Washington State Health Care Authority

Positron Emission Tomography (PET) scans for lymphoma – re-review

Final evidence report

October 15, 2018

Health Technology Assessment Program (HTA)
Washington State Health Care Authority
PO Box 42712
Olympia, WA 98504-2712
(360) 725-5126
hca.wa.gov/hta
shtap@hca.wa.gov

Positron Emission Tomography for Lymphoma

– Re-review –

Provided by:

Aggregate Analytics, Inc.

Prepared by:

Andrea C. Skelly, PhD, MPH
Erika D. Brodt, BS
Naomi Schwartz, MPH
Dena J. Fischer, DDS, MSD, MS
Aaron J. R. Ferguson, BA
Shelby Kantner, BA

October 15, 2018

This technology assessment report is based on research conducted by a contracted technology assessment center, with updates as contracted by the Washington State Health Care Authority. This report is an independent assessment of the technology question(s) described based on accepted methodological principles. The findings and conclusions contained herein are those of the investigators and authors who are responsible for the content. These findings and conclusions may not necessarily represent the views of the HCA/Agency and thus, no statement in this report shall be construed as an official position or policy of the HCA/Agency.

The information in this assessment is intended to assist health care decision makers, clinicians, patients and policy makers in making sound evidence-based decisions that may improve the quality and cost-effectiveness of health care services. Information in this report is not a substitute for sound clinical judgment. Those making decisions regarding the provision of health care services should consider this report in a manner similar to any other medical reference, integrating the information with all other pertinent information to make decisions within the context of individual patient circumstances and resource availability.

TABLE OF CONTENTS

TABLES

FIGURES

Abbreviations

ABVD: adriamycin, bleomycin, vinblastine, dacarbazine
AE: adverse event
AlloSCT: allogenic stem cell transplantation
aNHL: aggressive non-Hodgkin lymphoma
AutoSCT: autologous stem cell transplantation
BEACOPP: bleomycin, etoposide, adriamycin, cyclophosphamide, vincristine, procarbazine, prednisone
BEAM: BCNU, etoposide, cytarabine, melphalan
CHOP: cyclophosphamide, adriamycin, vincristine, prednisone
CI: confidence interval
CT: computed tomography
DFS: disease-free survival
eBEACOPP: escalated doses of bleomycin, etoposide, adriamycin, cyclophosphamide, vincristine, procarbazine, prednisone
EFS: event-free survival
^{18}FDG: 18-fluorodeoxy-D-glucose
F/U: follow-up
HL: Hodgkin lymphoma
HR: hazards ratio
iNHL: indolent non-Hodgkin lumphoma
NC: not calculable
NHL: Non-Hodgkin lymphoma
NPV: negative predictive value
NR: not reported
NS: not statistically significant
OR: odds ratio
OS: overall survival
PET: positron emission tomography
PET/CT: positron emission tomography/computed tomography
PFS: progression-free survival
PPV: positive predictive value
R-BEAM: rituximab + BCNU, etoposide, cytarabine, melphalan
R-CHOP: rituximab + cyclophosphamide, adriamycin, vincristine, prednisone
RCT: randomized controlled trial
RD: risk difference
RFS: relapse-free survival
RoB: risk of bias
RR: risk ratio
RT: radiation therapy
SD: standard deviation

Executive Summary

Introduction

Lymphoma is a heterogeneous group of cancers that affect cells of the immune system, primarily those involved in the lymphatic system, a complex network of vessels, tissues and organs which carry a fluid called lymph throughout the body. The affected cells are a type of white blood cell called a lymphocyte, which account for 20% to 40% of the total number of white blood cells in adults, and play an integral role in the human immune response by recognizing and destroying infectious organisms and abnormal cells. Lymphoma most often starts in the lymph nodes but can easily spread to other areas if not treated such as the spleen, tonsils, thymus gland and bone marrow, and occasionally even spread to organs outside the lymphatic system. Lymphomas are divided into two major categories: Hodgkin lymphoma (HL, previously called Hodgkin's disease) and non-Hodgkin lymphoma (NHL, all other lymphomas). While HL and NHL occur in the same locations, may be associated with the same symptoms (e.g., swollen lymph nodes), and often have similar appearance on physical examination they can be easily differentiated histologically by the specific type of lymphocyte each involves.

Each type of lymphoma behaves, spreads, and responds to treatment differently, so an accurate diagnosis is essential to determining the appropriate treatment strategy, expected response to treatment, and monitoring for recurrence. Lymphoma is diagnosed based on physical exam, lymph node biopsy and blood tests. After diagnosis, patients undergoing typical management of lymphoma receive initial staging, treatment, restaging after treatment and subsequent surveillance or further treatment and restaging depending on responsiveness to treatment (generally determined by Computed Tomography (CT)-assessed size reduction of enlarged lymph nodes, extent of bone marrow involvement, immunohistochemistry, flow cytometry and findings on PET scans). Treatment options can include combinations of chemotherapy (including the possibility of higher doses when paired with stem cell transplantation), radiation therapy, targeted therapy (e.g. inhibitor drugs), immunotherapy (including monoclonal antibodies), and in rare cases, surgery. Treatment for indolent forms of NHL ranges from watchful monitoring to aggressive therapy. Treatments are based on type, disease stage, age, co-existent medical conditions and prognostic factors. All treatments have been associated with a broad range side effects depending on the therapies used, and may include fertility issues, damage to the thyroid, heart and lungs as well as increased risk for infection, stroke and secondary cancers.

