Ontario
Cancer Care Ontario
Action Cancer Ontario

**Evidence Summary 27-4**

## A Quality Initiative of the Program in Evidence-Based Care (PEBC), Cancer Care Ontario (CCO)

# Positron Emission Tomography in Hodgkin Lymphoma Patients Undergoing Curative-Intent Treatment

*T Baetz, J Dudebout, J Salerno, J Dobranowski, E Eisenhauer, D Langer, L Jimenez-Juan, U Metser, M O'Malley, A Singnurkar*

**Report Date: March 22, 2016**

**Table of Contents**



For information about this document, please contact T Baetz or J Dudebout, the lead authors, through the PEBC via:
Phone: 905-527-4322 ext. 42822 Fax: 905 526-6775 E-mail: ccopgi@mcmaster.ca

For information about the PEBC and the most current version of all reports, please visit the CCO website at http://www.cancercare.on.ca/ or contact the PEBC office at:

Phone: 905-527-4322 ext. 42822 Fax: 905 526-6775 E-mail: ccopgi@mcmaster.ca

**PEBC Report Citation (Vancouver Style)**: Baetz T, Dudebout J, Salerno J, Dobranowski J, Eisenhauer E, Langer D, Jimenez-Juan L, et al. Positron Emission Tomography in Hodgkin Lymphoma Patients Undergoing Curative-Intent Treatment. Toronto (ON): Cancer Care Ontario; *Year Month Day*. Program in Evidence-Based Care Evidence Summary No.: X-XX.>, available on the CCO website: <insert URL when known>.

**Evidence Summary 27-4**

# Positron Emission Tomography in Hodgkin Lymphoma Patients Undergoing Curative-Intent Treatment: Evidence Summary

## THE PROGRAM IN EVIDENCE-BASED CARE

The Program in Evidence-Based Care (PEBC) is an initiative of the Ontario provincial cancer system, Cancer Care Ontario (CCO). The PEBC mandate is to improve the lives of Ontarians affected by cancer through the development, dissemination, and evaluation of evidence-based products designed to facilitate clinical, planning, and policy decisions about cancer control.

The PEBC is a provincial initiative of CCO supported by the Ontario Ministry of Health and Long-Term Care (OMHLTC). All work produced by the PEBC is editorially independent from the OMHLTC.

## INTRODUCTION

Lymphoma is a cancer that affects the network of organs and nodes that comprises the body's immune system, known as the lymph system [1]. There are two main categories of lymphoma: Hodgkin lymphoma (HL) and non-Hodgkin lymphoma (NHL). Treatment schema and outcomes differ between HL and NHL; therefore, this work focuses on HL.

HL has two major subtypes, classical HL and nodular lymphocyte-predominant HL, of which the former subtype accounts for the majority of cases (95%), and is also the focus of the current work. Furthermore, classical HL can be additionally divided into four entities including nodular sclerosis, mixed cellularity, lymphocyte depletion, and lymphocyte-rich. The five-year survival and prognosis of HL is approximately 80% [2].

In Canada, only among women does HL rank as one of the top 20 cancers ($20^{th}$: 0.5%) in terms of the number of new cases of cancer among all new cases of cancer. HL is relatively less frequent among men. HL does not appear in the top 20 causes of death due to cancer as a result of having a high cure rate with modern day therapies. The age-standardized incidence rate of HL is 2.8 per 100,000 [3].

## POSITRON EMISSION TOMOGRAPHY

### Technology

Positron emission tomography (PET) is a functional imaging technique that allows localization and characterization of tumour metabolism. This is accomplished through the use of a radiolabeled glucose analogue, $^{18}$F-fluorodeoxyglucose (FDG), which is transported into and trapped in neoplastic tissue through cell surface glucose transporters. More recently, co-registration of PET and computed tomography (CT) (PET/CT) has emerged as a powerful new hybrid modality that allows more accurate localization of lymphoma and quantification of lymphoma metabolism [4].

### Key Relationships: Biology, Treatment and Interpretive Criteria

Interpretation of FDG-PET scans or FDG-PET/CT scans (PET is used interchangeably with FDG-PET/CT and FDG-PET) can be complex due to a number of factors including the application or choice of interpretive criteria, as well as factors related to the biological response of lymphoma over the course of therapy.

Considering the abovementioned tracer, the biological functionality of FDG in lymphoma is such that its uptake is expected to decline with effective therapy in

chemotherapy-sensitive lymphoma patients. However, the biological response of lymphoma and corresponding treatment effectiveness are intimately linked with the PET interpretive criteria that are used in the assessment of disease status. A number of methods of PET interpretation criteria are currently in use, including absolute quantitative thresholds for lymphoma metabolism or comparison of lymphoma metabolic activity to that in normal structures such as the mediastinum or liver. Most interpretive criteria are considered qualitative (i.e., binary: 0 [disease absence] and 1 [disease presence]; or ranked on a scale: 1–5 such as in the Deauville [five-point)] criteria, with higher scores indicating increasing likelihood of lymphoma). Combining visual analysis of PET imaging with the choice of interpretive criteria lends itself to variation in the accuracy of the lymphoma status for patients [5]. FDG uptake not related to the presence of lymphoma (sometimes referred to as false-positive FDG uptake) may occur for a number of other biological or technical factors such as treatment-related inflammation or timing of imaging. A fundamental principle of imaging is that assessment of disease status is performed with consideration of inflammatory processes that may lead to erroneous results [6]. Residual uptake in tumours responsive to therapy (e.g., due to post-therapy inflammation) has led to a practice pattern that suggests that end-of-treatment PET scans should be performed at a minimum of three weeks (up to six to eight weeks) after the completion of therapy, and that mid-treatment PET scans should be performed just prior to the next chemotherapy cycle [5]. The proposed timing of mid-treatment and end-of-treatment PET scanning is summarized in Appendix 1.

In summary, the relationships among lymphoma biology, treatment, and interpretive criteria in lymphoma are complex. The measurement of viable lymphomatous masses using PET technology during the course of induction chemotherapy reflects the chemo-sensitivity of lymphoma, which may be influenced by a number of heterogeneity factors including lymphoma subtypes, chemotherapy regimens, and the timing of scans. Moreover, the criteria used to interpret the scans and the timing of the study can also affect the *apparent* impact of treatment on the lymphoma [7].

**CLINICAL AND BIOLOGICAL RATIONALE**

PET imaging in lymphoma can be performed at multiple time points along the continuum of care for lymphoma patients, such as: (i) to stage disease prior to initiating treatment; (ii) when patients are undergoing first-line treatment where use of PET imaging is to monitor mid-treatment response ('interim PET'); (iii) at the point in time when re-staging of disease is necessary such as after the completion of treatment or end of treatment ('final' or 'end-of-treatment' PET); and (iv) to detect disease recurrence or relapse among survivors of lymphoma.

Interim PET is aimed at predicting the outcome of a given therapy regime. Early (interim) PET offers a window into the chemo-sensitivity of the lymphoma with the goal of determining whether patients are potential responders to therapy. This information can then be used to alter treatment or to examine alternative strategies to improve complete remissions early in the course of therapy. Information obtained from interim PET scans can avoid potential under-treatment and the risk of relapse, or over-treatment and exposure to unnecessary toxicity and complications [8]. In contrast, when PET is conducted upon the completion of treatment, the aim of its use is to determine remission and residual disease status [6]. In HL, where cure is the goal of front-line therapy, a complete response with a "negative" PET scan is the desired outcome [9].

Due to the sensitivity of PET to key relationships (biological processes and treatment response), use of interim or final PET may better identify patients who would benefit from adaptive treatment or altered clinical management (i.e., positive or negative scan status). For interim PET status, this may include escalating or de-escalating therapy or investigational

approaches. For final PET status, this may include additional treatment (e.g., radiation therapy), biopsy, or repeat scans. However, the prognostic value of PET may depend on the stage of disease, timing of imaging, and chemotherapy regime [10].

## CURRENT STATUS OF TECHNOLOGY IN ONTARIO

Since October 1, 2009, Ontario's evidence-based PET program has provided support for PET scanning through a combination of publicly insured and funded uninsured services, depending on the specific clinical scenario (PET Scans Ontario: www.petscansontario.ca). For lymphoma, PET is an insured service indicated for the evaluation of residual mass(es) following chemotherapy in a patient with HL or NHL when further potentially curative therapy (such as radiation or stem cell transplantation) is being considered; or for the assessment of response in early stage HL following two or three cycles of chemotherapy when chemotherapy is being considered as the definitive single modality therapy. This is aligned with published PEBC work, where it was recommended that a FDG-PET/CT scan be performed for the assessment of early response in early stage (I or II) HL following two or three cycles of chemotherapy when chemotherapy is being considered as the definitive single-modality therapy, to inform completion of therapy or if more therapy is warranted [11]. These recommendations are further echoed in the PEBC PET monitoring reports [12] (Appendix 2). According to recently published PEBC work in HL, the use of a negative interim PET scan to identify patients with early stage HL for whom radiotherapy can be omitted without a reduction in progression-free survival is not recommended [13].

When using PET, the current standard of clinical practice is to "wait as long as possible" after treatment before performing a scan (interim or final) considering what is known about the key relationships in lymphoma: biological response and treatment. However, the exact time interval to "wait" is ambiguous and in practice standardization may be difficult to achieve for a number of patient, institutional, or treatment factors including individual patient treatment plans as to "when" they may be eligible for a PET scan and PET scan availability/booking schedules. In addition, some patients may travel significant distances for their PET scan, further complicating scheduling. These limitations potentially contribute to clinical practice variation, possible inappropriateness and inefficiencies, and decreased patient satisfaction if another PET scan has to be performed or rescheduled after having incurred travel time and costs.

Reporting of findings from PET scans also show lack of standardization, and it is sometimes unclear as to the PET reporting criteria used and/or are ambiguous regarding PET status. Additionally, from the standpoint of the clinician, there is uncertainty in which clinical activities should take place following, in particular, a borderline positive (i.e., Deauville [five-point] criteria score of 3) PET result, which could include an approach of watchful waiting, imaging follow-up, or further evaluation such as biopsy or change in treatment.

## OBJECTIVE

Sponsored by the Cancer Imaging Program of CCO, the aim of this document is to provide a synthesis and summary of the evidence in the area of PET imaging for lymphoma treatment evaluation, with a focus on making conclusions regarding its use (i.e., timing and reporting) in patients undergoing first-line curative-intent treatment. Specifically, we aim to summarize the evidence with respect to the appropriate timing and reporting of interim and end-of-therapy PET use, and the prospective clinical activities following a positive end-of-therapy scan. By addressing the timing, reporting, and clinical activities surrounding PET use, this evidence summary may help address the potential clinical practice variation, potential inappropriateness, and possible inefficiencies and patient dissatisfaction related to PET use in lymphoma, and help guide CCO's Cancer Imaging Program in its future endeavours.

## RESEARCH QUESTIONS

The following research questions were developed to direct the search for available evidence on HL.

**Question 1: What is the ideal timing and reporting of FDG-PET/CT or FDG-PET?**

a) For HL patients during treatment? (i.e., interim PET)
    i. Interim defined as scans performed during first-line curative-intent treatment (i.e., single modality or combined-modality therapy)

b) For HL patients following chemotherapy ± radiation therapy? (i.e., end-of-therapy or final PET)
    i. End-of-treatment PET defined as scans performed immediately upon the completion of treatment (i.e., immediate post-treatment or end of treatment).

**Question 2: What are the clinical activities that should be performed following an end-of-therapy PET(+) scan?**

a) Clinical activities are defined as the number of PET scans, number of CT scans, biopsy, and change in treatment.

## TARGET POPULATION

Patients aged ≥16 years with HL at any stage of disease who are undergoing first-line curative-intent treatment.

## INTENDED PURPOSE

The objective of this work is to provide an evidence summary on PET imaging in the care of HL patients undergoing first-line curative-intent treatment.

## INTENDED USERS

Clinicians involved in the care of HL patients.

**PROSPERO registration number:** CRD42015024431

## METHODS

This evidence summary was developed by a Working Group consisting of medical oncologists, radiologists, nuclear medicine specialists, a hematologist, and a health research methodologist at the request of the Cancer Imaging Program of CCO. The Working Group was responsible for reviewing the identified evidence and drafting the summary. Conflict of interest declarations for all authors are summarized in Appendix 3, and were managed in accordance with the *PEBC Conflict of Interest Policy*.

This evidence review was conducted in two planned stages; a search for systematic reviews followed by a search for primary literature. These stages are described in subsequent sections. Prior to this, a search for existing clinical practice guidelines on the use of PET imaging in HL was undertaken. This was performed to avoid potential duplication of efforts and to evaluate the existing summarized evidence on the topic.

### Search for Existing Clinical Practice Guidelines

A search was conducted for existing clinical practice guidelines. The following criteria were used to select potentially relevant publications: English language, relevant to the research questions based on title screen, publication year between 2007 and 2015 (past eight

years), methods were well-described, and recommendations were articulated. The following sources were searched for existing clinical practice guidelines that addressed the research questions:

- Practice guideline databases: Standards and Guidelines Evidence (SAGE), National Guidelines Clearinghouse
- Guideline developer websites, known publications, and other international groups or associations including:
    - National Comprehensive Cancer Network
    - European Society of Medical Oncology
    - International Harmonization Project
    - International Working Group
    - Malignant Lymphomas Imaging Workshop Group
    - American College of Radiology
    - Program in Evidence-Based Care/Cancer Care Ontario

If appropriate, guidelines that were considered relevant to the objectives and the research questions were then evaluated for quality using the AGREE II instrument [14].

