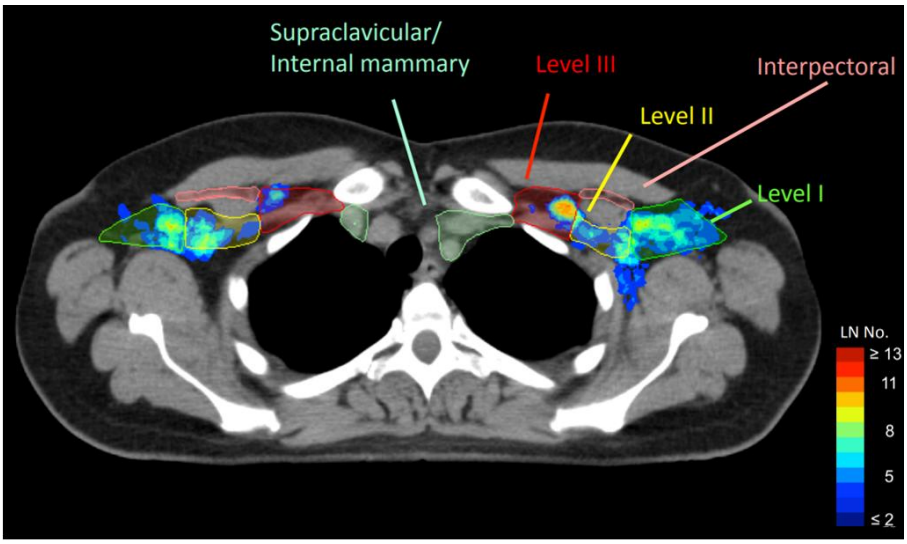

Figure 1: Example of a patient CT with the different locoregional lymph node levels delineated.

Despite these benefits, patients may encounter side effects during radiotherapy, such as skin irritation, fatigue, and breast edema. Potential long-term complications include changes to the breast, arm and shoulder pain, and moderate to severe fibrosis, which can affect the quality of life substantially [4,5]. Less frequently, but potentially more serious long-term complications include heart and lung diseases and the development of secondary cancers [6]. The probability and severity of these side effects are often directly correlated to the volume of irradiated tissue, emphasizing the delicate balance in defining appropriate radiotherapy target volumes.

The objective of this research project is to optimize the target volumes for locoregional radiotherapy in patients with early-stage breast cancer, through big data analysis and modelling of lymph node involvement using Hidden Markov Models.

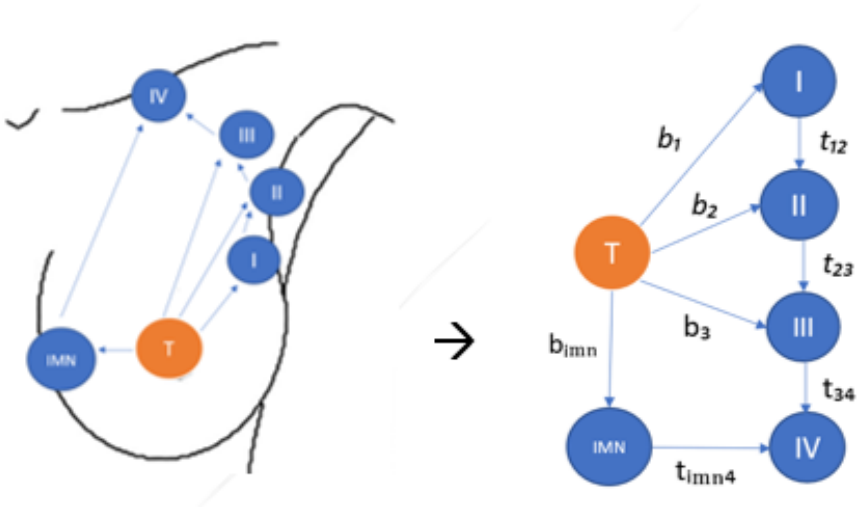

Figure 2: Lymphatic drainage pathways in the breast and corresponding schematic of the Hidden Markov Model with base- (b) and transition- (t) probabilities.

Specific Tasks

- Perform a literature study to get familiar with the different topics.
- Perform descriptive and statistical data analysis on a large (N>500), retrospective cohort of early-stage breast cancer patients, similar to [7]. The data will include patient and disease characteristics (clinical data), treatment-related data and lymph node involvement obtained from different modalities (imaging and pathology). Data analysis will be carried out in python using pandas library, and libraries for data visualization.
- Get familiar with an already implemented Hidden Markov Model (in Python), including the dataset used for fitting.
- Adapt the Hidden Markov Model to include patient- or disease-specific parameters, such as primary tumour location.
- Evaluation of new model performance after adaptation and refitting the model.

Contact

- Dr. Ir. Jennifer Dhont (jennifer.dhont@hubruxelles.be), Head of Data Science & AI Research Unit at Hopital Universitaire de Bruxelles (Erasme campus)

References

1. Waks, A. G., et al. Breast Cancer Treatment: A Review. JAMA 321, 288–300 (2019).
2. Bray, F., et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 68, 394–424 (2018).

3. Abe, O., et al. Effects of radiotherapy and of differences in the extent of surgery for early breast cancer on local recurrence and 15-year survival. Lancet 366, 2087–2106 (2005).
4. Hopwood, P., et al. Comparison of patient-reported breast, arm, and shoulder symptoms and body image after radiotherapy for early breast cancer. Lancet Oncol 11, 231–240 (2010).
5. Lyngholm, C. D., et al. Long-term follow-up of late morbidity, cosmetic outcome and body image after breast conserving therapy. Acta Oncol (Madr) 52, 259–269 (2013).
6. Piroth, M. D., et al. Heart toxicity from breast cancer radiotherapy: Current findings, assessment, and prevention. Strahlentherapie Und Onkologie 195, 1 (2019).
7. Ludwig, R., et al. Detailed patient-individual reporting of lymph node involvement in oropharyngeal squamous cell carcinoma with an online interface. Radiother Oncol 169, 1-7 (2022).