Positron emission tomography (PET) is a type of nuclear medicine imaging that utilizes small amounts of radioactive materials called radiotracers to examine and measure physiological functions in the body. The more energy a group of cells needs, the more the radiotracer will build up in that location. For lymphoma, the radioactive particle most commonly used for PET is ^{18}fluorine, which binds to glucose to form fluorodeoxy-D-glucose (FDG). Most types of lymphoma are metabolically active and use more glucose compared with normal structures (i.e. are termed FDG avid). This results in greater uptake of the radioactive FDG creating a "hot" spot on the PET image. In general, ^{18}FDG-PET is not routinely used for initial diagnosis of lymphoma as histologic samples obtained via biopsy are required; however, ^{18}FDG-PET may assist in identification of the best place for biopsy (e.g. most metabolically active site, locations where biopsy may be difficult, previous non-diagnostic needle biopsy). ^{18}FDG-PET has become widely used as an imaging tool for staging of lymphoma after histological diagnosis, for interim evaluation (e.g.

restaging and evaluation of treatment response during the course of treatment), and for end-of-treatment assessment, although there is some debate regarding the value of interim PET based on the quality of evidence available and inconsistency regarding criteria for PET interpretation in the literature. Some guidelines recommend interim PET and it is generally performed at least once and has become a standard for assessment of treatment response for most lymphomas. Interim PET/CT imaging may facilitate the ability to discriminate between those for whom additional or intensified treatment may be important and those for whom additional therapy may not be necessary or some forms of therapy (e.g. radiation) may not be needed. This may permit optimization of therapy and maintenance of treatment efficacy while decreasing or avoiding treatment-related side-effects or sequelae. Escalation of treatment in appropriate patients may improve survival. De-escalation of treatment may reduce treatment-related toxicity and result in cost-savings. Recently, studies have emerged which explore the role of interim PET findings for adapting therapy with these goals in mind. To evaluate the value of PET/CT, studies comparing treatment given based on PET/CT results with treatment that would have been given without PET/CT information are of primary interest. Interim PET is not generally performed for indolent lymphoma. Relative to evaluation at the end of primary treatment, PET results may help inform discontinuation of ineffective treatment and consideration of salvage therapy in patients for whom there is a clinical suspicion of disease progression. PET is generally not used for routine surveillance following treatment completion, largely due to concerns regarding false-positive findings and lack of impact on patient outcomes.

There are data suggesting that ^{18}FDG-PET may be a predictor of prognosis when performed early during treatment. Across 13 studies summarized in a recent narrative review, interim PET (performed after 1 to 4 cycles of treatment) generally predicted better progression-free survival (PFS) in those with negative interim results (PFS range 71% to 95%) a lower PFS in those with positive results (range 13% to 100%)[41] depending on type and stage of lymphoma.

Today, most PET scans are performed on combination PET/CT scanners. The combined PET/CT scans provide images that pinpoint the anatomic location of abnormal metabolic activity within the body and provide more accurate staging and evaluation than the two scans performed separately. Most clinical guidelines recommend the use of PET in conjunction with CT, whether each test is done separately or via an integrated PET/CT scanner for initial staging and re-staging at critical points after treatment. The combination of PET with diffusion weighted MRI is an emerging technology for the evaluation of lymphoma. While PET/CT involves radiation exposure, the risk of adverse events specifically related to its use is generally considered to be low. In contrast, treatment-related adverse events are common and may be severe.

The value of a medical test relates to its ability identify persons for whom there are appropriate and effective treatments. In the absence of studies randomizing patients to receive PET/CT or no PET/CT and following patients through treatments to final clinical outcomes, the impact of PET on clinical decision making leading to effective treatment and outcomes can only be assessed indirectly. The outcomes observed are a function of PET/CT results in combination with the treatments received on the basis of those results. Thus, our approach considers the clinical outcomes of overall survival (OS) and progression free survival (PFS) in relation to the modifications of treatment received based on PET results, along with treatment related-toxicity. In other words, for this report the value of PET for assisting with decisions to escalate, de-escalate or continue therapy is indirectly assessed based upon the impact of the PET-adapted treatment.

Policy Context

This topic was originally reviewed in 2011. It is being re-reviewed in 2018 due to newly available published evidence.

Objectives

The aim of this report is to update the 2011 HTA on positron emission tomography (PET) for lymphoma by summarizing information on diagnostic accuracy (e.g., sensitivity, specificity, predictive values) for context and systematically reviewing, critically appraising and analyzing new research evidence evaluating the clinical effectiveness (i.e., the ability of PET to stage, and influence therapeutic decisions, clinical management and clinical outcomes), safety, differential efficacy and safety in subpopulations, and cost-effectiveness of PET for lymphoma in adult and pediatric patients. The combination of PET with diffusion weighted MRI is an emerging technology for the evaluation of lymphoma. Currently this combination is not widely used, so the focus of this report will be on PET/CT as it is the current standard of care. Evidence on PET/MRI will be included as appropriate.

Contextual Questions

In patients with histologically proven HL and NHL included in the report, what are the accuracy and reliability of 18FDG-PET alone or in combination with CT for initial staging, interim evaluation (including monitoring during treatment, evaluation of treatment response and prognosis), re-staging (e.g., at end of treatment or other times), end of treatment assessment, evaluation of disease recurrence and surveillance of patients in remission? Specifically, provide a summary of the:

- Sensitivity and specificity and prognostic value (positive and negative predictive values)
- Inter- and intra-rater reliability (reproducibility)

In addition, a brief summary of the diagnostic accuracy and use of PET/CT for initial diagnosis will be provided. Summaries of accuracy will be based on highest quality systematic reviews which critically appraise included studies. Contextual information on the combination of PET with MRI will be presented. Contextual questions are not systematically reviewed, and use a “best evidence” approach.