**Search for Existing Systematic Reviews**

A search was conducted for existing systematic reviews. In brief, the following search criteria were used:

- Years: 1995 to October 22, 2015
- Databases searched: MEDLINE, EMBASE, Cochrane Database of Systematic Reviews
- Search terms: FDG-PET/CT, FDG-PET, lymphoma, treatment

Identified systematic reviews were evaluated based on their clinical content and relevance. Relevant systematic reviews were assessed using the 11-item Assessment of Multiple Systematic Reviews (AMSTAR) [15] tool to determine whether existing systematic reviews met a minimum threshold for completeness of reporting and could be considered for inclusion in the evidence base.

**Search for Primary Literature**

For this evidence summary, a search for primary literature was conducted. There was a need for a primary literature search to identify relevant studies that could then be synthesized with respect to the research questions.

***Literature Search Strategy***

A literature search was conducted on October 27, 2015 using MEDLINE and EMBASE and other databases. Details of the literature search can be found in Appendix 4. The literature search was part of a larger initiative addressing lymphoma patients (HL and NHL). For the current work, only studies on HL were considered for reasons that include less biological heterogeneity and less technical variability when using PET, compared with NHL.

***Study Selection Criteria and Process***

*Inclusion Criteria*

- Patients with HL at any stage of disease
- Newly diagnosed HL patients undergoing first-line curative-intent treatment (e.g., chemotherapy ± radiotherapy)
- Biopsy- or histological-confirmed diagnosis
- Minimum study size of 30 patients

- Studies using visual PET analysis
- Observational studies, randomized controlled trials, or systematic reviews/meta-analysis
- Studies that are relevant to the research questions
- Studies on treatment, FDG-PET/CT or FDG-PET, and HL
- Studies written in the English language
- Studies reporting at least one clinical/patient outcome of interest (e.g., progression-free survival, overall survival) (Question 1 only)

*Exclusion Criteria*
- Patients ≤15 years of age
- Case reports (N=1), narrative reviews, in vitro or animal studies, letters, editorials, conference abstracts
- Studies on the technical aspects of FDG-PET/CT or FDG-PET
- Studies on mixed populations of HL and NHL
- Studies on survivors of HL undergoing imaging surveillance for disease relapse or recurrence
- Studies on the clinical effectiveness of different imaging modalities (e.g., PET/CT vs. PET/magnetic resonance imaging)
- Studies using other hybrid technologies or tracers

A review of the titles and abstracts that resulted from the search was conducted by one reviewer (JS). For those items that warranted full-text review, one reviewer reviewed each item (JS).

***Data Extraction and Assessment of Study Quality and Potential for Bias***

Data abstraction was performed by one abstractor (JS) and audited by an independent reviewer. Abstracted data included study variables such as author, publication year, study location, study design, length of follow-up, and sample size; patient characteristics such as age and clinical stage; and PET scanning variables such as type of scan used. Studies were categorized as having a retrospective study design if the PET and outcome information had been collected in the past and prior to the start of the current study (e.g., historical medical record review study). A study was categorized as having a prospective study design if the PET information was collected at baseline and the outcome information for patients was collected at a point in time beyond the starting point of the study (e.g., clinical follow-up for treatment response). Studies that had PET scan interpretations performed prospectively with PET re-interpretation done retrospectively were categorized as having a mixed study design.

Details on the timing of PET use were abstracted including the sequence (i.e., interim or end-of-treatment), after which cycle, and the time interval. The criteria used for evaluating and interpreting PET scans (i.e., interim and end-of-treatment) were abstracted. Specific attention was given as to whether a Deauville score of 3 was classified as positive or negative, and these data were abstracted. Additional details on the treatment regimen were also abstracted.

Outcome information abstracted included performance metrics (sensitivity, specificity, positive predictive value [PPV], negative predictive value [NPV]). A priori, the clinically meaningful performance metric of interest was the NPV (defined as true negatives/true negatives + false negatives) [16] as reported in the included studies. NPV was selected according to the precautionary principle of “do no harm” when planning on de-escalating treatment based on a negative PET scan. For performance metrics, the following

thresholds and interpretations will be used: excellent (80% to 100%), substantial (60% to 80%), moderate (40% to 60%), fair (20% to 40%), and poor (0 to 20%) [17].

The clinical activities following a positive end-of-treatment scan abstracted included whether a repeat PET scan was reported, treatment change occurred, or biopsy was performed. For studies that performed a biopsy on patients with positive end-of-treatment PET scans, the results of those biopsies were abstracted, including clinical/patient outcomes (e.g., progression-free survival) in relation to biopsy status (i.e., positive/negative). All extracted data and information were audited by an independent auditor.

### *Heterogeneity*

A priori, therapy regimen and timing of PET were considered sources of heterogeneity.

### *Risk of Bias*

Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) [18] was used to evaluate the risk of bias for studies that contributed data to performance metrics outcomes. Non-diagnostic accuracy outcomes were evaluated using ACROBAT-NRSI [19].

## Synthesizing the Evidence

The basis of this evidence synthesis is on the data/information abstracted from included studies (i.e., studies that met the inclusion criteria) [20]. Among included studies with suitable data/information, the abstracted data/information were synthesized qualitatively: presented descriptively or narratively in summary tables. According to the research questions, the summary tables were then examined for consensus/trends. Review of the summary tables, therefore, helped direct the conclusions. When there was an absence of data/information or consensus/trends to inform the conclusions, the conclusions were then based on the author's consensus expert clinical opinion; however, this is not regarded as evidence informed [20].

The evidence synthesis involved separating interim PET scanning from end-of-treatment PET scanning. For interim PET scanning, risk-adapted therapy was used as a stratification variable and early HL (stages I and II) was summarized separately from advanced HL (stages III and IV). Established PET criteria was defined as criteria that have been previously published (i.e., Deauville [five-point], International Harmonization Project [IHP], and Gallamini criteria). The criteria used for evaluating and interpreting PET scans were examined in relation to the definition of positivity of PET scans. When the Deauville (five-point) criteria were used, whether a score of 3 led to a positive PET scan classification or a negative PET scan classification was examined. The clinical stage of HL was the stage category with the highest proportion of patients. Age was reported as the median or mean age in years among the study population. For studies where this information was absent, the age category with the highest proportion of patients was reported.

## RESULTS

### Search for Existing Clinical Practice Guidelines and Systematic Reviews

A search for existing clinical practice guidelines revealed a number of publications, including those that are well known in the area of imaging and lymphoma [9,21-23]. However, upon review, these publications did not yield an appropriate source document on which to build an evidence base. Therefore, the AGREE II instrument was not used. The search for existing systematic reviews identified 97 citations (duplicates removed). Full-text review did not identify any systematic reviews that met the inclusion/exclusion criteria and would be able to answer our research questions. As no existing guideline or systematic review was

identified that could be used as a source document, a search for primary literature was performed to answer the proposed research questions.

## Search for Primary Literature

### *Literature Search Results*

A total of 3407 citations (after removal of duplicates) were identified. After title and abstract screening, there were 185 papers that underwent full-text review. After full-text review, 20 studies on HL were identified (Appendix 5).

### *Study Design*

There were 20 included studies on HL [24-43], of which three (15.0%) were clinical trials; two were phase II non-randomized clinical trials [25,42], and the other was a phase III randomized clinical trial [24]. The remaining 17 studies (85.0%) were observational in nature, with 13 of 17 studies (76.5%) of a retrospective design [26-33,35,38-41], two of 17 (11.8%) studies of a prospective design [37,43], and two of 17 (11.8%) studies of a combination of retrospective and prospective designs (mixed) [34,36]. The sample sizes of included studies ranged from less than 50 patients in three studies [26,40,41], to greater than 250 patients in two studies [24,28]. The clinical stage of study populations were predominately low risk (stage I/II, 14 of 20 studies, 70.0%) [24-29,32,35,36,38-40,42,43], with fewer studies including high-risk (stage III/IV, four of 20 studies, 20.0%) populations [30,31,33,37]. One study had equal numbers of low-risk and high-risk patients (50% stage I/II, 50% stage III/IV) [34]. For one study, although the included results were reported for HL separately, the clinical stage information was reported for the entire population (including NHL); therefore, the clinical stage of HL patients for this study was unable to be abstracted [41]. All studies had a histologically confirmed diagnosis of HL as evidenced by staging (Appendix 6, Table 1).

## Quality Assessment

The QUADAS-2 risk of bias tool was used to assess study quality for the 11 studies that contributed summary data to the outcome of performance metrics (Appendix 7, Table 1). In terms of patient selection, all studies included HL patients. Each study included biopsy- or histologically confirmed HL; therefore, the risk of bias due to ill-defined disease status is low. Limiting our work to HL patients promoted homogeneity, and also promoted homogeneity with respect to corresponding treatment strategies. The main risk of bias stems from the study design, where not all studies enrolled patients consecutively or used a random sample of eligible patients, suggesting the potential for selection bias. Overall, patient selection as a source of bias was judged to be mixed (low-high risk).

In terms of bias surrounding the use of the index test, there were five of 11 (45.5%) studies in which the nuclear medicine specialist interpreting the PET scans was reported not to have been blinded to treatment response information during evaluation, which may lead to a high risk of bias among retrospective study designs. Different studies used different criteria to score and interpret PET scans and this lack of standardization of PET use may have introduced bias. However, reporting criteria and thresholds were pre-specified across studies. Overall, the index test of PET scanning as a source of bias was judged to be mixed (low-high risk).

The reference gold standard of clinical follow-up with or without pathology confirmation including laboratory examinations, PET imaging, non-PET imaging, or symptoms were all considered in light of established treatment response criteria (Appendix 8, Table 1). All studies mentioned some aspect of treatment response; however, reporting of the details of the methodology across studies was not consistent or clear. The manner in which

treatment response criteria were used in individual studies was not clear for some studies, with the potential of misclassification of patients and subsequent bias. There was a lack of standardization when evaluating treatment response across studies, such as definitions used. Additionally, all studies lacked information about whether assessors of treatment response were blinded to PET scan results. Overall, use of the reference standard was judged to be inadequate for most studies, thereby suggesting a high risk of bias.

In summary, there are a number of methodological concerns. Lack of blinding of both the PET assessors and assessors of treatment response suggests that information bias cannot be ruled out. Studies were classified as to whether an interim PET or a final PET scan were performed; however, in the absence of biochemical data to support the ideal timing of PET scan use (i.e., the best interval of time by which to achieve a treatment response), it is unclear as to whether the timing of PET scans may have impacted the ability to correctly classify patients. Two studies differed in terms of treatment, which may have led to an inferior treatment response among patients. A main limitation of included studies is the lack of information and unclear standardization around determining patient treatment response including the criteria or definitions used, examinations performed, imaging used, and when treatment response was assessed. In addition, the lack of details around the time interval by which to perform a PET scan, either in the interim or end of treatment, was poorly reported in included studies. Although there are no applicability concerns, in summary, there are a number of methodological concerns that would contribute to a high risk of bias according to QUADAS-2 criteria for the outcome of performance metrics. Overall, the evidence for performance metrics was judged to be fair.

**Question 1: Timing and Reporting of FDG-PET/CT and FDG-PET Scanning**

***Timing of Interim PET Scans***

Fourteen of 20 studies (70.0%) [24-28,30-37,43] reported on the timing of interim PET scanning, with more studies on interim PET in early stage HL (Appendix 6, Tables 2a and 2b). Results show that a majority of studies had an interim PET scan performed after the second cycle of chemotherapy (n=9 studies, 64.3%) [25-28,30,32,33,35,36], followed by after the fourth cycle of chemotherapy (n=3 studies, 21.4%) [31,34,37]. The fewest number of studies reported an interim PET scan after the third cycle of chemotherapy (n=2 studies, 14.3%) [24,32] or after the first cycle of chemotherapy (n=1 study, 7.1%) [43]. The main difference between early stage and advanced stage HL is that no study performed interim PET after cycle 1 or cycle 3 among advanced stage HL studies, whereas this occurrence among early stage HL studies was infrequent. Details on the exact timing of interim PET use such as the interval of time after a given cycle, or prior to the next cycle was mixed, and is shown to be variable (e.g., zero to six days prior to the next cycle; or a range of 10 days after, and 24 to 28 days after the last specified cycle).

***Timing of End-of-Therapy PET Scans***

Fourteen of 20 studies (70.0%) [26,27,29,31-33,35-42] reported on the use of end-of-therapy PET scanning (Appendix 6, Table 3). A majority of studies performed final PET after at least six cycles of therapy (n=3 studies, 21.4%) [31,37,42]; however, not all studies reported after which cycle the final PET scan was performed. The range of the exact timing of end-of-therapy PET scan use is reported to be between two and six weeks after chemotherapy (six studies) [29,31,37-39,42] and within or around the two to three month time period (six studies) [27,29,33,38,40,41]. For five of six studies, the final PET scan was performed at two to three months after combined modality therapy [27,29,33,38,41]. In one study, the time interval after radiation therapy was not specified, only that a portion of

patients had radiation therapy in addition to chemotherapy and that the average time to final PET after chemotherapy was approximately two months [40]. The exact timing of end-of-therapy PET scanning was not reported in detail for four studies [26,32,35,36].

