Chapter 9

Summary and discussion

Summary and discussion

Positron Emission Tomography (PET) is a relatively new imaging technique that allows the visualization of biochemical processes in tissue. PET has the ability to complement the traditional imaging modes such as computed tomography (CT) which provide information only on anatomical structures. Especially in malignant lymphoma, the differentiation between active disease and scar tissue is of utmost importance in response monitoring. Although the use of PET in the management of lymphoma seems to be established already, much about the true potentials and drawbacks of FDG-PET in this disease are still unknown. The PET studies described in this thesis may help to improve the interpretation of PET results in patients with untreated or relapsed lymphoma and help to design future PET guided clinical studies.

In the **Introduction**, the technical aspects of PET are described and an update of the role of FDG-PET in malignant lymphoma is provided as a starting point for our studies. During the last years, clinical risk scores, e.g., the International prognostic Index (IPI)[1] and biological characteristics including gene profile patterns, have been set up to identify patient groups with different prognostic features. In addition, there is an ongoing search for more accurate prognostic factors in the management of malignant lymphoma and midtreatment PET may be one of them. Combination of both midtreatment PET and clinical characteristics may further improve the prognostic classification.

When, in 1996, a dedicated PET scanner was planned in the VU University Medical Center (VUMC), the potential additional value of PET in staging of Hodgkin's lymphoma was discussed. Extended subtotal node irradiation (STNI) had been the treatment of choice for Hodgkin's lymphoma patients in clinical stages I-II with lymphadenopathy above the diaphragm, for many years until the mid '90. However, up to one third of all patients relapsed outside the radiation field, indicating that there must have been occult disease not properly detected by conventional staging, including CT scanning. **Chapter 2** describes a retrospective evaluation of the cohort of 106 consecutive patients treated with STNI at the VUMC from 1975 until 1995, for relapse free survival and the incidence of second cancers. With a median follow up of more than eleven years, the relapse-free interval was 78% at 5 years and 72% at 10 years. In 60%, the relapse was located outside the radiation field, indicating the insensitivity of conventional staging.

In **Chapter 3** the results of a systematic review are described of the literature on the diagnostic performance of FDG-PET in the evaluation of first-line therapy of Hodgkin's

and aggressive non-Hodgkin's lymphoma including its sensitivity and specificity. Eligibility criteria included the use of FDG as a tracer, a dedicated PET camera and sufficient data to allow calculations of sensitivity and specificity. Fifteen studies involving 705 patients met the inclusion criteria. Even though we identified several methodological deficiencies, the study results consistently show that FDG-PET has a very high specificity in this setting with a pooled sensitivity of 72% and 84% for aggressive NHL and HL, respectively. Using the obtained point estimates of sensitivity and specificity, we calculated the predicted post-PET probabilities of viable lymphoma as a function of the pre-test probability (prevalence). These data suggest that in the case of a 50% probability of persisting viable tumor after first-line therapy for aggressive NHL, the probability of having persistent viable tumor after a positive scan is 97%, versus 22% in the case of a negative PET result. Applying a 15% pre-test probability of relapse in Hodgkin's lymphoma, the projected probability of relapse in the case of a negative PET result is 3%, versus 60% in the case of a positive PET. The reviewed studies investigated diagnostic accuracy and not the impact of PET results on management or on patients' outcome resulting from management changes. Nowadays, several studies on Hodgkin's and non-Hodgkin's lymphoma are PET tailored and it is very fascinating to investigate whether PET based therapy can influence patient outcome positively.

In current literature, PET interpretation and assessment is typically visual, and therefore subject to observer variation. In most studies, the aspect of reproducibility is not assessed. In **Chapter 4** we describe the results of an observer variation study in staging and therapy evaluation in malignant lymphoma. Eleven nuclear medicine physicians with different levels of PET experience independently reviewed 37 PET scans, 10 obtained at presentation and 27 during or after therapy. They were requested to identify and localize suspicious lymphoma sites, to assign a stage to the baseline scans, and to interpret the remaining scans for the presence of viable lymphoma. Individual (extra-)nodal regions were assessed for the likelihood of malignancy as positive, negative or equivocal, and these results were compared with expert readings as the gold standard. In a sensitivity analysis, we compared the results of conservative (i.e. equivocal scores allocated by the observers considered to be benign) and sensitive reading (assigning equivocal scores to malignant categories), respectively. In the assessment of baseline scans, 61% and 56% of the scans were scored in accordance with the experts, using sensitive and conservative reading respectively. Errors occurred in either direction. We found a trend towards overstaging by the more experienced readers and towards understaging for the less experienced ones. The experts classified 14 out of 27 scans obtained during or after therapy as positive for viable tumor, and 82% versus 94% were scored in accordance with the experts, using sensitive and conservative reading,