Key Questions

The focus of this portion of the report is on the clinical impact of 18FDG PET/CT as this is the current standard of care. Except where noted the term PET will be construed to mean PET/CT. Information related to the PET/MRI combination will be included if relevant. In patients with histologically proven HL or NHL undergoing PET for initial staging, interim evaluation (including monitoring during treatment, evaluation of treatment response and prognosis), re-staging (e.g. at end of treatment or other times), end of treatment assessment, evaluation of disease recurrence and surveillance of patients in remission:

1. What is the evidence of clinical effectiveness of ^{18}FDG imaging in combination with CT (PET/CT) results?
    a. How do ^{18}FDG PET/CT results impact therapeutic decisions or clinical management? Do test results lead to use of effective treatment strategies (e.g. including initial treatment following staging or treatment acceleration, deceleration or termination at interim imaging) compared with treatment strategies not using such test results?
    b. How do clinical outcomes (e.g. overall survival) differ based on PET/CT-related treatment decisions compared with decisions made in the absence of such test results?

   c. Does the use of ^{18}FDG PET for treatment decisions lead to reduction in treatment-related adverse events/sequelae in general compared with treatment decisions that do not involve PET/CT?
   d. Is there a reduction in the need for other tests?
   e. How do end of treatment ^{18}FDG PET/CT results impact clinical decision making compared with clinical decisions that do not involve PET/CT??
   f. How does surveillance ^{18}FDG PET/CT results impact clinical decision making compared with clinical decisions that do not involve PET/CT??

2. What is the safety profile of ^{18}FDG PET/CT for lymphoma?
   a. What adverse events are reported: type and frequency directly attributable to ^{18}PET/CT (mortality, major morbidity, radiation exposure, other)?

3. What is the evidence that ^{18}FDG PET/CT imaging in patients with known lymphoma has differential efficacy or safety in subpopulations? Including consideration of:
   a. Patient age, sex, characteristics or evidence-based patient selection criteria
   b. Type of scanning machine and software, reader training, and other operational factors
   c. Provider type, setting or other provider characteristics
   d. Health care system type, including worker's compensation, Medicaid, state employees

4. What is the evidence of short and long-term cost-effectiveness of ^{18}FDG PE/CT for patients with lymphoma compared with other imaging or clinical care not involving ^{18}FDG PET/CT?

Inclusion and exclusion criteria are summarized as follows and are detailed in the full report. Briefly, included studies met the following requirements with respect to participants, intervention, comparators, outcomes, and study design:

- **Population**: Adults, adolescents or children with biopsy-proven HL or NHL. Diagnoses of interest (to be evaluated separately) include HL or aggressive NHL, indolent NHL and other NHL with focus on most common types.
- **Interventions:** Focus is on positron emission tomography (PET) to measure glucose metabolism (^{18}FDG-PET) in addition to computed tomography (CT) (including combined PET/CT equipment). Combination of PET with MRI (including diffusion weighted MRI) will be included as appropriate.
- **Comparators:** Other imaging (CT alone, MRI, including diffusion weighted MRI); standard clinical protocols or standard prior tests/evaluations (including history and physical examination, laboratory studies, biopsy) that do not involve ^{18}FDG PET.
- **Outcomes:** Primary outcomes are 1) improvement in clinical outcomes based on PET/CT-derived clinical decision making (focus is on overall survival, progression- or event-free survival, morbidity and mortality, and 2) quality of life. (See full report for secondary and indirect outcomes). Safety outcomes are type and frequency of adverse events directly attributable to PET/CT (e.g., incidental findings, repeat/additional procedures, radiation exposure, other). Economic outcomes are cost-effectiveness (e.g., cost per improved outcome), cost-utility (e.g., cost per quality adjusted life year (QALY), incremental cost effectiveness ratio (ICER) outcomes.
- **Timing:** PET performed at any time after initial diagnosis including for initial staging, interim evaluation, end-of-treatment evaluation, evaluation of recurrence and surveillance.

- **Studies:** For efficacy and effectiveness, the focus will be on comparative PET/CT studies with the least potential for bias (i.e. RCTs). For safety, the focus will be on studies characterizing direct PET/CT harms (including incidental findings, repeat biopsy, radiation safety). For differential efficacy or effectiveness, only studies which stratify on patient or other characteristics and formally evaluate statistical interaction (effect modification) will be considered. For cost-effectiveness, only full, formal economic studies (i.e., cost-effectiveness, cost-utility, cost-minimization, and cost-benefit studies) will be included. Given that this is re-review, only studies published subsequent to the 2011 report (Search dates February 2011 to May 2018) will be considered.

Methods

The scope of this report and final key questions were refined based on input from clinical experts and public comments received on draft key questions. Clinical expert input was sought to confirm critical outcomes on which to focus.

The 2011 report used a rapid review methodology of selected literature sources and focused on the diagnostic accuracy of PET reported in prior systematic reviews. The focus of this re-review is on studies evaluating the clinical impact of PET/CT on the primary clinical outcomes of overall survival, progression-free survival, and treatment-related toxicity necessitating a broader search methodology. A formal, structured systematic search of the peer-reviewed literature was performed across a number of databases including PubMed to identify relevant peer reviewed literature as well as other sources (National Guideline Clearinghouse, Center for Reviews and Dissemination Database) to identify pertinent clinical guidelines and previously performed assessments.

For the systematic review portion of the report, studies were selected for inclusion based on pre-specified criteria detailed in the full report. All records were screened by two independent reviewers. Selection criteria included a focus on studies with the least potential for bias that were written in English and published in the peer-reviewed literature.