***Reporting of PET Scans (Interim and End of Therapy)***

All 20 studies used some sort of qualitative criteria to interpret and report PET scans [24-43] (Appendix 6, Tables 4a and 4b). Considering established reporting criteria (Table 4a), the PET reporting criteria used most often was the Deauville (five-point) criteria (12 instances, nine studies) [24,25,27-30,33,34,43] followed by the IHP criteria (seven instances, seven studies) [25,26,29,31,34,36,37]. The Gallamini criteria were used the least (three instances, three studies) [29,34,36]. Frequently used were other customized criteria for individual studies (10 instances, nine studies) [27,32,34-36,38-41] (Table 4b).

As shown in Appendix 6, Table 5, when the five-point scale of the Deauville criteria was used, a score of 3 was considered negative more often than positive (nine of 12 instances [75.0%] versus three of 12 instances [25.0%]). Notably, some studies interpreted the Deauville criteria more than one way [28,29].

**Outcomes**

***Performance Metrics***

The performance metrics that were considered included sensitivity, specificity, PPV, and NPV from studies meeting the inclusion criteria.

Eleven of 20 studies (55.0%) examined PET in relation to one or more performance metrics [25,27,29-31,34,36,37,40,42,43] (Appendix 6, Tables 6a and 6b). There were eight studies on interim PET scans [25,27,30,31,34,36,37,43], and four studies on final PET scans [29,31,40,42] in relation to one or more performance metrics of sensitivity, specificity, PPV, and NPV. Markova et al (2012) [31] examined both interim and end-of-therapy scanning in relation to performance metrics. All studies used clinical follow-up with or without pathology confirmation as the reference standard. Notably, the criteria used to assess PET positivity varied, and so did the timing of PET scanning. Within studies that used more than one criterion [25,27,29,34,36], the trend shown was that when a more stringent criteria was used (e.g., Deauville criteria [five-point], only a score of 5 is positive), there were fewer patients categorized as PET-positive compared with when less stringent criteria were used to score PET scans as positive.

***Interim PET Scans***

The range of PET positivity for interim scans ranged from 6.3% to 34.4%. The results show that the NPV ranged from 81.9% to 98.0%. The specificity (i.e., true negatives among true negatives and false PET-positives) ranged from 78.0% to 97.0%. The sensitivity ranged from 33.0% to 68.0%. The PPV ranged from 14.0% to 73.0%. Among interim PET scan studies that specified interim PET was performed after cycle 2, the NPV was >80% (5 studies) (Appendix 6, Table 6a).

***End-of-Therapy PET Scans***

The range of PET positivity for end-of-therapy scans ranged from 13.4% to 25.0%. The results show that the NPV ranged from 86.1% to 98.1%. The specificity (i.e., true negatives among true negatives and false PET-positives) ranged from 89.1% to 98.2%. The sensitivity ranged from 55.0% to 90.9%. The PPV ranged from 62.5% to 92.0% (Appendix 6, Table 6b).

**Question 2: Clinical Activities Following a Positive End-of-Treatment FDG-PET/CT or FDG-PET Scan**

***Clinical Activities***

Nine of 20 (45.0%) studies [26,29,31,36-39,41,42] reported on the clinical activities following a positive end-of-therapy PET scan. A majority of the time, the clinical activity that was reported was a change in treatment (four instances) or a biopsy (six instances). To a lesser extent, a repeat PET scan was the clinical activity (one instance) (Appendix 6, Table 7). For the six studies [29,36,38,39,41,42] reporting a biopsy following a positive end-of-therapy scan, the results of those biopsies are described in Appendix 6, Table 8. No studies reported on clinical/patient outcomes such as progression-free survival or overall survival in relation to biopsy results. There was a lack of evidence on clinical/patient outcomes.

**Heterogeneity**

The following variables were considered a priori as sources of heterogeneity: therapy regimen and timing of PET. Most studies used comparable and standard regimens of chemotherapy plus radiotherapy [44], except in one study that did not include radiation therapy [25]. One study used anthracycline-based chemotherapy, not otherwise specified. Therefore, it appears that most studies provided appropriate therapy. Seven studies reported a change in therapy management due to interim PET scan results; however, the influence of using a risk-adapted therapeutic approach among those studies is not evident as it relates to our work. Details of the therapeutic management are shown in Appendix 9, Table 1. Timing of interim PET scan use was reported poorly in the included studies, and was shown to be variable. Timing of final PET scan use was reported more consistently but again lacked consistent reporting across studies.

**DISCUSSION**

The exact timing (interim and final) of PET scanning was shown to be variable among included studies; however, the NPV across different sequences and timing of PET scanning, interpretation criteria, and treatment regimens was shown to be high (>80%), even among studies that specified interim PET was performed after cycle 2, although based on fair quality of evidence. It appeared that when using established reporting criteria, the Deauville (five-point) criteria and IHP predominated, with a tendency to report a Deauville score of 3 as negative. Clinical activities following a positive final scan were predominately biopsy and treatment change; however, there was a lack of evidence on clinical/patient outcomes. Due to the lack of high-quality evidence in the form of randomized controlled trials, the conclusions of this evidence synthesis rely heavily on the authors' consensus based expert clinical opinion.

**Timing of PET Scans**

***Interim PET Scans***

For interim PET scanning, the evidence synthesis showed that there is no firm time; however, generally occurs in the days immediately leading up to the next cycle of chemotherapy. Our evidence synthesis builds upon current clinical practice including current guidelines and government policy (Appendices 1 and 2), and what is known about key relationships in lymphoma (biological response, treatment, and interpretive criteria). We showed that performing an interim PET scan after cycle 2 predominates in the literature (nine studies); however, the exact time interval was not shown from the evidence.

Therefore, in support of current clinical practice and based on the authors' consensus expert clinical opinion, the following benchmarks are supported:

- An interim PET scan should generally be performed as far away from cycle 2 and as close as possible to cycle 3 as is feasible.
- However, as a general qualifying statement, the role of interim PET is still evolving. When acquired, findings from interim PET should be used with the utmost regard for optimal standards of practice.

***End-of-Therapy PET Scans***

When considering final PET after chemotherapy, our evidence supports an interval of time of a minimum of two weeks (14 days) (six studies). However, preferably, and in agreement with published guidelines (Appendix 2), the time interval to perform a final PET scan after chemotherapy is preferably at least three weeks (21 days). When considering final PET after combined modality treatment, our evidence supports an interval of time of two to three months (six studies). Again, this is in agreement with published guidelines (Appendix 2).

Therefore, in support of published guidelines and based on the authors' consensus expert clinical opinion, the following benchmarks are supported:

- End-of-therapy PET may be performed at a minimum of 14 days after chemotherapy but preferably 21 days from chemotherapy end of treatment.
- End-of-therapy PET may be performed at two to three months following radiotherapy.

**Reporting of PET Scans**

The reporting of PET scanning was variable. When using the Deauville (five-point) criteria, the current standard of practice is to categorize a score of 3 as negative; however, ultimately, decisions with respect to patient care rest with the treating physician and up-to-date knowledge of the correlation between PET results and the clinical circumstances. The evidence tends to support the current clinical practice of scoring a Deauville score of 3 as negative; however, the evidence was mixed with regard to the reporting criteria used (established versus other). Our consensus expert clinical opinion continues to endorse the use of the Deauville (five-point) criteria as part of current standard of practice. However, there is a gap of provincial consensus recommendations on the interpretation of Deauville scores of 3.

Therefore, our consensus expert clinical opinion supports using the Deauville (five-point) interpretation criteria for use in Ontario, as described by an international group in the publication by Barrington et al (2014):

- Deauville scores of 1 and 2 are considered to represent complete metabolic response (i.e., PET-negative scan and cancer-free). Deauville scores of 3 also likely represents complete metabolic response at interim and good prognosis at completion of standard treatment. However, in trials where de-escalation is based on PET response, it may be preferable to consider a Deauville score of 3 as inadequate response to avoid under-treatment.
- The Deauville (five-point) scale for reporting PET/CT results may be interpreted in the context of the anticipated prognosis, clinical findings, and other markers of response. Deauville scores of 1 and 2 represent complete metabolic response, whereas Deauville scores of 3 also probably represent complete metabolic response in patients receiving standard treatment.

**Clinical Activities Following a Positive End-of-Treatment PET Scan**

There were no studies that reported on clinical/patient outcomes such as progression-free survival or overall survival in relation to biopsy results when biopsy was the choice of

clinical management following a PET-positive scan. There was a lack of evidence on clinical/patient outcomes. More work in the field of lymphoma in the form of high-quality studies is needed to address our proposed question 2.

Therefore, based on the authors' consensus expert clinical opinion, future work such as consensus recommendations regarding the appropriate clinical action following end-of-therapy PET-positive scans would be beneficial.

**Overall Strengths and Limitations**

This evidence summary focused only on HL patients, thereby addressing one of the major criticisms of previous work of including mixed populations of HL and NHL patients in reviews. The included studies were homogeneous with respect to the main histological type of HL, being classical HL, which makes up 95% of all cases and includes the following subgroups: nodular sclerosis, mixed cellularity, lymphocyte depleted, and lymphocyte rich. However, the included studies were heterogeneous in terms of the timing of PET scan use, the use of reporting and interpretation criteria for PET scanning, and determination of treatment response, thus making the interpretation of the results based solely on the evidence unclear. Therefore, our conclusions rely heavily on our consensus expert clinical opinion with the use of the data/information gathered and synthesized here as a starting point. Future work may continue to assemble and synthesize the evidence base on the topic.

For question 1, the inclusion criteria specified that studies were to be included if they reported on at least one clinical/patient outcome; however, at the time of data abstraction, the authors' consensus-based expert clinical opinion was that these outcomes were no longer a focus of the current evidence summary as specified a priori for question 1. Studies reporting on PET timing or performance metrics only, without clinical/patient outcomes, were not included in the current evidence base (see inclusion criteria). Accordingly, a larger than usual number of studies underwent full-text review (185 of 414, 45%) to carefully include studies that would answer question 2 while preserving the originally stated inclusion criteria pertaining to question 1. As a result, this evidence summary represents a subset of the available body of literature had the inclusion criteria been much broader.

Future work should address standardized protocols for the assessment of treatment response (reference standard) and PET imaging (index test). The included body of literature represents current and contemporary evidence that was assembled using systematic review methodology and examined in an in-depth and comprehensive manner in relation to the research questions. Included studies were evaluated in light of the quality of the evidence and expert clinical opinion filled in gaps where the evidence was lacking. Furthermore, the body of literature on the topic was synthesized in an expedited fashion. Overall, given the vast heterogeneity of included studies in terms of the timing of PET and to a lesser extent the reporting of PET, our conclusions reflect a blended synthesis of systematic review methodology with consensus expert clinical opinion, and the conclusions are heavily based on consensus expert clinical opinion. The evidence base from the systematic review should be interpreted as fair, with the necessary limitations in mind.

**CONCLUSIONS**

In first-line treatment of HL, interim PET scanning is most commonly performed after cycle 2. Interim PET scanning should generally be performed as close to the next chemotherapy cycle (i.e., as far away from cycle 2 and as close as possible to cycle 3) as is feasible. End-of-therapy PET scanning may be performed at least 14 days and preferably 21 days following chemotherapy and two to three months following radiation therapy. Consensus recommendations on interpretation of Deauville 3 scores and appropriate clinical action for PET-positive scans would be beneficial.

## INTERNAL REVIEW

The evidence summary was reviewed by the Director of the PEBC. The Working Group is responsible for ensuring the necessary changes are made.

### Approval by Cancer Imaging Program

After internal review, the report was presented to the Cancer Imaging Program. The Cancer Imaging Program reviewed the document and formally approved the document (March 14th, 2016).


## ACKNOWLEDGEMENTS

The Cancer Imaging Program and the Working Group would like to thank the following individuals for their assistance in developing this report:

- Melissa Brouwers, Sheila McNair, Hans Messersmith, Erin Kennedy for providing feedback on draft versions.
- Kristy Yiu for conducting a data audit.
- Sara Miller for copy editing.