respectively. The 13 negative scans were scored in agreement with the experts in only 45% of the cases. False positivity pertained especially to the neck, periclavicular and axillary lymph nodes, mediastinum, lung and bone marrow. More experienced observers tended to have fewer false negative scores. So there were substantial disparities among nuclear medicine physicians' interpretation of FDG PET scans of lymphoma patients, which may affect patient care and results of multi-center clinical trials. A standardized method of reporting is urgently needed to improve consistency. Moreover, PET-CT technology might also improve the specificity of interpretations.

The International Prognostic Index (IPI)[1;2] and the International prognostic Score (IPS)[3] are currently used clinical prognostic indices for diffuse large B-cell lymphoma and Hodgkin's lymphoma, respectively. However, these prognostic models use static pre-treatment characteristics to predict likelihood of response and survival at diagnosis. During and directly after therapy, the conventional restaging using radiological methods like CT scanning do not accurately predict patient outcome, mainly because of incomplete or delayed shrinkage of malignant tissues. ^{67}Galliumscintigraphy is a metabolic study that relies on the accumulation of the isotope into viable lymphoma cells via binding to transferrin receptors. ^{67}Ga is useful in assessing response to lymphoma treatment, improving the specificity of CT. The value of early restaging with ^{67}Ga was demonstrated in a selected group of 30 patients with aggressive NHL and poor prognosis treated with high dose chemotherapy [4]. At the VUMC, we had a huge experience with ^{67}Ga scintigraphy for response monitoring, and with the introduction of the PET scanner in our hospital in 1997, we designed a prospective study comparing FDG PET and ^{67}Ga after 2 cycles of CHOP chemotherapy, to assess which nuclear imaging technique had the best characteristics for response monitoring in aggressive NHL **(Chapter 5)**. FDG PET and ^{67}Ga images were analysed visually by four experienced nuclear medicine physicians, who were blinded for the alternative scan technique and follow up. Twenty-six patients were included and 11 remained free from progression with a median follow up of more than 2 years, whereas 14 patients relapsed and one died of lung cancer. Time to progression was associated with PET positivity, but not with ^{67}Ga positivity. Sixty-four percent of patients with negative early restaging PET remained in clinical remission, whereas only 25% of PET positive patients remained disease free. ^{67}Ga appeared to have no predictive value for treatment outcome. Obviously, the sample size was too small to provide accurate estimates of predictive values.

In **Chapter 6**, the concept of assessment of chemo-sensitivity by interim FDG-PET was tested in a multi-center prospective cohort study in newly diagnosed patients with aggressive NHL. The PALET-study (**p**rognosis of **a**ggressive **l**ymphoma using **e**mission **t**omography) is a

co-operation between department of haematology and PET centers from Groningen (UMCG), Nijmegen (UMCN), Maastricht (AZM) and Amsterdam (VUMC). Patients with untreated aggressive NHL, planned to be treated with CHOP-like chemotherapy were eligible. One blinded FDG-PET scan was performed just before the fourth cycle of chemotherapy. One hundred and fourteen patients, of which 92 were evaluable, were included in the analysis. Of these patients, 36 showed disease progression after a median follow up of 34 months. 63% of the PET positive patients and 29% of the PET negative patients progressed. Two-year progression free survival (PFS) was 71% for the PET-negative patients versus 42% for PET positive patients. PET was the strongest independent prognostic factor in a multivariate regression analysis. Using a prognostic model incorporating interim PET, interim CT and LDH at diagnosis, 4 risk groups with a progression free survival varying from 88% to 30% could be identified. These results indicate that interim PET in addition to clinical risk factors may be used for a more accurate treatment stratification.