Included studies reporting on primary outcomes of interest were critically appraised independently by two reviewers evaluating the methodological quality, study limitations and potential for bias based on study design as well as factors which may bias studies. An overall Strength of Evidence (SoE) combines the appraisal of study limitations with consideration of the number of studies and the consistency across them, directness and precision of the findings to describe an overall confidence regarding the stability of estimates as further research is available. The SoE was assessed by two researchers following the principles for adapting GRADE (Grades of Recommendation Assessment, Development and Evaluation)[1] as outlined by the Agency for Healthcare Research and Quality (AHRQ).[10] The strength of evidence was based on the highest quality evidence available for the primary outcomes. Briefly, bodies of evidence consisting of RCTs were initially considered as High strength of evidence. The strength of evidence could be downgraded based on the limitations (i.e., risk of bias, consistency of effect, directness of outcome, precision of effect estimate, and reporting/publication bias). When assessing the SoE for studies performing subgroup analysis, we also considered whether the subgroup analysis was preplanned (*a priori*) and whether a test for homogeneity or interaction was done. There are also situations where nonrandomized studies could be upgraded if large magnitude of effect (strength of association) or a dose-response relationship is observed if all known confounders were adjusted for and there was no

downgrade for any of the domains. The final strength of evidence was assigned an overall grade of high, moderate, low, or insufficient, which are defined as follows:

- High - Very confident that effect size estimates lie close to the true effect for this outcome; there are few or no deficiencies in the body of evidence; we believe the findings are stable.
- Moderate – Moderately confident that effect size estimates lie close to the true effect for this outcome; some deficiencies in the body of evidence; we believe the findings are likely to be stable but some doubt remains.
- Low – Limited confidence that effect size estimates lie close to the true effect for this outcome; major or numerous deficiencies in the body of evidence; we believe that additional evidence is needed before concluding that findings are stable or that the estimate is close to the true effect.
- Insufficient – We have no evidence, are unable to estimate an effect or have no confidence in the effect estimate for this outcome; OR no available evidence or the body of evidence has unacceptable efficiencies precluding judgment.

We summarized evidence separately for Hodgkin lymphoma and non-Hodgkin lymphoma. Hodgkin lymphoma was further divided into early-stage, advanced-stage, and relapsed/refractory disease, and non-Hodgkin lymphoma into indolent and aggressive subtypes. The intervention of interest was positron emission tomography (PET) to measure glucose metabolism (^{18}FDG-PET) in addition to computed tomography (CT) (including combined PET/CT equipment), referred to in this report simply as PET. PET alone or using other tracers was excluded as an intervention for evaluation of primary outcomes. Combination of PET with MRI (including diffusion weighted MRI), though not the focus of this report, was included as appropriate with the contextual question.

Due to heterogeneity across studies with regard to designs, patient populations, treatments and clinical methods meta-analysis was not performed.

There is heterogeneity across included studies and guidelines with regard to how the timing of PET is defined and classified. For the purposes of this report the 2014 Consensus statement of the International Conference on Malignant Lymphomas Imaging Working Group[8] is used as feasible wherein 'interim PET' (iPET) is conceptualized as early (mid-treatment) assessment during the course of first-line treatment, and end of treatment PET (EoT-PET) is done for remission assessment (or response to treatment) after that initial induction therapy and prior to potential salvage therapy or consolidation therapy. For purposes of this report, PET/CT evaluation of patients in remission done in asymptomatic patients will be referred to as surveillance while PET/CT done if there is clinical suspicion of recurrence or relapse will be noted as such.

For the contextual question, data on diagnostic accuracy were summarized based on systematic reviews published subsequent to 2011 report to the extent possible. For completeness, data from primary studies were reported for lymphoma types included in the report if systematic review information was not available. These primary studies were identified by searching the EndNote library containing search results for our systematic review. For some combinations of lymphoma type and PET use, neither primary studies nor systematic reviews were available. To provide a general context for quality of included systematic reviews, a modified AMSTAR tool was used and for single studies of diagnostic accuracy and reliability, a modification of the QUADAS tool was used (Appendix D). Full formal evaluation of risk of bias was not done and overall strength of evidence was not assessed.

Results

Below is a summary of key findings from the 2011 and 2018 updated reports (Tables A-D).

Findings from the 2011 report were based on a rapid review methodology and focused on evidence summarized from systematic reviews and health technology assessments. Evidence from these sources centered on evaluation of diagnostic accuracy. Updated information on test accuracy is described in the contextual question results section. Subsequent to the 2011 report, more recent evidence has evaluated the role of PET/CT for clinical decision making and adapting treatment based on PET findings. These studies provide the primary focus for this report with emphasis on studies with the least potential for bias (see Key Question section below).

Contextual questions

Summary data in the tables below were taken from Table 1 and Table 2 in the prior report. In the prior report, some conclusions and related strength of evidence appear to be across HL and aNHL populations and are noted in the tables. Information related to specific lymphoma types are noted as provided by authors.

Data for the current report were obtained primarily from systematic reviews published subsequent to 2011 report to provide information to answer the contextual question for this re-review. Overall, 22 SRs and eight primary studies are summarized in the full report to provide context regarding the diagnostic accuracy and reliability of use PET/CT in patient. Formal risk of bias assessment was not done on individual studies and strength of evidence was not assessed for the contextual question. Additional detail and context are available in the full report.

A summary of results for the contextual question on diagnostic accuracy for this re-review are found together with a summary of findings from the previous report. (Tables A-D) There was substantial heterogeneity with regard study designs, PET/CT criteria and thresholds for positivity, patient populations and outcomes reported; the overall quality of systematic reviews and primary studies ranged from moderately good to poor.