## REFERENCES


1. National Cancer Institute. Lymphoma [Accessed December 16, 2015]. Available from: http://www.cancer.gov/types/lymphoma.
2. King RL, Howard MT, Bagg A. Hodgkin lymphoma: pathology, pathogenesis, and a plethora of potential prognostic predictors. Adv Anat Pathol. 2014;21(1):12-25.
3. Canadian Cancer Society's Advisory. Committee on Cancer Statistics. Canadian Cancer Statistics 2015. Toronto ON, Canadian Cancer Society; 2015.
4. Gallamini A, Borra A. Role of PET in lymphoma. Curr Treat Options Oncol. 2014;15(2):248-61.
5. Meignan M, Hutchings M, Schwartz LH. Imaging in lymphoma: The key role of fluorodeoxyglucose-positron emission tomography. Oncologist. 2015;20(8):890-5.
6. Coughlan M, Elstrom R. The use of FDG-PET in diffuse large B cell lymphoma (DLBCL): predicting outcome following first line therapy. Cancer Imaging. 2014;14:34.
7. Cashen AF, Dehdashti F, Luo J, Homb A, Siegel BA, Bartlett NL. 18F-FDG PET/CT for early response assessment in diffuse large B-cell lymphoma: Poor predictive value of international harmonization project interpretation. J Nucl Med. 2011;52(3):386-92.
8. Uslu L, Donig J, Link M, Rosenberg J, Quon A, Daldrup-Link HE. Value of 18F-FDG PET and PET/CT for evaluation of pediatric malignancies. J Nucl Med. 2015;56(2):274-86.
9. Cheson BD, Pfistner B, Juweid ME, Gascoyne RD, Specht L, Horning SJ, et al. Revised response criteria for malignant lymphoma. J Clin Oncol. 2007;25(5):579-86.
10. Schwenzer NF, Pfannenberg AC. PET/CT, MR, and PET/MR in lymphoma and melanoma. Semin Nucl Med. 2015;45(4):322-31.
11. Kouroukis C, Cheung M, Sussman J, Hodgson D, Freeman M, Kellett S. The Clinical Utility of Positron Emission Tomography in the Diagnosis, Staging, and Clinical Management of Patients with Lymphoma: Recommendation Report. 2015 [Accessed March 3, 2016]. Available from: https://www.cancercare.on.ca/common/pages/UserFile.aspx?fileId=334671.
12. Poon R, and The Program in Evidence Based Care Disease Site Group Reviewers. Evidence from Primary Studies and Systematic Reviews and Recommendations from Clinical Practice Guidelines January to June 2015. 2015 [Accessed March 3, 2016]. Available from: https://www.cancercare.on.ca/common/pages/UserFile.aspx?fileId=348344.
13. Herst J, Crump M, Baldassarre F, MacEachern J, Sussman J, Hodgson D, et al. Management of Early-Stage Hodgkin Lymphoma. 2015 [Accessed March 3, 2016]. Available from: https://www.cancercare.on.ca/common/pages/UserFile.aspx?fileId=350545.
14. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. CMAJ. 2010;182(18):E839-42.
15. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol. 2007;7:10.
16. Sackett DLH, B.R.; Tugwell, P. The Interpretation of Diagnostic Data, in Clinical Epidemiology: A Basic Science for Clinical Medicine. Boston/Toronto: Little Brown and Company; 1985.
17. Szklo M, Nieto JF. Chapter 8 Quality Assurance and Control. Epidemiology Beyond the Basics. Gaithersburg, Maryland, USA: Aspen Publishers, Inc.; 2000. p. 343-404.

18. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-36.
19. Sterne JAC, Higgins JPT, Reeves BC on behalf of the development group for ACROBATNRSI. A Cochrane Risk Of Bias Assessment Tool: for Non-Randomized Studies of Interventions (ACROBATNRSI), Version 1.0.0, 24 September 2014. Available from http://www.riskofbias.info [Accessed July, 2015].
20. Goodman CS, Ahn R. Methodological approaches of health technology assessment. Int J Med Inform. 1999;56(1-3):97-105.
21. Barrington SF, Mikhaeel NG, Kostakoglu L, Meignan M, Hutchings M, Mueller SP, et al. Role of imaging in the staging and response assessment of lymphoma: consensus of the International Conference on Malignant Lymphomas Imaging Working Group. J Clin Oncol. 2014;32(27):3048-58.
22. Cheson BD, Fisher RI, Barrington SF, Cavalli F, Schwartz LH, Lister TA. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. J Clin Oncol. 2014;32(27):3059-68.
23. Juweid ME, Stroobants S, Hoekstra OS, Mottaghy FM, Dietlein M, Guermazi A, et al. Use of positron emission tomography for response assessment of lymphoma: consensus of the Imaging Subcommittee of International Harmonization Project in Lymphoma. J Clin Oncol. 2007;25(5):571-8.
24. Radford J, Illidge T, Counsell N, Hancock B, Pettengell R, Johnson P, et al. Results of a trial of PET-directed therapy for early-stage Hodgkin's lymphoma. N Engl J Med. 2015;372(17):1598-607.
25. Kostakoglu L, Schoder H, Johnson JL, Hall NC, Schwartz LH, Straus DJ, et al. Interim [18F]fluorodeoxyglucose positron emission tomography imaging in stage III non-bulky Hodgkin lymphoma: would using combined positron emission tomography and computed tomography criteria better predict response than each test alone? Leuk Lymphoma. 2012;53(11):2143-50.
26. Iltis A, Eder V, Blasco H, Colombat P, Senecal D. Decisional early interim F-fluoro-2-deoxy-D-glucose positron emission tomography after two cycles of chemotherapy in de novo Hodgkin lymphoma. Acta Haematol. 2015;133(2):172-8.
27. Rigacci L, Puccini B, Zinzani PL, Biggi A, Castagnoli A, Merli F, et al. The prognostic value of positron emission tomography performed after two courses (INTERIM-PET) of standard therapy on treatment outcome in early stage Hodgkin lymphoma: a multicentric study by the Fondazione Italiana Linfomi (FIL). Am J Hematol. 2015;90(6):499-503.
28. Simontacchi G, Filippi AR, Ciammella P, Buglione M, Saieva C, Magrini SM, et al. Interim PET after two ABVD cycles in early-stage Hodgkin lymphoma: outcomes following the continuation of chemotherapy plus radiotherapy. Int J Radiat Oncol Biol Phys. 2015;92(5):1077-83.
29. Kajary K, Molnar Z, Gyorke T, Szakall Jr S, Molnar P, Lengyel Z. Comparison of the International Harmonization Project, London and Gallamini criteria in the interpretation of 18F-FDG PET/CT examinations after first-line treatment in Hodgkin's lymphoma. Nucl Med Commun. 2014;35(2):169-75.
30. Rossi C, Kanoun S, Berriolo-Riedinger A, Dygai-Cochet I, Humbert O, Legouge C, et al. Interim 18F-FDG PET SUVmax reduction is superior to visual analysis in predicting outcome early in Hodgkin lymphoma patients. J Nucl Med. 2014;55(4):569-73.
31. Markova J, Kahraman D, Kobe C, Skopalova M, Mocikova H, Klaskova K, et al. Role of [18F]-fluoro-2-deoxy-d-glucose positron emission tomography in early and late therapy

assessment of patients with advanced Hodgkin lymphoma treated with bleomycin, etoposide, adriamycin, cyclophosphamide, vincristine, procarbazine and prednisone. Leuk Lymphoma. 2012;53(1):64-70.
32. Barnes JA, LaCasce AS, Zukotynski K, Israel D, Feng Y, Neuberg D, et al. End-of-treatment but not interim PET scan predicts outcome in nonbulky limited-stage Hodgkin's lymphoma. Ann Oncol. 2011;22(4):910-5.
33. Gallamini A, Patti C, Viviani S, Rossi A, Fiore F, Di Raimondo F, et al. Early chemotherapy intensification with BEACOPP in advanced-stage Hodgkin lymphoma patients with a interim-PET positive after two ABVD courses. Br J Haematol. 2011;152(5):551-60.
34. Le Roux PY, Gastinne T, Le Gouill S, Nowak E, Bodet-Milin C, Querellou S, et al. Prognostic value of interim FDG PET/CT in Hodgkin's lymphoma patients treated with interim response-adapted strategy: comparison of International Harmonization Project (IHP), Gallamini and London criteria. Eur J Nucl Med Mol Imaging. 2011;38(6):1064-71.
35. Luminari S, Cesaretti M, Tomasello C, Guida A, Bagni B, Merli F, et al. Use of 2-[18F]fluoro-2-deoxy-D-glucose positron emission tomography in patients with Hodgkin lymphoma in daily practice: a population-based study from Northern Italy. Leuk Lymphoma. 2011;52(9):1689-96.
36. Dann EJ, Bar-Shalom R, Tamir A, Epelbaum R, Avivi I, Ben-Shachar M, et al. A functional dynamic scoring model to elucidate the significance of post-induction interim fluorine-18-fluorodeoxyglucose positron emission tomography findings in patients with Hodgkin's lymphoma. Haematologica. 2010;95(7):1198-206.
37. Markova J, Kobe C, Skopalova M, Klaskova K, Dedeckova K, Plutschow A, et al. FDG-PET for assessment of early treatment response after four cycles of chemotherapy in patients with advanced-stage Hodgkin's lymphoma has a high negative predictive value. Ann Oncol. 2009;20(7):1270-4.
38. Advani R, Maeda L, Lavori P, Quon A, Hoppe R, Breslin S, et al. Impact of positive positron emission tomography on prediction of freedom from progression after Stanford V chemotherapy in Hodgkin's disease. J Clin Oncol. 2007;25(25):3902-7.
39. Hueltenschmidt B, Sautter-Bihl ML, Lang O, Maul FD, Fischer J, Mergenthaler HG, et al. Whole body positron emission tomography in the treatment of Hodgkin disease. Cancer. 2001;91(2):302-10.
40. Guay C, Lepine M, Verreault J, Benard F. Prognostic value of PET using 18F-FDG in Hodgkin's disease for posttreatment evaluation. J Nucl Med. 2003;44(8):1225-31.
41. Zinzani PL, Fanti S, Battista G, Tani M, Castelucci P, Stefoni V, et al. Predictive role of positron emission tomography (PET) in the outcome of lymphoma patients. Br J Cancer. 2004;91(5):850-4.
42. Straus DJ, Johnson JL, LaCasce AS, Bartlett NL, Kostakoglu L, Hsi ED, et al. Doxorubicin, vinblastine, and gemcitabine (CALGB 50203) for stage I/II nonbulky Hodgkin lymphoma: pretreatment prognostic factors and interim PET. Blood. 2011;117(20):5314-20.
43. Hutchings M, Kostakoglu L, Zaucha JM, Malkowski B, Biggi A, Danielewicz I, et al. In vivo treatment sensitivity testing with positron emission tomography/computed tomography after one cycle of chemotherapy for Hodgkin lymphoma. J Clin Oncol. 2014;32(25):2705-11.
44. Uhm J, Kuruvilla J. Treatment of newly diagnosed advanced stage Hodgkin lymphoma. Blood Rev. 2012;26(4):167-74.

## Appendix 1: Published Guidelines on the Timing of PET Use

**Juweid et al (2007) – International Harmonization Project**

- PET should not be performed before at least three weeks after chemotherapy and preferably eight to 12 weeks after completion of radiotherapy.
- PET should be performed as close as possible (i.e., within four days) before the subsequent cycle.

**Cheson et al (2007) – International Harmonization Project**

- PET scans should not be performed for at least three weeks, and preferably six to eight weeks after the completion of therapy.

## Appendix 2: Current Ontario Recommendations

### Evidence from Primary Studies and Systematic Reviews and Recommendations from Clinical Practice Guidelines January to June 2015

R. Poon and the Program in Evidence-Based Care Disease Site Group Reviewers
(https://www.cancercare.on.ca/common/pages/UserFile.aspx?fileId=348344)
Program in Evidence-Based Care (PEBC), Cancer Care Ontario (CCO)
Report Date: October 13, 2015

Current Recommendations for the Utilization of PET/CT in Hematologic Cancer

- When functional imaging is considered to be important in situations where anatomical imaging is equivocal and/or in potentially curable cases, a FDG-PET/CT scan is recommended.
- When functional imaging is considered to be important in situations where anatomical imaging is equivocal and treatment choices may be affected in limited-stage indolent lymphomas, a FDG-PET/CT scan is recommended.
- An FDG-PET/CT scan is recommended for the assessment of early response in early stage (I or II) Hodgkin lymphoma following two or three cycles of chemotherapy when chemotherapy is being considered as the definitive single-modality therapy, to inform completion of therapy or whether more therapy is warranted.
- In potentially curable cases, when functional imaging is considered to be important and conventional imaging is equivocal, a FDG-PET/CT scan is recommended to investigate recurrence of Hodgkin lymphoma or non-Hodgkin lymphoma.
- An FDG-PET/CT scan is recommended for the evaluation of residual mass(es) following chemotherapy in a patient with Hodgkin or non-Hodgkin lymphoma when further potentially curative therapy (such as radiation or stem cell transplantation) is being considered and when biopsy cannot be safely or readily performed.
- An FDG-PET/CT scan is not recommended for the routine monitoring and surveillance of lymphoma.

## Appendix 3: Membership

**Table 1. Members of the Lymphoma Imaging Working Group**

| Name | Affiliation | Declarations of interest |
|---|---|---|
| Tara Baetz (Medical oncology) | Cancer Centre at Kingston General Hospital and Queen's University<br>Kingston, ON | None declared |
| Jill Dudebout (Hematology) | Cancer Centre at Kingston General Hospital and Queen's University<br>Kingston, ON | None declared |
| Jennifer Salerno (Methodologist) | Program in Evidence-Based Care<br>Hamilton, ON | None declared |
| Julian Dobranowski (Radiology) | St. Joseph's Healthcare Hamilton and Cancer Care Ontario<br>Hamilton, ON | None declared |
| Elizabeth Eisenhauer (Medical oncology) | Cancer Centre at Kingston General Hospital and Queen's University<br>Kingston, ON | None declared |
| Laura Jimenez-Juan (Radiology) | Sunnybrook Health Sciences Centre<br>Toronto, ON | None declared |
| Ur Metser (Radiology/Nuclear medicine) | University Health Network, Princess Margaret Hospital<br>Toronto, ON | None declared |
| Martin O'Malley (Radiology) | University Health Network, Princess Margaret Hospital<br>Toronto, ON | None declared |
| Amit Singnurkar (Nuclear medicine) | Juravinski Hospital and Cancer Centre<br>Hamilton, ON | None declared |

**Conflict of Interest**

In accordance with the PEBC Conflict of Interest (COI) Policy PEBC Conflict of Interest (COI) Policy, the authors were asked to disclose potential conflicts of interest. As of March 2016, there were no conflicts of interest. To obtain a copy of the policy, please contact the PEBC office by email at ccopgi.mcmaster.ca.