Patients with a relapse of malignant lymphoma are treated with second line chemotherapy, currently consisting of three cycles of reinduction chemotherapy (DHAP-VIM-DHAP) plus Rituximab followed by stem cell mobilisation, myeloablative chemotherapy (BEAM) and autologous stem cell reinfusion (ASCT). Since the likelihood of response to chemotherapy is determined by a number of factors, we questioned whether the combined use of clinical parameters and PET response might be of value.

Chapter 7 gives a description of 101 relapsed lymphoma patients (78 NHL and 23 HL) from the haematology centers of UMCG and VUMC. These patients were treated according to the above mentioned protocol, resulting in transplant for 77% of the NHL and 87% of the HL patients. A FDG-PET scan was scheduled just before treatment and after the second course of induction chemotherapy (DHAP-VIM). Of NHL patients, 29% had non-responsive PET scans; these patients had a 2-year PFS of only 10%. Of HL patients, 21% had non-responsive PET-scans and a 2-year PFS of 40%.

In a multivariate regression analysis, both clinical risk score and PET response after DHAP-VIM appeared to be independent predictive factors for PFS. The combined use of these factors led to a model with four different risk groups with a PFS varying between 17% for patients with insufficient PET response and unfavourable clinical risk factors and 95% for patients with a complete PET response and favourable clinical risk factors ($p<.001$). These results applied to relapsed NHL and the combined group of lymphoma patients, but not to Hodgkin's patients alone, mainly due to the limited number of patients. So in relapsed lymphoma, the prognostic value of midtreatment PET is defined more clearly, especially if used in conjunction with clinical risk factors. In relapsed lymphoma, a favourable prognosis

group and a poor prognosis group can be identified, which may have implications for the design of future clinical trials.

During the last few years, the use of FDG-PET for evaluation of malignant lymphoma has increased dramatically. The widely used International Working Group criteria for response assessment of lymphoma, published in 1999, are based predominantly on CT scanning and do not include PET as part of response assessment[5]. Considering the more recent widespread use of PET, it became clear that the International Working Group criteria warranted revision. The Competence Network Malignant Lymphoma convened an International Harmonization Project and charged the Imaging subcommittee to develop guidelines for performing and interpreting FDG-PET for treatment assessment in lymphoma, to ensure the reliability of the method both in the context of clinical trials and in clinical practice. **Chapter 8** describes these consensus recommendations based on published PET literature and the collective expertise of the members of the Imaging subcommittee. The value of PET for response assessment at the conclusion of therapy in Hodgkin's lymphoma (HL) and diffuse large B-cell lymphoma (DLBCL) is t clearest. Its ability to distinguish between viable tumor and necrosis or fibrosis in residual masses often present after treatment is of great value, because conventional imaging modalities are unable to make this distinction. PET after completion of therapy should be performed at least 3 weeks, and preferably 6 to 8 weeks, after chemotherapy or chemo-immunotherapy, and 8 to 12 weeks after radiation or chemo-radiotherapy. A pre-therapy PET is recommended for HL, DLBCL, follicular lymphoma and mantle-cell lymphoma, because it can facilitate the interpretation of the post-therapy PET. For variably FDG-avid lymphomas, a pre-therapy PET is mandatory if PET is used to assess their response to treatment. For the interpretation of PET scans after therapy, visual assessment alone appears to be adequate for determining whether PET is positive or negative. The generally used definition of a positive PET finding as focal or diffuse FDG uptake above background in a location incompatible with normal anatomy/physiology seems to be appropriate, with mediastinal blood pool activity as the reference background activity for residual mass $\geq$ 2 cm in greatest transverse diameter. It is anticipated that these guidelines will promote continuing dialogue between haematologists and imaging physicians to achieve the greatest benefit from PET imaging in patients with lymphoma.