Regarding reliability, eight studies evaluating the reliability of PET/CT for various types of lymphoma are summarized in the full report. Across studies and timing of PET/CT, interrater reliability in adult populations ranged from moderate to substantial based on the Landis and Koch criteria when the 5-Point Scale (5-PS) criteria were used. Interrater reliability may vary based on what threshold for positivity of the 5-PS is used and based on reader experience interpretation of oncology-related images. Results across three studies in pediatric patients suggest that interrater reliability ranged from fair to substantial. One study reported that the probability of concordance was highest for extreme values (>50% for 1 and 5) and lowest for middle values (25% for 2, 36% for 3), concluding that the binary distinction for 5-PS is the most reliable criterion for clinical decision making Studies are detailed in the full report.

Table A. Comparison of previous report and re-review: Summary of diagnostic accuracy of PET for HL

Prior 2011 HTA* (KQ1)		This 2018 Re-review HTA (Contextual Question, No SOE)	
Accuracy data	**Conclusions and SOE**	**Accuracy data**	**Conclusions**
Initial staging: NR *Routine staging after primary treatment* • Sensitivity: 84% (95% CI 71% to 91%) • Specificity: 90% (95% CI 84% to 94%) *Monitoring of treatment (mid-cycle,* • Sensitivity: 81% • Specificity: 97% • PPV: 15% • NPV: 100% • LR+: 28.4 • LR-: 0.19 *Evaluation of residual mass* • Sensitivity: 43% to 100% • Specificity: 67% to 100% **Prognosis:** *Estimate of prognosis after primary treatment (PFS)* • Sensitivity: 81% • Specificity: 97% • PPV: 0% • NPV: 63% *Estimate of prognosis after secondary treatment: NR* **Pediatric Populations: NR**	**Initial staging: NR†** **Re-staging (3 SRs): Moderate SOE;** PET/CT has higher sensitivity and specificity than conventional staging; recommend use **Monitoring treatment (*1 SR, 1 primary)*: Moderate SOE** Diagnostic efficacy considered uncertain. Recommend use if part of a clinical trial, not for routine use. **Residual Mass: (3 SRs)** Heterogeneous results across studies; wide range of sensitivity, specificity **Surveillance: Low SOE‡** Recommend against use **Estimation of prognosis: Low SOE** PET/CT outperforms CT in predicting subsequent outcome; recommend use **Pediatric Populations: NR**	**Initial staging (2 primary studies vs. CT):** • Sensitivity: 95%% to 100% • Specificity: 91%% to 99% **Interim PET: no new SRs identified** **SEnd of treatment (1 SR)** • FP: 23.1% (95% CI 4.7% to 64.5%) **Post-treatment surveillance (1 primary study)** • Sensitivity: 100% (95% CI 91.2% to 100%) • Specificity: 86.3% (95% CI 78.7% to 91.5%) • PPV: 72.7% (95% CI 59.0% to 83.8%) • NPV: 100% (95% CI 96.2% to 100%) • LR+: 7.3 (95% CI (4.5 to 11.7) • LR-: 0.0 **Prognosis (3 SRs):** *Estimate of prognosis during primary treatment **(interim)** (1 SR of 14 studies)* • Sensitivity: 67% (95% CI 57% to 76%) • Specificity: 89% (95% CI 84% to 93%) • LR+: 6.2 (95% CI 3.9 to 10.0) • LR-: 0.37 (95% CI 0.27 to 0.50) • NPV: 93% (95% CI 85% to 100%) • AUC: 0.85 (95% CI 0.92 to 0.88) • FP rate 10% *Estimate of prognosis **after** primary treatment, pre-ASCT (PFS) (1 SR)* • Sensitivity: 67.2% (95% CI 58.2% to 75.3%) • Specificity: 70.7% (95% CI 64.2% to 76.5%) *Estimate of prognosis **after** primary treatment (OS) (1SR)* • Sensitivity: 74.4% (95% CI 58.8% to 86.5%) • Specificity: 58.0% (95% CI 49.3% to 66.3%) **End of treatment Confirmed Relapse:(1SR)** • Proportion with FDG-PET-negative residual mass: 6.8% (95% CI 2.6% to 12.5%)	**Initial staging:** PET/CT is accurate for initial staging **End of treatment: y:** Previous HTA suggests that PET/CT outperforms both CT and FDG-PET alone; new data suggests that there may be a large number of false-positive results. Authors suggest that use of end of treatment PET be reconsidered. This new SR does not specify if there was a clinical suspicion of disease progression. **Post-treatment surveillance:** Compared with US/chest radiography, PET/CT had lower specificity and PPV in patients who had achieved PET/CT confirmed complete response following first-line treatment; authors suggest using US/chest radiography for surveillance in these patients. **Prognosis:** Diagnostic accuracy of interim PET may be limited; inconsistency in interpretation criteria may contribute to low estimates and wide CIs; the high NPV may permit stratification based on negative findings. Pre-ASCT PET diagnostic accuracy may be low; Authors observe that patients with positive PET have worse prognosis than those with negative PET

Prior 2011 HTA* (KQ1)		This 2018 Re-review HTA (Contextual Question, No SOE)	
Accuracy data	**Conclusions and SOE**	**Accuracy data**	**Conclusions**
		<u>Pediatric Populations: Pediatric HL (1 primary study, N=57) (based on Deauville criteria)</u> **Initial Staging**: PET vs. CT • Number of detected sites: 219 vs. 152 Relapse occurred in 4/57 **Interim PET** (relapse) • Sensitivity: 0% (95% CI 0% to 60.4%) • Specificity: 91.4% (95% CI 78.7% to 97.2%) • PFS in PET +: 93.3±4.1% • PFS in PET - : 89.6±3.8% **End of Treatment**: (relapse) • Sensitivity: 25% (95% CI 13.1% to 78%) • Specificity: 95.7% (95% CI 84.2% to 99.2%)	**End of treatment, Relapse** One SR evaluating relapse: FDG-PET-negative residual masses are not proven to be associated with worse outcomes compared to post-treatment PET–based complete remission without a residual mass **Pediatric Populations: Pediatric HL** More sites were detected with PET vs. CT for initial staging; For the outcome of detecting relapse, data from this small study are limited, precluding development of firm conclusions. Authors conclude that there is insufficient evidence to support escalation therapy based on interim PET in pediatric HL

ASCT = autologous stem cell transplant; US = ultrasound

*The prior report used a rapid review methodology; Evidence base consisted of 6 SRs, 2 primary studies for the full report.