## Appendix 4: Literature Search Strategy

Databases: EBM Reviews - Cochrane Central Register of Controlled Trials September 2015, Ovid MEDLINE® In-Process & Other Non-Indexed Citations and Ovid MEDLINE® 1946 to Present, Embase 1974 to 2015 October 26

**Searches**

| | |
|---|---|
| 1 | exp Lymphoma/ |
| 2 | lymphoma$.mp. |
| 3 | exp hodgkin disease/ |
| 4 | hodgkin$.mp. |
| 5 | exp nonhodgkin lymphoma/ |
| 6 | exp B cell lymphoma/ |
| 7 | (nonhodgkin$ or non?hodgkin$ or B?cell lymphoma$ or diffuse large B?cell lymphoma$ or DLBCL).mp. |
| 8 | or/1-7 |
| 9 | exp deoxyglucose/ |
| 10 | (deoxyglucose or desoxyglucose or deoxy-glucose or desoxy-glucose or deoxy-d-glucose or desoxy-d-glucose or 2deoxyglucose or 2deoxy-d-glucose or fluorodeoxyglucose or fluorodesoxyglucose or fludeoxyglucose or fluordeoxyglucose or fluordesoxyglucose or fluoro-2-deoxy-d-glucose or 2fluoro-2deoxyglucose or 2-fluoro-2-deoxy-d-glucose or 2-fluoro-2-deoxyglucose or 18fluorodeoxyglucose or 18fluorodesoxyglucose or fluoro-d-glucose or 18fluordeoxyglucose or 18fluordesoxyglucose or fluorine-18-fluorodeoxyglucose or fdg$ or 18fdg$ or 18f-dg$ or f-18-dg or 18f-fdg).mp. |
| 11 | (glucose and (fluor or 2fluor$ or fluoro or fluorodeoxy or fludeoxy or fluorine or 18f or 18flu$)).mp. |
| 12 | exp fluorodeoxyglucose/ |
| 13 | exp Fluorodeoxyglucose F18/ |
| 14 | fluorodeoxyglucose F18.mp. |
| 15 | exp emission tomography/ |
| 16 | exp positron emission tomography/ |
| 17 | exp tomography, emission-computed/ |
| 18 | (positron emission tomography or PET?scan$ or PET-FDG or FDG-PET or PET-CT or PET$CT).mp. |
| 19 | (emission and (tomography or tomograph$ or tomographic$ or tomographies)).mp. |
| 20 | exp tomography, x-ray/ |
| 21 | computer assisted tomography/ |
| 22 | (comput$ adj1 (tomography or tomograph$ or tomograhpic$ or tomographies)).mp. |
| 23 | (CT or CT?scan).mp. |
| 24 | biopsy.mp. |
| 25 | ((treatment adj1 change) or (treatment adj1 management) or (treatment adj1 modification$) or (management adj1 strateg$)).mp. |

26 or/9-14

27 or/15-19

28 26 and 27

29 8 and 28

30 (early or early?PET or interim$ or interim?PET or I?PET or mid?therapy or (treatment adj1 monitoring) or final?PET or F?PET or end?of?therapy or end?of?treatment or post?therapy or treatment).mp.

31 29 and 30

32 or/20-25

33 31 and 32

34 31 or 33

35 (comment or letter or editorial or note or erratum or short survey or news or newspaper article or patient education handout or case report or historical article).pt.

36 exp Animal/ not Human/

37 35 or 36

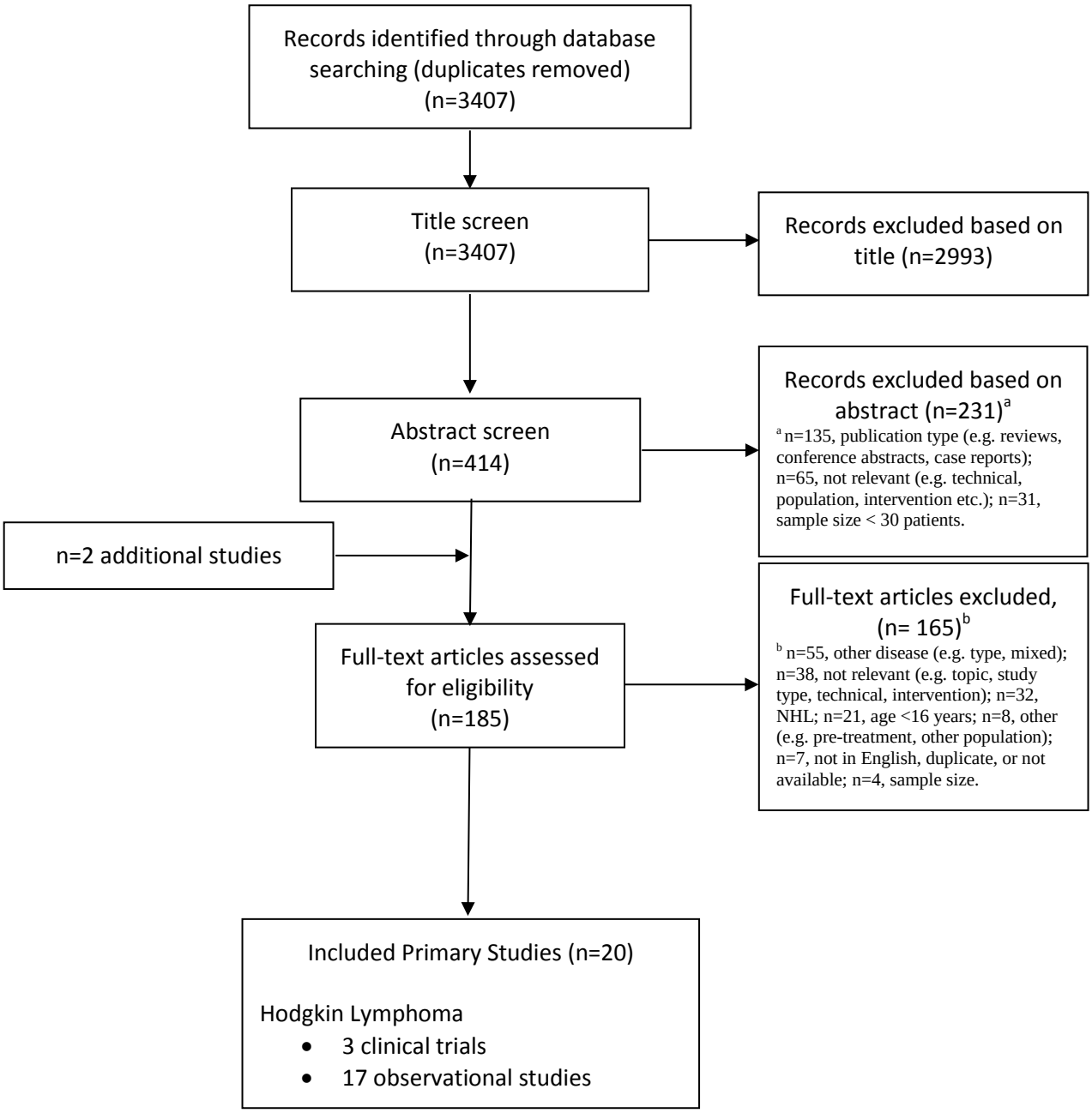
38 34 not 37

39 limit 38 to yr="1995 -Current"

40 remove duplicates from 39

## Appendix 5: PRISMA Flow Diagram

**Figure 1. Citation Flow Chart**



## Appendix 6: Summary Tables

**Table 1. Characteristics of Included Studies (N=20 Studies)**

| Author (Year) | Study, Country | Study Design | Clinical Stage[1] (% pts) | Age (Years) | PET Type | PET Use | | | M-FU (months) | SS[2] |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | BL | I-PET | F-PET | | |
| **CLINICAL TRIALS** | | | | | | | | | | |
| Radford et al (2015)[3] | RAPID, UK | Phase 3 NI | 33% stage IA, 67% stage IIA | 34.7 | PET, PET/CT | - | X | - | 60.0 | 565 |
| Kostakoglu et al (2012)* | CALGB, USA | Phase 2 | 72% stage IIA | <40.0[4] | PET, PET/CT | X | X | - | 39.6[5] | 88 |
| Straus et al (2011)* | CALGB, USA | Phase 2 | 71% stage IIA | 37.0 | PET, PET/CT | X | X | X | 39.6[5] | 99 |
| **OBSERVATIONAL STUDIES** | | | | | | | | | | |
| ***Stages I-II (early HL)*** | | | | | | | | | | |
| Iltis et al (2015) | France | RCS | 52% stage II | 37.5 | PET/CT | X | X | X | 62.4 | 48 |
| Rigacci et al (2015) | Italy | RCS | 91% stage II | 30.0 | PET, PET/CT | X | X | X | 46.0 | 246 |
| Simontacchi et al (2015) | Italy | RCS | 80% stage IIA | ≤50.0[4] | PET | X | X | - | 56.0 | 257 |
| Hutchings et al (2014) | USA, Europe | PCS | 54% stages I-II | 34.1 | PET/CT | X | X | - | 29.1 | 126 |
| Kajary et al (2014) | Hungary | RCS | 68% stages I-II | 32.5 | PET/CT | X | - | X | 54.0 | 66 |
| Barnes et al (2011) | USA | RCS | 88% stage II | 34.0 | PET | - | X | X | 46.0 | 96 |
| Le Roux et al (2011)** | France | Mixed | 50% stages I-II | 31.2 | PET/CT | X | X | - | 49.0 | 90 |
| Luminari et al (2011) | Italy | RCS | 53% stage II | 38.0 | PET/CT | X | X | X | 30.0 | 136 |
| Dann et al (2010) | Israel | Mixed | 54% stages I-II | 30.0 | PET/CT | X | X | X | 59.0 | 96 |
| Advani et al (2007) | USA | RCS | 73% stage I-II | 29.0 | PET, PET/CT | X | - | X | 48.0 | 81 |
| Guay et al (2003) | Canada | RCS | 52% stages I-II | 38.0 | PET | - | - | X | 16.2 | 48 |
| Hueltenschmidt et al (2001)[6] | Germany | RCS | 52% stage II | 38.1 | PET | X | - | X | 20.4 | 81 |
| ***Stages III-IV (advanced HL)*** | | | | | | | | | | |
| Rossi et al (2014) | France | RCS | 63% stages III-IV | 35.5 | PET/CT | X | X | - | 50.0 | 59 |
| Markova et al (2012)[7] | Czech Republic | RCS | 62% stages III-IV | 30.7 | PET | X | X | X | 52.0 | 69 |
| Gallamini et al (2011) | Italy, USA | RCS | 53% stage III+ | 34.0 | PET, PET/CT | X | X | X | 34.0 | 165 |
| Le Roux et al (2011)** | France | Mixed | 50% stages III-IV | 31.2 | PET/CT | X | X | - | 49.0 | 90 |
| Markova et al (2009) | Czech Republic | PCS | 40% stage IV | <50.0[4] | PET | - | X | X | NR | 50 |
| ***Stage Unknown*** | | | | | | | | | | |
| Zinzani et al (2004)[8] | Italy | RCS | NR | 41.0 | PET | - | - | X | NR | 41[5] |

Abbreviations: BL, baseline; CALGB, Cancer and Leukemia Group B; CT, computed tomography; F-PET, final PET; FU, follow-up; HL, Hodgkin lymphoma; I-PET, interim PET; IFRT, involved field radiation therapy; M, median; NI, non-inferiority; NR, not reported; PCS, prospective cohort study; PET, positron emission tomography; RAPID, Randomized Phase III Trial to Determine the Role of FDG-PET Imaging in Clinical Stages IA/IIA Hodgkin Disease; RCS, retrospective cohort study; SS, sample size.

1 Clinical stage reported as the stage category that represents the majority of patients in the study population. Rounded to the nearest whole number.

2 After exclusions. Reported for Hodgkin lymphoma only for Zinzani (2004).

3 For 420 patients with PET(-) results that then underwent randomization to receive no further treatment vs. patients to receive IFRT. Plus, the 145 patients who had PET(+) results.

4 Age reported as median or mean for the entire study population or as the age category with the highest proportion of patients.

5 Median follow-up based on non-progressing patients.

5 Stage based on 25 patients that had baseline staging; mean follow-up time.

6 Staging based on 69 patients as the denominator; age based on mean or median, not clear.

7 Sample size refers to the subgroup of Hodgkin patients for which abstracted information is based upon and reported.

*Note: Straus (2011) is the larger trial upon which Kostakoglu (2012) is based, with different and overlapping methodology, therefore both studies were included.