Discussion and future perspective

This thesis addresses the use of FDG-PET in response monitoring at the conclusion of therapy as well as early during therapy, to assess chemosensitivity in primary aggressive NHL and relapsed aggressive lymphoma. In newly diagnosed HL, interim FDG-PET has proven to be a valuable tool to predict outcome. Recent publications have shown that FDG-PET after two cycles of chemotherapy has a strong correlation with outcome [6;7]. These results have led to clinical studies such as the EORTC H10 study and the GHSG HD18 study in which treatment decisions are guided by the results of midtreatment PET. To avoid an indiscriminate overtreatment for a substantial fraction of the patients, a risk-adapted therapy tailored to the individual patient is now the subject of new trials. The new EORTC/GELA Intergroup H10 trial is designed for patients with early stage HL. This Phase III trial aims at reducing treatment related toxicity in selected groups of patients who are "overtreated" with the current combined modality treatment. It is hypothesized that early response to treatment can lead to a decrease in the burden of treatment, e.g. omitting radiotherapy. Patients with a negative FDG-PET after 2 cycles of ABVD chemotherapy constitute the favourable risk group, independent of classical prognostic factors. The primary objective for this favourable risk group is to evaluate whether chemotherapy alone is as effective as combined modality therapy, thus potentially reducing toxicity. The group of patients with PET-positivity after two cycles of chemotherapy is considered to be the unfavourable risk group. The objective for this group is to evaluate whether treatment intensification with escalated BEACOPP will improve the outcome when compared with standard therapy. Notably, in the standard arm FDG-PET scan after 2 cycles of ABVD is mandatory as well, but the results will be used for documentation purposes only. This is the first randomised intergroup trial on early treatment adaptation guided by the results of FDG-PET scanning. The German Hodgkin Study Group will soon start the HD18 study for advanced stage HL, with PET guided therapy modifications after two cycles of escalated BEACOPP. So for HL patients, the important question as to whether PET guided treatment modifications will improve outcome and lower treatment related toxicity is likely to be answered in the next decade.

During the last few years, treatment results of aggressive B-cell NHL patients have improved dramatically by the introduction of Rituximab in combination with chemotherapy [8;9]. But as a result of the improved first-line therapy, relapses after chemo-immunotherapy are more difficult to cure with second-line chemotherapy and autologous stem cell transplantation [10;11]. In parallel to HL, mid-treatment PET could also play an important role in treatment stratification in future studies for aggressive NHL. However, aggressive NHL consists of a

far more heterogeneous group of patients than HL. New trials will address the efficacy of new therapeutics like monoclonal antibodies and other biological immunomodifiers. These agents could directly interfere with the kinetics and biodistribution of FDG. In this setting, FDG-PET assessments may have recognized limitations. These include unreliable evaluation of bone marrow status, difficulty in distinguishing benign post-therapeutic and intercurrent inflammatory responses, inability to consistently and reliably characterize lesions at the subcentimeter level, and last but not least, the variability in interpretation and analysis between readers, and the variability in signal acquisition and image reconstruction between instrument platforms. Whether semi-quantitative assessment using the Standardized Uptake Value (SUV) of the affected area will improve the visual PET assessment has to be determined in future prospective studies following standardization of the PET acquisition. In case of multicenter studies, phantom studies have to be performed to determine intercenter correction factors [12]. Recently, the HOVON Imaging Group has determined a standard to be used for future studies of the Dutch-Belgium Haemato-oncology Cooperative Group (HOVON).

Since the advent of the PET scanner, improvements in the concept of the scanner and other technologies have seen the light. After a decennium of whole body PET, the integrated or hybrid PET-CT scanner has been introduced. Several studies have claimed that that the hardware fused anatomical (CT) and functional (PET) images have superior accuracy compared to visual fusion (side-by-side reading) [13;14]. Nowadays, many centers have access to a combined PET-CT scanner. With the increasing number of PET-CT facilities, the key challenge is to implement PET-CT in multicenter studies. FDG-PET exploits the reliance of tumor cells on glycolytic metabolism to image lymphoma; therefore, there is a strong mechanistic rationale for the adoption of this endpoint as a meaningful surrogate for tumor response to therapy. The development of standardized FDG-PET image acquisition and analysis procedures will certainly provide better, earlier and more reliable clinical assessments of when to switch to second-line therapies in the face of ineffective treatment. Other radiopharmaceuticals like F-18-fluoro-3'-deoxy-L-fluorothymidine (FLT) have been developed recently. Uptake of this tracer is directly related to the rate of DNA synthesis and could reflect cell proliferation. Today, the role of FLT for therapy monitoring has not yet been evaluated in detail.
For multicenter studies, the calibration of PET and PET-CT systems is of the utmost importance, if there is a need to quantify the intensity of tracer uptake. In the very near future, a validated PET-CT procedure holds the potential to provide early efficacy evaluation of many novel anticancer agents, and to shorten the overall duration of Phase I and Phase II trials. Validated imaging endpoint information should enable an evidence-based regulatory review that will