†Recommendation/SOE appear to have been based on consideration of mixed populations of aNHL and HL together; SOE for the recommendation of Initial staging with PET/CT was moderate for mixed aNHL and HL populations

‡Data basis unclear

Table B. Comparison of previous report and re-review: Summary of diagnostic accuracy of PET for aggressive NHL (including transformed indolent NHL)

Prior 2011 HTA* (KQ1)		This 2018 Re-review HTA	(Contextual Question, No SOE)
Accuracy data	**Conclusions, SOE**	**Accuracy data**	**Conclusions**
Diagnosis: No evidence **Initial staging:** NR *Routine staging after primary treatment* • Sensitivity: 72% (95% CI 61% to 82%) • Specificity: 100% (95% CI 97% to 100%) *Monitoring of treatment (mid-cycle)* • Sensitivity: 78% • Specificity: 87% • PPV: 80% • NPV: 89% • LR+: 5.9 • LR-: 0.26 *Evaluation of residual mass* • Sensitivity: 60% to 78% • Specificity: 94% to 100% **Prognosis:** *Estimate of prognosis after primary treatment (PFS)* • Sensitivity: 78% • Specificity: 87% • PPV: 0% • NPV: 63% *Estimate of prognosis after secondary treatment:* NR **Pediatric populations: NR**	**Diagnosis: NR** **Initial staging: NR†** **Re-staging (3 SRs): Moderate SOE** Recommend use **Monitoring treatment (1 SR, 1 Primary): Moderate SOE** Heterogeneous results across studies; wide range of sensitivity, specificity; Recommend use if part of a clinical trial, not for routine use **Surveillance: Low SOE‡** Recommend against use **Estimation of prognosis (1 SR, 2 primary): Low SOE** PET/CT outperforms CT in predicting subsequent outcome; recommend use **Pediatric populations: NR**	**Diagnosis (2 SRs):** Immune competent people • Sensitivity: 88% (95% CI 80% to 94%) • Specificity: 86% (95% CI 73% to 94%) • AUC 0.9192 HIV patients • Sensitivity: 100% • Specificity: 75% to 100%) **Initial staging (1 SR with patient-level data):** • Sensitivity: 95% (95% CI 89% to 98%) • Specificity: 40% (95% CI 9% to 78%) • DOR: 13.28 (95% CI 2.60 to 67.82) • LR+: 1.63 (95% CI 0.91 to 2.91) • LR-: 0.13 (95% CI 0.05 to 0.38) • AUC: 0.8537 Interim PET (1SR, various NHL types) • FP rate: Pooled 83% (95%CI 72% to 90.2%) End of treatment (1SR, various NHL types) • FP rate: Pooled 31.5% (95%CI 3.9 to 83.9%) **Post-treatment surveillance (2 primary studies)** DLBCL patients in complete remission after primary therapy • Sensitivity: 100% (95%CI 30.5%-100%) • Specificity: 86.1%(95% CI 79.8 to 91.0) • FP rate vs. CT: 13.7 vs. 1.7% • PPV vs. CT: 11.5% vs. 30.3% • NPV vs. CT: 100% vs. 100% Transformed indolent NHL • Sensitivity: 83% • Specificity: 94% • PPV: 63% • NPV: 98% • FP 2% to 5%	**Diagnosis:** PET is reasonably accurate for aNHL diagnosis **Initial staging** PET has high sensitivity but poor specificity **Interim and end of treatment**: 1 SR reports a high frequency of false positives at both time frames; Confidence intervals around estimates are wide, particularly for end of treatment PET.: **Post-treatment surveillance** Both studies conclude that surveillance is not indicated DLBCL patients who were in complete remission after primary therapy*:* PET/CT had higher FP rate vs. CT, resulting in unneeded procedures and that PET/CT should be reserved for evaluation of clinically suspected relapse **Prognosis** Diagnostic accuracy of PET/CT for prognosis varied substantially across systematic reviews and results across them conflicting; some suggest PET is accurate for predicting prognosis, others do not. Differences in treatments, heterogeneity of populations within and across studies, use of different criteria for

Prior 2011 HTA* (KQ1)		This 2018 Re-review HTA	(Contextual Question, No SOE)
Accuracy data	**Conclusions, SOE**	**Accuracy data**	**Conclusions**
		Prognosis (6 SRs): *Estimate of prognosis during primary treatment (interim PET) (PFS) (3 SRs)* • Sensitivity: 21.2% to 89.7% • Specificity: 37.4% to 90.7% • AUC: 0.651 • HR (PET/CT positive vs. PET/CT negative): 2.91 to 4.4 *Estimate of prognosis during primary treatment (OS) (3 SRs)* • Sensitivity: 8.3% (newly diagnosed to 82.2%) • Specificity: 30.5% (newly diagnosed to 91.5%) • AUC: 0.817 • HR (PET/CT positive vs. PET/CT negative): range 1.99 to 5.91 *Estimate of prognosis after primary treatment (PFS) (3 SRs)* • HR (PET/CT positive vs. PET/CT negative): range 4.05 to 6.75 *Estimate of prognosis after primary treatment (OS) (3 new SRs, 3 SRs from prior report)* • HR (PET/CT positive vs. PET/CT negative): 5.91 (95% CI 3.15 to 11.09) *Estimate of prognosis after secondary treatment:* NR **<u>Pediatric: Pediatric nonlymphoblastic NHL (1 primary study, N=34)</u>** **Initial Staging**: PET vs. CT • Number of detected sites: 130 vs. 114 **Interim PET** (monitoring during treatment) • Sensitivity: 77.8% (95% CI 44.9% to 95.9%) • Specificity: 54.5% (95% CI 41.1% to 62.0%) **End of Treatment**: • Sensitivity: 75.0% (95% CI 23.5% to 98.7%) • Specificity: 75.0% (95% CI 66.4% to 78.9%)	positivity and timing of PET all likely contribute to the variation. In general, prognosis for individuals with a positive PET is worse than those with a negative PET. (See full report for detail). **Pediatric populations: Pediatric nonlymphoblastic NHL** Confidence intervals for diagnostic accuracy measures are wide due to the small sample size in one study making conclusions regarding accuracy difficult. Authors conclude that PET/CT may be better than CT for initial staging, that interim PET/CT findings cannot be used to prognosticate for PFS or OS and post-treatment PET/CT help prognosticate PFS but not OS