**Note: Le Roux (2011) is listed twice as its study population is evenly split between stages I-II (50%) and stages III-IV (50%) patients.

**Table 2a. Sequence of Interim PET in Stages I-II (Early) Hodgkin Lymphoma (N=10 Studies)**

| Author (Year) | Description of Interim PET Timing | Cycle | | | |
|---|---|---|---|---|---|
| | | C1 | C2 | C3 | C4 |
| **Clinical Trials** | | | | | |
| ***Change in Treatment***[1] | | | | | |
| Radford et al (2015) | During the 2 wks after day 15 of ABVD cycle 3. | | | X | |
| ***No Change in Treatment*** | | | | | |
| Kostakoglu et al (2012)* | After 2 cycles of AVG, performed 0-6 d prior to cycle 3. | | X | | |
| **Observational Studies** | | | | | |
| ***Change in Treatment*** | | | | | |
| Iltis et al (2015) | After 2 cycles of first-line CT. For ABVD, PET was performed from 11 to 28 d after the 2nd cycle. For VABEM, PET was performed from 24 to 28 d after the 2nd cycle. | | X | | |
| Simontacchi et al (2015) | After 2 cycles of ABVD therapy, approximately 10 d after completion of the 2nd ABVD cycle. | | X | | |
| Le Roux et al (2011)** | After 4 cycles of ABVD. | | | | X |
| Luminari et al (2011) | Performed after 2-3 cycles of therapy (majority with ABVD/ABVD-like CT). | | X[2] | | |
| Dann et al (2010) | After the 1st or 2nd cycle of BEACOPP. | | X[3] | | |
| ***No Change in Treatment*** | | | | | |
| Rigacci et al (2015) | After 2 cycles of ABVD, after the end of the 2nd cycle nearest the first part of the 3rd cycle. | | X | | |
| Hutchings et al (2014) | Within the last 5 days of the 1st and/or 2nd CT courses. | X[4] | | | |
| Barnes et al (2011) | Performed after 2-4 cycles of ABVD. | | X[5] | X | |

Abbreviations: ABVD, adriamycin, bleomycin, vinblastine, dacarbazine; AVG, doxorubicin, vinblastine, gemcitabine; BCNU, high-dose carmustine; BEACOPP, bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, prednisone; C, cycle; CT, chemotherapy; d, days; PET, positron emission tomography; VABEM, vinblastine, doxorubicin, BCNU, etoposide, methylprednisone; wks, weeks.

1 Change in treatment defined as changes to chemotherapy or radiotherapy, as reported in original study, treatment changes based on interim PET scan result.

2 Majority of cases (C2: 80 vs. C3: 9 patients).

3 Majority of cases (C2: 81 vs. C1: 15 patients).

4 All patients had an interim PET performed after cycle 1. A subset of the study population (70.6%) had an interim PET scan after both cycle 1 and cycle 2, based on pre-specified criteria.

5 A majority of patients had interim PET after cycle 2B (43%) or cycle 3B (47%). Only 6% of patients had interim PET after cycle 3A and only 4% of patients had interim PET after cycle 4A.

*Note: Kostakoglu (2012) and Straus (2011) had similar interim PET schedules, therefore only one study is reported here.

**Note: Le Roux (2011) is listed twice as its study population is evenly split between stages I-II (50%) and stages III-IV (50%) patients.

**Table 2b. Sequence of Interim PET in Stages III-IV (Advanced) Hodgkin Lymphoma (N=5 Studies)**

| Author (Year) | Description of Interim PET Timing | Cycle | | |
|---|---|---|---|---|
| | | C2 | C3 | C4 |
| **Obervational Studies** | | | | |
| ***Change in Treatment*** | | | | |
| Gallamini et al (2011) | After the 2nd ABVD cycle, a few days before the 3rd cycle. | X | | |
| Le Roux et al (2011)** | After 4 cycles of ABVD. | | | X |
| ***No Change in Treatment*** | | | | |
| Rossi et al (2014) | Performed after 2 cycles of anthracycline-based CT for all pts. | X | | |
| Markova et al (2012) | Performed after 4 cycles of BEACOPP (as close as possible to the 5th cycle). | | | X |
| Markova et al (2009) | After 4 cycles of BEACOPP, as close as possible to the 5th cycle of CT. | | | X |

Abbreviations: ABVD, adriamycin, bleomycin, vinblastine, dacarbazine; BEACOPP, bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, prednisone; C, cycle; CT, chemotherapy; PET, positron emission tomography; pts, patients.
**Note: Le Roux (2011) is listed twice as its study population is evenly split between stages I-II (50%) and stages III-IV (50%) patients.

**Table 3. Sequence of Final PET in Hodgkin Lymphoma**

| Author (Year) | Cycle | Description of Final PET Timing |
|---|---|---|
| **Studies with Details (N=10)** | | |
| Rigacci et al (2015) | 4 | At the end of 4 cycles of ABVD including RT, never performed earlier than 90 d after IFRT. |
| Kajary et al (2014) | - | CT alone: min 24 d after the last day of CT (betw. 24-76 d, m: 35 d).<br>CT+RT: at least 8 wks after RT for 38 pts and betw. 5-8 wks for 8 pts (overall betw. 5-35 wks, m: 12 wks). |
| Markova et al (2012) | 6-8 | Within 2-6 wks of the last application of CT. |
| Gallamini et al (2011) | - | No less than 3 mo. after the end of CT, or consolidation RT. |
| Straus et al (2011) | 6 | 1-2 wks after completion of cycle 6 of AVG. |
| Markova et al (2009) | 6-8 | Within 2-6 wks of the last CT application. |
| Advani et al (2007) | - | 1-2 wks after CT, and post-RT was at least 2 mo. after completion of RT. |
| Zinzani et al (2004) | - | At least 1 mo. after the end of CT and 3 mo. after any RT. |
| Guay et al (2003) | - | At the completion of CT, MOPP or ABVD (17 pts had PET after RT, 14 pts had PET prior to RT). Median time between last course of CT and PET was 58 d (~8 wks). |
| Hueltenschmidt et al (2001) | - | Within 4-6 wks after completion of primary therapy. |
| **Studies without Details (N=4)** | | |
| Iltis (2015)[1] | - | End-of-treatment |
| Barnes (2011) | - | |
| Luminari et al (2011) | - | |
| Dann et al (2010) | - | |

Abbreviations: ABVD, adriamycin, bleomycin, vinblastine, dacarbazine; AVG, doxorubicin, vinblastine, gemcitabine; CT, chemotherapy; d, days; IFRT, involved field radiation therapy; m, median; mo., months; MOPP, mustargen, viscristine (oncovin), procarbazine, prednisone; PET, positron emission tomography; pts, patients; RT, radiotherapy; wks, weeks.

1 Among responders. After cycle 2 of salvage therapy in non-responders.

**Table 4a. Details of PET Interpretation and Scoring for Studies Using Established Reporting Criteria**

| Author (Year) | Details |
|---|---|
| ***Deauville (Five-Point) Criteria (N=9 Studies)*** | |
| Radford et al (2015) | Negative: 1-2, positive: 3-5 |
| Rigacci et al (2015) | Negative: 1-3, positive: 4-5 |
| Simontacchi et al (2015)* | Negative: 1-2, positive: ≥ 3; negative: 1-3, positive: ≥ 4 |
| Hutchings et al (2014) | Negative: 1-3, positive: 4-5 |
| Kajary et al (2014)* | The threshold for positivity is MBP activity (score = 3); the threshold for positivity is liver activity (score = 4); negative: 1-4, positive: 5 |
| Rossi et al (2014) | Negative: 1-3, positive: 4-5 |
| Kostakoglu et al (2012) | Negative: 1-3, positive: 4-5 |
| Gallamini et al (2011) | Negative: 1-3, positive: 4-5 |
| Le Roux et al (2011) | Negative: 1-4, positive: 5 |
| ***International Harmonization Project Criteria (IHP) (N=7 Studies**)*** | |
| Iltis et al (2015) | Negative: based on minimal residual uptake, defined as slightly and diffusely increased FDG uptake at the site of moderate-sized or large residual masses (i.e. ≥ 2 cm in diameter), regardless of location, w. intensity greater than or equal to that of the MBP structures. Positive: the activity of MBP was used as the reference background activity for residual masses ≥ 2 cm at the widest transverse diameter, regardless of location. A smaller residual mass or a normal-sized lymph node (i.e. ≤ 1 × 1 cm in diameter) was considered positive if its activity was above that of the surrounding background. |
| Kajary et al (2014) | Positive: if FDG uptake is more intense than the MBP in lesions > 2 cm and more than background uptake in lesions < 2 cm that are consistent with involvement on CT. |
| Kostakoglu et al (2012)[1] and Straus et al (2011)** | Positive: focal or diffuse FDG uptake above background in a location incompatible with normal anatomy or physiology, w/o. a specific standardized uptake value cut-off. |
| Markova et al (2012) | Negative: a mild and diffusely increased uptake at the site of the residual mass w. an intensity lower than or equal to the MBP. |
| Le Roux et al (2011) | MBP is the reference background to determine positivity for a residual mass ≥ 2 cm in greatest transverse diameter, regardless of its location. For, smaller residual lesions, positive if activity is above that of surrounding background. |
| Dann et al (2010) | Compared to MBP, negative: 0-2, positive: 3-4 |
| Markova et al (2009) | Negative: a mild and diffusely increased uptake at the site of the residual mass w. an intensity lower than or equal to the MBP. Positive: if focal or diffuse uptake was seen above background in a location incompatible with normal anatomy or physiology, w/o. a specific standardized uptake cut-off value. |
| ***Gallamini Criteria (N=3 Studies)*** | |
| Kajary et al (2014) | Regardless of size, FDG activity greater relative to MBP is positive. |
| Le Roux et al (2011) | Irrespective of size, positivity only in the presence of focal uptake outside the physiological uptake areas with clearly increased activity relative to the MBP. |
| Dann et al (2010) | Compared to liver, negative: 0-2, positive: 3-4 |

Abbreviations: CT, computed tomography; FDG, 2-[18F]fluoro-2-deoxy-D-glucose; IHP, International Harmonization Project; MBP, mediastinal blood pool; PET, positron emission tomography; w, with; w/o, without.

1 As specified in the study, taken from IHP and Cheson et al (2007).

*Note: indicating studies that used more than one interpretation.

**Note: Kostakoglu (2012) and Straus (2011) represent the same study in this regard, and only counted once.

**Table 4b. Details of PET Interpretation and Scoring for Studies Using Other Reporting Criteria**

| Author (Year) | Details |
|---|---|
| ***Other criteria (N=9 Studies)*** | |
| Rigacci et al (2015) | Negative: absence of increased FDG uptake outside the physiological sites of radionucleotide concentration. Positive: presence of focal concentration of FDG outside the areas of physiological uptake with a value increased relative to background. |
| Barnes et al (2011) | Negative: 0-1, positive: 2-4 |
| Le Roux et al (2011) | Negative: no FDG uptake above local background in sites previously involved. Positive: all other findings were considered positive, including faint residual uptake. |
| Luminari et al (2011) | Scans were recoded as positive or negative if the presence/absence was clearly indicated in the report. Indefinite wording were considered inconclusive and not further analyzed. |
| Dann et al (2010)* | Negative: no abnormal uptake. Positive: any focus of abnormal uptake (not related to physiological or benign tracer uptake; compared with baseline, negative: 0-2, positive: 3-4. |
| Advani et al (2007) | Negative or positive for residual disease based on visual analysis. |
| Zinzani et al (2004) | Areas of focal uptake were interpreted as positive for lymphoma unless they were at the sites of known accumulation. |
| Guay et al (2003) | All foci of elevated FDG uptake not explainable by physiologic uptake represented viable lymphoma. Scans were classified as positive or negative. |
| Hueltenschmidt et al (2001) | Any focus of FDG uptake exceeding the normal FDG uptake in the respective area was considered to represent lymphoma involvement. Equivocal FDG-PET readings were classified as negative. |

Abbreviations: CT, computed tomography; FDG, 2-[18F]fluoro-2-deoxy-D-glucose; IHP, International Harmonization Project; MBP, mediastinal blood pool; PET, positron emission tomography; w/o, without.
*Note: indicating studies that used more than one interpretation.

**Table 5. Deauville (Five-Point) Criteria Definition of Positivity (N=9 Studies)**

| | | Score of 3 | |
|---|---|---|---|
| **Author (Year)** | **Overall PET Scoring** | **Negative** | **Positive** |
| Radford et al (2015) | Negative: 1-2, positive: 3-5 | - | ✓ |
| Rigacci et al (2015) | Negative: 1-3, positive: 4-5 | ✓ | - |
| Simontacchi et al (2015)* | Negative: 1-2, positive: ≥3 | - | ✓ |
| | Negative: 1-3, positive: ≥4 | ✓ | - |
| Hutchings et al (2014) | Negative: 1-3, positive: 4-5 | ✓ | - |
| Kajary et al (2014)* | The threshold for positivity is MBP activity (score=3) | - | ✓ |
| | The threshold for positivity is liver activity (score=4) | ✓ | - |
| | Negative: 1-4, positive: 5 | ✓ | - |
| Rossi et al (2014) | Negative: 1-3, positive: 4-5 | ✓ | - |
| Kostakoglu et al (2012) | Negative: 1-3, positive: 4-5 | ✓ | - |
| Gallamini et al (2011) | Negative: 1-3, positive: 4-5 | ✓ | - |
| Le Roux et al (2011) | Negative: 1-4, positive: 5 | ✓ | - |

Abbreviations: MBP, mediastinal blood pool; PET, positron emission tomography.
*Note: indicating studies that used more than one interpretation.