favour the accelerated approval of new therapeutic strategies for chemo-resistant lymphoma patients.

It remains the challenge for future trials to exploit validated PET techniques in a multicenter setting for monitoring response in the progress of new lymphoma trials. Close cooperation between haematologists and nuclear medicine physicians in study groups like HOVON is vital to meet these challenges.

Reference List

1. The International Non-Hodgkin's Lymphoma Prognostic Factors Project. A predictive model for aggressive non-Hodgkin's lymphoma. N.Engl.J Med. 1993;329:987-994.
2. Shipp MA. Prognostic factors in aggressive non-Hodgkin's lymphoma: who has "high- risk" disease? Blood 1994;83:1165-1173.
3. Hasenclever D, Diehl V. A prognostic score for advanced Hodgkin's disease. International Prognostic Factors Project on Advanced Hodgkin's Disease. N.Engl.J.Med. 1998;339:1506-1514.
4. Janicek M, Kaplan W, Neuberg D et al. Early restaging gallium scans predict outcome in poor-prognosis patients with aggressive non-Hodgkin's lymphoma treated with high-dose CHOP chemotherapy. J.Clin.Oncol. 1997;15:1631-1637.
5. Cheson BD, Horning SJ, Coiffier B et al. Report of an international workshop to standardize response criteria for non-Hodgkin's lymphomas. NCI Sponsored International Working Group. J.Clin.Oncol. 1999;17:1244.
6. Hutchings M, Loft A, Hansen M et al. FDG-PET after two cycles of chemotherapy predicts treatment failure and progression-free survival in Hodgkin lymphoma. Blood 2006;107:52-59.
7. Gallamini A, Hutchings M, Rigacci L et al. Early interim 2-[18F]fluoro-2-deoxy-D-glucose positron emission tomography is prognostically superior to international prognostic score in advanced-stage Hodgkin's lymphoma: a report from a joint Italian-Danish study. J.Clin.Oncol. 2007;25:3746-3752.
8. Coiffier B, Lepage E, Briere J et al. CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. N.Engl.J.Med. 2002;346:235-242.
9. Sehn LH, Donaldson J, Chhanabhai M et al. Introduction of combined CHOP plus rituximab therapy dramatically improved outcome of diffuse large B-cell lymphoma in British Columbia. J.Clin.Oncol. 2005;23:5027-5033.
10. Vellenga E, van Putten WL, van ', V et al. Rituximab improves the treatment results of DHAP-VIM-DHAP and ASCT in relapsed/progressive aggressive CD20+ NHL: a prospective randomized HOVON trial. Blood 2008;111:537-543.
11. Hagberg H, Gisselbrecht C. Randomised phase III study of R-ICE versus R-DHAP in relapsed patients with CD20 diffuse large B-cell lymphoma (DLBCL) followed by high-dose therapy and a second randomisation to maintenance treatment with rituximab or not: an update of the CORAL study. Ann.Oncol. 2006;17 Suppl 4:iv31-iv32.
12. Westerterp M, Pruim J, Oyen W et al. Quantification of FDG PET studies using standardised uptake values in multi-centre trials: effects of image reconstruction, resolution and ROI definition parameters. Eur.J.Nucl.Med.Mol.Imaging 2007;34:392-404.
13. Brepoels L, Stroobants S, Verhoef G. PET and PET/CT for response evaluation in lymphoma: current practice and developments. Leuk.Lymphoma 2007;48:270-282.
14. Kasamon YL, Jones RJ, Wahl RL. Integrating PET and PET/CT into the risk-adapted therapy of lymphoma. J.Nucl.Med. 2007;48 Suppl 1:19S-27S.