Prior 2011 HTA* (KQ1)		This 2018 Re-review HTA	(Contextual Question, No SOE)
Accuracy data	**Conclusions, SOE**	**Accuracy data**	**Conclusions**
		Prognosis: (complete metabolic remission vs. not): Based on interim PET • PFS: 85.7% (95% CI 53.9% to 96.2%) • OS: 78.6% (95% CI 47.2% to 92.5%) Based on end of treatment PET • PFS: 94.7% (95% CI 68.1% to 99.2%) • OS: 84.2% (95% CI 58.6% to 94.6%)	

*The prior report used a rapid review methodology; Evidence base consisted of 6 SRs, 2 primary studies.

†Recommendation/SOE appear to have been based on consideration of mixed populations of aNHL and HL together; SOE for the recommendation of Initial staging with PET/CT was moderate for mixed aNHL and HL populations

‡Data unclear

Table C. Comparison of previous report and re-review: Summary of diagnostic accuracy of PET for indolent NHL

Prior 2011 HTA* (KQ1)		This 2018 Re-review HTA (Contextual Question)	
Accuracy data	**Conclusions, SOE**	**Accuracy data**	**Conclusions**
Initial staging: • Sensitivity: 99% **Routine staging after primary treatment: NR** **Monitoring of treatment (mid-cycle): NR** **Evaluation of residual mass: NR** **Prognosis:** *Estimate of prognosis after primary treatment (PFS)* • Sensitivity: 100% • Specificity: 88% • PPV: 62% • NPV: 100% *Estimate of prognosis after secondary treatment:* NR **Histological transformation of iNHL to aNHL** • Sensitivity: 91% to 93% • Specificity: 67% to 87% • LR+: 2.6 to 7.2 • LR-: 0.08 to 0.14	**Initial staging (1 SR, 6 Primary): Low SOE** PET/CT outperforms CT; recommend use **Re-staging: No evidence** **Monitoring treatment: No evidence** **Surveillance: No evidence** **Estimation of prognosis (2 primary): Low SOE;** PET/CT is reasonably accurate in predicting relapse of iNHL, negative PET/CT appears more valuable than positive PET/CT**; r**ecommend use **Histological transformation (1 SR, 2 primary):** High sensitivity, variable specificity	**Staging (2 SRs, 1 primary study):** • Sensitivity: 77% to 96% • Specificity: 98% (95% CI 97% to 99%) • Proportion of patients upstaged based on FDG-PET results: 18.7% (95% CI 10.8% to 30.4%) : **Assessment of treatment response (1 SR including 1 small study):** • Sensitivity: 72% (95% CI 47% to 90%) • Specificity: 100% (95% CI 87% to 100%) **Post-treatment surveillance:** NR **Prognosis (4 SRs):** *Estimate of prognosis after primary treatment (PFS) (2 SRs)* • HR (PET/CT positive vs. PET/CT negative): range 3.9 to 5.1 • PPV: 89% to 94% *Estimate of prognosis after primary treatment (OS) (1 SR)* • HR (PET/CT positive vs. PET/CT negative): 6.7 (95% CI 2.4 to 18.5) *Estimate of prognosis after secondary treatment: NR*	**Staging:** PET/CT may be as specific but more sensitive vs. CT for relapse; PET may have high sensitivity in newly diagnosed marginal zone B cell lymphoma. PET may improve staging. **Treatment response:** PET/CT has higher specificity but lower sensitivity than CT in patients with indolent NHL **Post-treatment surveillance: NR** **Prognosis** Conclusions varied across reviews. In general, prognosis for individuals with a positive PET is worse than those with a negative PET; although one SR of follicular lymphoma reports no significant difference in PFS in PET + vs. PET- patients and concludes use of interim PET is not supported by evidence. (See full report for detail) another concluded that PET-based response criteria are more predictive of progression free survival than CT-based response criteria.

*The prior report used a rapid review methodology; Evidence base consisted of 1 SR, 6 primary studies.