Deauville five-point scale as follows (Cheson et al., 2014; Barrington et al., 2014):
1 = no uptake.
2 = uptake ≤ mediastinum.
3 = uptake > mediastinum ≤ liver.
4 = uptake moderately higher than liver.
5 = uptake markedly higher than liver and/or new lesions.
X = new areas of uptake unlikely to be related to lymphoma.

**Table 6a. Outcome: Performance Metrics of Interim PET Use (N=8 Studies)**

| Author (Year) and Criteria | % +ve | Details of PET Use | Performance Metrics [% ± (95% CI)][1] | | | |
|---|---|---|---|---|---|---|
| | | | SN (%) | SP (%) | PPV (%) | NPV (%) |
| *Rigacci et al (2015) | | | | | | |
| Other | 14.6 | After cycle 2 of ABVD nearest the first part of cycle 3 (m: 12 d, R: 9-16 d) | 65.5 | 92.0 | 53.0 | 95.0 |
| Deauville (+ve ≥4) | 10.2 | | 68.0 | 97.0 | 73.0 | 96.0 |
| Hutchings et al (2014) | | | | | | |
| Deauville (+ve ≥4) | 30.3 | Performed after cycle 1 | - | - | 44.4 | 96.8 |
| Deauville (+ve ≥4) | 14.6 | Performed after cycle 2 | - | - | 61.5 | 92.1 |
| Rossi et al (2014) | | | | | | |
| Deauville (+ve ≥4) | 22.0 | Performed after 2 cycles of anthracycline-based CT for all pts | 46.0 | 84.0 | 46.0 | 85.0 |
| *Kostakoglu et al (2012) | | | | | | |
| IHP (PET, PET/CT) | 27.3 | After 2 cycles of AVG, performed 0-6 d prior to cycle 3 | 52.4 (30.0-74.0) | 80.6 (60.0-89.0) | 45.8 (26.0-67.0) | 84.4 (73.0-92.0) |
| Deauville (PET, PET/CT) (+ve ≥4) | 18.2 | | 38.1 (18.0-62.0) | 88.1 (78.0-95.0) | 50.0 (25.0-75.0) | 81.9 (71.0-90.0) |
| Deauville (PET/CT) (+ve ≥4) | 17.6 | | 41.2 (18.0-67.0) | 89.5 (78.5-96.0) | 53.8 (25.0-81.0) | 83.6 (72.0-92.0) |
| **Markova et al (2012), IHP | 26.1 | Performed after 4 cycles of BEACOPP (as close as possible to the 5th cycle) | - | - | - | 98.0 (94.0-100.0) |
| *Le Roux et al (2011) | | | | | | |
| Other | 34.4 | After 4 cycles of ABVD | - | - | 16.0 | 95.0 |
| IHP | 28.9 | | - | - | 19.0 | 95.0 |
| Gallamini | 22.2 | | - | - | 25.0 | 95.0 |
| Deauville (+ve=5) | 12.2 | | - | - | 45.0 | 96.0 |
| *Dann et al (2010) | | | | | | |
| Other | 25.0 | After the 1st or 2nd cycle of BEACOPP | 55.0 (23.0-88.0) | 78.0 (69.0-87.0) | 21.0 (5.0-37.0) | 94.0 (88.0-99.0) |
| IHP | 21.9 | | 44.0 (12.0-76.0) | 80.0 (72.0-88.0) | 19.0 (2.0-36.0) | 93.0 (87.0-99.0) |
| Gallamini | 16.7 | | 33.0 (2.0-64.0) | 85.0 (77.0-92.0) | 19.0 (0-38.0) | 92.0 (86.0-98.0) |
| Other | 6.3 | | 33.0 (2.0-64.0) | 96.0 (93.0-100.0) | 50.0 (10.0-90.0) | 93.0 (88.0-98.0) |
| Markova et al (2009), IHP | 28.0 | After 4 cycles of BEACOPP, as close as possible to the 5th cycle of CT | - | - | 14.0 (12.0-16.0) | 97.0 (94.0-100.0) |
| RANGE: | | | 33.0 to 68.0 | 78.0 to 97.0 | 14.0 to 73.0 | 81.9 to 98.0 |

Abbreviations: ABVD, adriamycin, bleomycin, vinblastine, dacarbazine; AVG, doxorubicin, vinblastine, gemcitabine; BEACOPP, bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, prednisone; CI, confidence interval; CT, computed tomography; d, days; IHP, International Harmonization Project; m, median; NPV, negative predictive value; PET, positron emission tomography; PPV, positive predictive value; pts, patients; R, range; SN, sensitivity; SP, specificity.

1 All studies used clinical follow-up with or without pathology confirmation as the reference standard.

*Note: indicating studies that used more than one interpretation criteria.

**Note: indicating studies that included both interim PET and final PET.

**Table 6b. Outcome: Performance Metrics of Final PET Use (N=4 Studies)**

| Author (Year) and Criteria | % +ve | Details of PET Use | Performance Metrics [% ± (95% CI)][1] SN (%) | SP (%) | PPV (%) | NPV (%) |
|---|---|---|---|---|---|---|
| *Kajary et al (2014) | | | | | | |
| IHP | 24.2 | For pts that had ABVD alone, PET was | 90.9 | 89.1 | 62.5 | 98.0 |
| Deauville (MBP) | 24.2 | performed at a minimum of 24 d after the | 90.9 | 89.1 | 62.5 | 98.0 |
| Deauville (Liver) | 21.2 | last day of CT (betw. 24-76 d, m: 35 d). For | 90.9 | 92.7 | 71.4 | 98.1 |
| Deauville (+ve=5) | 13.6 | pts that had CT plus RT, PET was performed | 72.7 | 98.2 | 88.9 | 94.7 |
| Gallamini | 19.7 | at least 8 wks after RT for 38 pts and betw. 5-8 wks for 8 pts (m: 12 wks, R: 5-35 wks) | 81.8 | 92.7 | 69.2 | 96.2 |
| **Markova et al (2012), IHP | 13.4 | After 6/8 cycles (within 2-6 wks of the last application of CT) | - | - | - | 95.0 (87.0-100.0) |
| Straus et al (2011), IHP | 18.8 | 1-2 wks after 6 cycles of AVG | 55.0 (31.5-76.9) | 93.3 (83.8-98.2) | 73.3 (44.9-92.2) | 86.1 (75.3-93.5) |
| Guay et al (2003), Other | 25.0 | After the completion of therapy | 79.0 | 97.0 | 92.0 | 92.0 |
| RANGE: | | | 55.0 to 90.9 | 89.1 to 98.2 | 62.5 to 92.0 | 86.1 to 98.1 |

Abbreviations: ABVD, adriamycin, bleomycin, vinblastine, dacarbazine; btw., between; CI, confidence interval; CT, computed tomography; d, days; IHP, International Harmonization Project; m, median; MBP, mediastinal blood pool; NPV, negative predictive value; PET, positron emission tomography; PPV, positive predictive value; pts, patients; R, range; RT, radiotherapy; SN, sensitivity; SP, specificity; wks, weeks.
1 All studies used clinical follow-up with or without pathology confirmation as the reference standard.
*Note: indicating studies that used more than one interpretation criteria.
**Note: indicating studies that included both interim PET and final PET.

**Table 7. Clinical Activities Following a Positive End-of-Treatment PET Scan (N=9 Studies)**

| Author (Year) | Clinical Activities | | | Details |
|---|---|---|---|---|
| | Repeat PET Scan | Biopsy | Treatment Change | |
| Iltis et al (2015) | - | - | ✓ | After salvage therapy, another PET scan was done and PET-positive pts after 2 cycles of salvage therapy received third-line therapy until CR and were auto- and/or allografted. (After salvage therapy, PET-negative pts received BCNU, BEAM for autografting). |
| Kajary et al (2014) | - | ✓ | - | A PET-positive scan in a pt who presented any evidence of disease, progression or relapse was confirmed as progression on CT and/or cytology or histology and/or clinical symptoms. |
| Markova et al (2012) | - | - | ✓ | Local RT (30 Gy) restricted to pts who had PR w. residual mass ≥ 2.5 cm after CT and PET(+) at the completion of 6-8 cycles BEACOPP therapy. |
| Straus et al (2011) | - | ✓ | - | Residual disease (after 6 cycles of CT) was assessed among PET(+) patients by biopsy, if clinically feasible (or by following until relapse). |
| Dann et al (2010) | - | ✓ | - | If FDG uptake was present in only a single site at the end of therapy, major attempts were made to obtain tissue specimens for histological examination prior to taking a decision about therapy failure. |
| Markova et al (2009) | - | - | ✓ | Local RT (30 Gy) was restricted to those pts who had a partial remission w. residual mass ≥ 2.5 cm after CT and who were F-PET(+). |
| Advani et al (2007) | - | ✓ | ✓ | After completion of RT, suspected relapses were confirmed in all cases with histologic dx, and if confirmed, pts were treated with high-dose therapy with autologous stem-cell support. |
| Zinzani et al (2004) | - | ✓ | - | Among the 5 pts who were CT(-)/PET(+) after treatment, all were submitted to a lymph node biopsy. |
| Hueltenschmidt et al (2001) | ✓ | ✓ | - | Abnormal PET foci were judged as true-positives/false-positive for lymphoma involvement, provided this finding was confirmed/excluded by biopsy and/or the further course of the disease. There were 5/51 (9.8%) pts that had a subsequent scan. |

Abbreviations: BEACOPP, bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, prednisone; BEAM, etoposide, cytosine, arabinosid, melphalan; BCNU, high-dose carmustine; CR, complete response; CT, computed tomography; dx, diagnosis; F-PET, final PET; FDG, 2-[18F]fluoro-2-deoxy-D-glucose; GY, gray; PET, positron emission tomography; PR, partial response; pts, patients; RT, radiotherapy; w, with; w/o, without.

**Table 8. Results of Biopsies Following a Positive End-of-Treatment PET Scan (N=6 Studies)**

| Author (Year) | Results |
|---|---|
| **Kajary et al (2014)** | Of 65 patients who were followed for a minimum of 20 months of follow-up, 10/65 patients experienced therapy failure (any evidence of disease, progression, or relapse) during follow-up care, which was confirmed with cytology or histology in only 4/10 patients. The remaining 6/10 patients showed obvious progression on CT, and 3/6 patients also had serious clinical symptoms, therefore the treating physician decided to start additional therapy without histological confirmation. |
| **Straus et al (2011)** | Of 23 relapses, 18 were documented by biopsy [15 in primary sites, 7 in both primary and new sites, 1 in a new site only]. |
| **Dann et al (2010)** | There were 2/3 patients who were treated initially with escalated BEACOPP and had a positive interim PET scan showing a single residual mediastinal mass and therefore received four further cycles of escalated BEACOPP followed by radiation therapy. Further, these two patients had a positive end-of-therapy scan taken three months after radiation therapy. Biopsies from these two patients appeared negative and they had no evidence of disease progression. |
| **Advani et al (2007)** | Among 81 patients treated with combined-modality therapy including the Stanford V regimen, 7/81 patients who had an end-of-therapy PET scan experienced a biopsy-confirmed relapse (median time of relapse: 14 months for positive end-of-therapy PET scans vs. 16 months negative end-of-therapy PET scans). |
| **Zinzani et al (2004)** | Among the five patients that underwent lymph node biopsy following an end-of-therapy positive PET scan, histological confirmation of HL was made in one patient, and this patient also relapsed. |
| **Hueltenschmidt et al (2001)** | Among 51 patients who had an end-of-therapy PET scan, including repeat scans in some patients, there were 63 end-of-therapy PET scans performed, of which three scans were determined to be incorrect (patient 1: thymus uptake was revealed by histology to be thymus hyperplasia; patient 2: inflammatory lung process was revealed by histology to be thyroid adenoma; patient 3: negative CT scan was revealed by histology to be active disease and the patient was scheduled for involved field irradiation). |

Abbreviations: BEACOPP, bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, prednisone; CT, chemotherapy; HL, Hodgkin lymphoma; PET, positron emission tomography.

