Lymphoma is a heterogeneous disease on several levels, which can be present in any of the over 500 lymph nodes in the body, or other lymphatic organs such as the bone marrow and spleen. Patients with lymphoma are frequently assessed using whole-body FDG PET/CT imaging for staging and treatment response assessment, which is largely performed visually by physicians. However, by extracting quantitative features from PET/CT images, we can access more in-depth information about the disease, improving prognostic power and risk stratification.

In diseases characterized by a high disease burden such as lymphoma, the full potential of quantitative PET/CT imaging metrics is not achieved due to the difficulty of identifying and assessing all disease. A manual workflow can take up to 30-45 minutes per patient for difficult cases and is prone to subjectivity. In addition, even with a quantitative assessment, studies are often limited by only considering a patient-level assessment rather than considering characteristics of individual lesions. Thus, the overall goal of this research was to unlock the potential of quantitative imaging metrics for improving the assessment of patients with lymphoma through robust, automated image analysis techniques.

To achieve this, a fully automated, end-to-end assessment workflow was developed with three phases: detection, segmentation, and quantification. The detection method implemented uses multi-resolution pathway 3D CNNs (DeepMedic), which were trained to create probability maps of where diseased lymph nodes are located by learning contours created by physicians. By varying the threshold of these probability maps, detection performance can be varied for increased sensitivity or specificity. Results for the detection phase were promising, with performance similar to the differences between two physicians performing the same task.

For the segmentation phase, 11 automated segmentation methods were assessed for their ability to consistently and accurately contour lymphoma lesions on PET. Wide variation in performance was found for all algorithms as well as between physicians, indicating the difficulty

of lymphoma lesion segmentation. DeepMedic was among the top performing segmentation algorithms, and was the only algorithm robust to the initial algorithm input.

Based on these findings, the end-to-end workflow implemented for quantification used DeepMedic for both detection and segmentation. Overall quantification performance showed excellent agreement with physician patient-level quantitation for both adult and pediatric lymphoma patient applications, with relatively few outliers present. The extraction of SUV_{max} was highly sensitive to detection specificity, while SUV_{mean}, SUV_{total}, and MTV were less sensitive as the majority of false positives were small and relatively low uptake.

In addition to patient-level disease assessment on baseline images, a technique was developed to track lesions from baseline to follow-up images. This allowed for the calculation of a lesion-level PET imaging metric that has not been assessed in lymphoma: lesion response heterogeneity. Lesions were classified according to ±30% for the definition of progressing and partially responding disease, with results showing that the majority of lesions show complete or partial response to therapy.

The work presented in this thesis presents many opportunities in which automation and quantification can improve the assessment of lymphoma patients. By shifting from a visual and qualitative description to an automated and quantitative approach, assessment at baseline and follow-up can be standardized and easily adopted into a clinical setting. In addition, the extraction of both patient-level and individual lesion-level quantitative metrics allows for a deeper understanding of disease characteristics and how treatments are working. These improvements, among many other opportunities, can therefore facilitate the implementation of personalized treatment decisions for lymphoma patients.