## Appendix 7: Quality of Evidence

**Table 1. Quality Assessment by QUADAS-2[a] (N=11 studies)**

| | Risk of Bias* | | | | Applicability Concerns | | |
|---|---|---|---|---|---|---|---|
| Study | Patient Selection | Index Test | Reference Standard | Flow and Timing | Patient Selection | Index Test | Reference Standard |
| Rigacci et al (2015) | ☺ | ☺/☹[1] | ☺/☹[1,2] | ☺/ ? | ☺ | ☺ | ☺ |
| Hutchings et al (2014) | ☹ | ☺ | ☺/☹[1,3,4] | ☺/ ? | ☺ | ☺ | ☺ |
| Kajary et al (2014) | ☺ | ☺ | ☺/☹[1,4,5] | ☺/ ? | ☺ | ☺ | ☺ |
| Rossi et al (2014) | ☺ | ☺ | ☺/☹[1,4] | ☹/ ?[6] | ☺ | ☺ | ☺ |
| Kostakoglu et al (2012) | ☹ | ☺ | ☺/☹[1,2] | ☹/ ?[7] | ☺ | ☺ | ☺ |
| Markova et al (2012) | ☺ | ☺/☹[1] | ☺/☹[1,4] | ☺/ ? | ☺ | ☺ | ☺ |
| Le Roux et al (2011) | ☹ | ☺[8] | ☺/☹[1,4] | ☺/ ? | ☺ | ☺ | ☺ |
| Straus et al (2011) | ☹ | ☺ | ☺/☹[1,2] | ☹/ ?[7] | ☺ | ☺ | ☺ |
| Dann et al (2010) | ☹ | ☺/☹[1,9] | ☺/☹[1,3,4] | ☺/ ? | ☺ | ☺ | ☺ |
| Markova et al (2009) | ☹[10] | ☺/☹[1] | ☺/☹[1,2] | ☺/ ? | ☺ | ☺ | ☺ |
| Guay et al (2003) | ☺ | ☺/☹[1] | ☺/☹[1,2,3] | ☺/ ? | ☺ | ☺ | ☺ |
| Overall: | ☺/☹ | ☺/☹ | ☹ | ☺/ ? | ☺ | ☺ | ☺ |

Abbreviations: FN, false negative; FP, false positive; GHSG, German Hodgkin Study Group; IWG, International Working Group; PET, positron emission tomography; TN, true negative; TP, true positive.

[a] ☺Low risk (No limitations), ☹High risk (Limitations), ? Unclear risk

1 Blinding of reader to treatment outcomes or index PET scans not reported.
2 Methods not clearly described e.g., definition of a 'disease-positive patient'.
3 Kajary (2014) and Dann (2010) confirmed treatment outcomes. Hutchings (2014) and Guay (2003) included biopsy-confirmed relapses.
4 Studies provided some sort of definition of TP, TN, FP, FP for their calculations.
5 Criteria used not specified, although definitions of TP, TN, FP, and FN were reported.
6 Treatment strategy included anthracycline-based CT not otherwise specified.
7 Treatment strategy included no radiation.
8 Prospective PET scan interpretation in light of baseline scan; retrospective PET scan interpretation was blinded to treatment outcomes.
9 Blinding of reader to treatment outcomes was not reported across all PET interpretation criteria.
10 Although patients were part of the GHSG study population, randomization was not maintained for the current study population.

*Specific risk of bias criteria used:
**Patient selection:** No limitations (☺): consecutive sampling in a given time period and population, random sample with method provided; Limitation (☹): patient selection not reported or not specified.
**Index test:** No limitations if answered yes; limitations if answer is no. For, (a) blinding of reader, (b) used a pre-specified criteria, (c) threshold or definition of positivity stated.
**Reference standard:** No limitations if answered yes; limitations if answer is no. For, (a) blinding of reader, (b) provided a description of the reference standard being clinical follow-up with or without pathology confirmation e.g., laboratory, PET imaging, non-PET imaging, or symptoms or (c) provided the response criteria for determining treatment failure.
**Flow and timing:** No limitations if answered yes; limitations if answer is no. For, (a) where there appropriate interventions or any interventions between the index test and the reference standard, such as treatment provided or the specific timing of the index text, (b) length of follow-up, (c) did all patients receive the same reference standard.

## Appendix 8: Published Guidelines on Treatment Response Criteria

**Table 1. Treatment Response Criteria and Evaluation**

| Treatment Response Criteria | Treatment Response Categories | Treatment Response Evaluation |
|---|---|---|
| Cheson et al (2014)<br><br>**Lugano Classification** | • Complete<br>• Partial<br>• No response or stable disease<br>• Progressive disease | • PET/CT based treatment responses using the Deauville five-point scale for FDG-avid histologies<br>• CT-based treatment responses for non-avid histologies |
| Cheson et al (2007)<br><br>**International Harmonization Project** | • Complete<br>• Partial remission<br>• Stable disease<br>• Relapsed or progressive disease | • PET/CT and/or CT based treatment responses by visual analysis |
| Cheson et al (1999)<br><br>**International Working Group** | • Complete<br>• Complete (unconfirmed)<br>• Partial remission<br>• Relapse/progression | • Clinical, radiologic (i.e., CT scan) and pathologic criteria |

Abbreviations: CT, computed tomography; FDG, 2-[18F]fluoro-2-deoxy-D-glucose; PET, positron emission tomography.

## Appendix 9: Sub-Analysis

**Table 1. Details of Therapy Management for Included Studies (N=20 Studies)**

| Author (Year) | Chemotherapy and Radiation Therapy Description | Radiation | Risk-Adapted Therapy |
|---|---|---|---|
| ***Clinical Trials*** | | | |
| Radford et al (2015) | Three cycles of standard ABVD treatment. I-PET(+) pts received a 4th ABVD cycle and IFRT. I-PET(-) pts were randomly assigned in a 1:1 ratio to receive 30 Gy IFRT or no further treatment. Second line treatment among I-PET(-) pts for recurrent disease included high-dose CT and subsequent autologous stem-cell transplantation. Second line treatment among I-PET(+) pts included transplantation (autologous [7 pts], allogeneic [1 pt]). | ✓ | ✓ |
| Kostakoglu et al (2012) | Each AVG cycle was administered by IV on days 1 and 15. Cycles were repeated every 28 days for a total of 6 cycles. | No | No |
| Straus et al (2011) | Each AVE cycle was administered by IV on days 1 and 15. Cycles were repeated every 28 days for a total of 6 cycles. | No | No |
| ***Observational Studies*** | | | |
| Iltis et al (2015) | Using the prognostic scoring system, pts categorized as early stage received 3 cycles of ABVD, pts categorized as intermediate risk received 6 cycles of ABVD or 3 cycles of VABEM, pts categorized as advanced stage received 8 cycles of ABVD or 3 cycles of VABEM. I-PET(+) pts received salvage therapy consisting of VABEM (for pts who were treated with ABVD as first-line CT) and platinum-containing regime for pts who were treated with VABEM as first-line CT. I-PET(-) pts received the originally planned therapy. After salvage therapy, another PET scan was done and PET-positive pts after 2 cycles of salvage therapy received third-line therapy until CR and were auto- and/or allografted. After salvage therapy, PET-negative pts received BCNU, BEAM for autografting. IFRT (20 or 30 Gy) was performed after the completion of CT in pts of CR on intial sites for early stage or bulky disease. There were 24 pts that received RT in CR after the completion of first-line CT. | ✓ | ✓ |
| Rigacci et al (2015) | All pts treated with 4 cycles ABVD followed by IFRT (30-32 Gy). ABVD was provided as standard therapy, NOS. | ✓ | No |
| Simontacchi et al (2015) | All pts received first 2 cycles of ABVD, with subsequenct cycles and modifications determined by treating hemato-oncologist. All pts received RT to initial sites of disease (involved field or involved nodal and site). All pts received RT after CT, within 4-6 weeks. | ✓ | ✓ |
| Hutchings et al (2014) | Early stage pts received 2-4 cycles of ABVD, days 1 and 15 of 28-day cycle, followed by RT to initially involved LN or nodal areas, or with 6 cycles of ABVD. Advanced stage pts received 6-8 cycles of ABVD ± RT. Five pts with advanced stage + adverse risk factors received 8 cycles of escBEACOPP. | ✓ | No |
| Kajary et al (2014) | CT alone was given to 20 pts of either 4 or 6 cycles of ABVD. CT + RT was given to 46 pts (4 or 6 cycles of ABVD). | ✓ | No |
| Rossi et al (2014) | Stage I or II: 4-6 cycles of anthracycline-based CT followed by 20-36 Gy IFRT. Stage III or IV: 8 cycles of anthracycline-based CT. Treatment was not changed on the basis of I-PET results. | ✓ | No |
| Markova et al (2012) | Treated according to GHSG trial. Arm A: 8 cycles of BEACOPP (escalated), Arm B: 6 cycles of BEACOPP (escalated), Arm C: 8 cycles of time-condensed BEACOPP14 (baseline). Local RT was restricted to those pts who had partial remission w. residual mass ≥ 2.5 cm after CT and who were PET positive after the completion of 6-8 cycles of BEACOPP. | ✓ | No |
| Barnes et al (2011) | Most patients received 6 cycles of ABVD (43%), or 6 cycles of ABVD plus IFRT (30%), or 4 cycles of ABVD plus IFRT (26%), or 4 cycles of ABVD (1%). | ✓ | No |

| | | | |
|---|---|---|---|
| Gallamini et al (2011) | Treatment with 2 cycles of ABVD. Then, I-PET(+) pts had CT intensification with BEACOPP, w. no changes in planned RT schedule. I-PET(-) pts continued with ABVD for a total of 6 cycles. For all pts, consolidation RT was given after CT (up to total dose of 36 Gy). | ✓ | ✓ |
| Le Roux et al (2011) | I-PET(+) but with complete response on CT: had IFRT for early favourable HL or an additional 4 cycles of ABVD for early unfavourable HL. The remaining I-PET(+) pts had escalation therapy by ASCT conditioned by BEAM high-dose CT. I-PET(-) and no progression on CT: had IFRT for early favourable HL or an additional 4 cycles of ABVD for early unfavourable HL. | ✓ | ✓ |
| Luminari et al (2011) | Treatment included ABVD/ABVD-like (n=116, 85%), intensified CT e.g. BEACOPP, COPP/EBV/CAD (n=11, 8%), CT w/o ADM e.g. VBM, MOPP (n=6, 4%), RT/palliative (n=3, 3%). RT at the end of CT was given to 72 pts. | ✓ | ✓ |
| Dann et al (2010) | Risk-adapted BEACOPP. If IPS ≤ 2, then pts treated with 2 initial cycles of std BEACOPP. If IPS ≥ 3, then pts treated with 2 initial cycles of escalated BEACOPP. I-PET(+) pts initially treated with std or escalated BEACOPP were given 4 cycles of escalated BEACOPP. I-PET(-) pts received 4 cycles of std BEACOPP. RT was given to pts with bulky mediastinal mass, I-PET(+), or both, and early disease (180 Gy). | ✓ | ✓ |
| Markova et al (2009) | 19 pts received 8 cycles BEACOPP (escalated), 16 pts received 6 cycles BEACOPP (escalated), or 15 pts received 8 cycles of time-condensed BEACOPP14 (baseline). Local RT (30 Gy) was restricted to those pts who had a partial remission w. residual mass ≥ 2.5 cm after CT and who were F-PET(+). | ✓ | No |
| Advani et al (2007) | Pts were treated with the Stanford V regime, as previously described, w. 8 or 12 weeks of CT depending on if pts had favourable stage I/II asymptomatic (nonbulky mediastinal) (8 weeks, 43.2% of pts) or those pts w. bulky mediastinal disease or stage III/IV disease (12 weeks, 56.8% of pts). RT at 30 Gy was given to involved sites in favourable stage I/II pts and 36 Gy to sites ≥ 5 cm and macroscopic splenic disease in all other pts. Some pts (6 pts) w. favourable stage I/II received 20 Gy to involved sites as part of a prior study. | ✓ | No[1] |
| Zinzani et al (2004) | HL patients were treated with 4 or 6 cycles of ABVD. RT was performed at the clinician's discretion (30-36 Gy). | ✓ | No |
| Guay et al (2003) | Chemotherapy alone was given to 17 pts (mean number of cycles was 7 cycles) with either MOPP or ABVD. There were 31 pts that had CT plus RT (17 pts had RT before PET, whereas 14 pts had RT after PET). | ✓ | No |
| Hueltenschmidt et al (2001) | Primary treatment before PET included CT for 48 pts (59.3%), radiotherapy for 7 pts (8.6%), and combined therapy (CT+RT) for 26 pts (32.1%). | ✓ | No |

Abbreviations: ABVD, adriamycin, bleomycin, vinblastine, dacarbazine; ADM, adriamycin; AVG, doxorubicin, vinblastine, gemcitabine; BEACOPP, bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, prednisone; BEAM, etoposide, cytosine, arabinosid, melphalan; BCNU, high-dose carmustine; CAD, lomustine, doxorubicin, vindesine; COPP, cyclophosphamide, vincristine, procarbazine, prednisone; CR, complete response; CT, chemotherapy; EBV, epirubicin, bleomycin, vinblastine; F-PET, final PET; GHSG, German Hodgkin Study Group; GY, gray; HL, Hodgkin lymphoma; I-PET, interim PET; IFRT, involved field radiotherpay; IV, intravenously; LN, lymph nodes; MOPP, chlormethnine, vincristine, procarbazine, prednisone or MOPP, mustargen, vincristine (oncovin), procarbazine, prednisone (Guay 2003); NOS, not otherwise specified; PET, positron emission tomography; pts, patients; RT, radiotherapy; VABEM, vinblastine, doxorubicin, BCNU, etoposide, methylprednisone; VBM, vinblastine; w, with.

1 Scan results after CT did not influence RT.