Table D. Comparison of previous report and re-review: Summary of diagnostic accuracy of PET for mixed populations of aggressive NHL and HL

Prior 2011 HTA* (KQ1)		This 2018 Re-review HTA (Contextual Question)	
Accuracy data	**Conclusions, SOE**	**Accuracy data**	**Conclusions**
Initial staging: • Sensitivity: 88% to 100% • Specificity: 90% to 100% *Routine staging after primary treatment*: NR *Monitoring of treatment (mid-cycle)* • PPV: 15% to 80% • NPV: 90% to 100% *Evaluation of residual mass* • NR **Prognosis:** *Estimate of prognosis after primary treatment:* NR *Estimate of prognosis after secondary treatment (PFS)* • Sensitivity: 69% • Specificity: 81% • LR+: 3.6 • LR-: 0.387 • HR: 3.23 (95% CI NR) **Pediatric populations: NR**	**Initial staging (4 SRs, 2 primary): Moderate SOE** Recommend use **Re-staging: Moderate SOE** Recommend use **Monitoring treatment: Moderate SOE** Recommend use if part of a clinical trial, not for routine use **Surveillance: Low SOE†** Recommend against use **Estimation of prognosis (2 SRs, 3 primary): Low SOE; report** PET/CT has lower sensitivity and specificity in predicting subsequent outcome after secondary treatment than after primary treatment; recommend use **Pediatric populations: NR**	**Initial staging:** No new systematic reviews During or end of treatment (various NHL types) (1 SR) • FP: range 7.7% to 90.5% **Post-treatment surveillance (2 SRs):** • Sensitivity: 75% to 100% • Specificity: 43% to 92% • PPV: 20% to 100% • NPV: 98% to 100% **Prognosis (1 SR):** *Estimate of prognosis after primary treatment Pre-ASCT (PFS) (1 SR)* • Overall HR (PET/CT positive vs. PET/CT negative): 2.32 (95% CI 1.52 to 3.12) • HL pooled HR (3.28 (95% CI 1.61 to 4.95) • NHL pooled HR 2.00 (95% CI 0.71 to 3.29) *Estimate of prognosis after secondary treatment (PFS) Post-ASCT (1 SR)* • HR (PET/CT positive vs. PET/CT negative): 4.61 (95% CI 1.59 to 7.64) *Estimate of prognosis after secondary treatment (OS) (1 SR)* • HR (PET/CT positive vs. PET/CT negative): 2.64 (95% CI 1.55 to 3.72) **<u>Pediatric populations: (1 SR of 5 studies in pediatric HL and NHL)</u>** **Initial Staging**: PET vs. conventional imaging • Sensitivity: 96% to 99% • Specificity: 95% to 100% **Interim PET** (PFS, PET vs. CT, therapeutic response) • Sensitivity: 77% to 100% • Specificity: 54.5% to 97.7%	**Initial staging: No new evidence** **During or after treatment PET**: PET/CT scans resulted in a wide range ofof false positives across lymphoma types and time frames; Authors conclude that the role of PET/CT be reconsidered. **Post-treatment surveillance** Two SRs concluded that there is a lack of evidence to support PET/CT in post-treatment surveillance **Prognosis** In general, prognosis for individuals with a positive PET is worse than those with a negative PET. Authors conclude that PET is a useful prognostic tool before ASCT in HL but not NHL. **Pediatric Populations:** Limited data are provided. The accuracy of PET for initial staging may be high; a broad range of sensitivity and specificity is noted for interim PET.

*The prior report used a rapid review methodology; Evidence base consisted of 3 SRs, 5 primary studies.

†Data unclear

Key Questions

For purposes of this report, the term "PET-adapted" refers to studies that explicitly describe what specific management changes or treatments were done or not done based on PET results, the number of patients impacted by the PET-based treatment decision(s) and describe the prognosis in terms of OS and PFS for those patients. Studies of PET-adapted treatment evaluated interim or end of treatment PET primarily. Studies that provided general information on PET's impact on staging (e.g. upstaging or down staging) or report non-specific treatment changes are briefly abstracted and summarized. Such outcomes were considered surrogates for treatment and therefore secondary and intermediate outcomes. Identified studies that were not PET-adapted were either prospective or retrospective cohort studies. Overall, these studies were considered at to be at moderately high or high risk of bias. Detailed risk of bias was not assessed for these studies; however selection bias and performance bias were generally noted across them during abstraction. They were not included in determination of overall strength of evidence.

Except where noted the term PET will be construed to mean PET/CT

KQ1. Summary of Results:

For efficacy and effectiveness, 18 nonrandomized observational studies provided data for initial staging with PET/CT[3,5-7,9,14,16,17,19,28-30,33,40,46,49,58,69]; only four of these studies were in pediatric populations[7,19,49,58] one of which provided data separately for children with HL and NHL.[19] For interim PET evaluation, a total of 26 studies (across 28 publications) were identified and included eight RCTs (10 publications) of PET-adapted therapy[4,11,12,15,24,31,37,39,54,55]; 15 observational studies of PET-adapted therapy[13,21,23,32,38,42-44,47,53,59,62-64,71]; and three clinical nonrandomized observational studies.[7,52,67] With the exception of one clinical observational study in children,[7] all studies evaluating interim PET were in adult populations. For end-of-treatment PET evaluation, a total of six studies (across 7 publications) were identified and included two RCTs of PET-adapted therapy[24,31]; two nonrandomized observational studies (3 publications) of PET-adapted therapy[25,26,70]; and two clinical observational studies.[7,61] With the exception of one clinical observational study in children,[7] all studies evaluating end-of treatment PET were in adult populations. For surveillance imaging with PET, a total of eight studies, one RCT[50] and seven nonrandomized observational studies (8 publications) provided data.[18,22,27,35,48,51,65,66]

Initial staging

Summary of evidence (strength of evidence not assessed for indirect/secondary outcomes)

- No studies of PET/CT- guided initial treatment based on incorporation of PET/CT results in initial staging were identified. No study provided detail on how specific changes in clinical management affected patient outcomes.
- A total of 18 observational studies at moderately high to high risk of bias compared initial staging with PET/CT or PET versus CT alone or with conventional staging in patients with lymphoma. Outcomes reported primarily included the proportions of patients upstaged or down staged and/or proportions of patients for whom the change in stage would impact treatment. These are considered surrogate outcomes for treatment modification and therefore indirect.

- In general across lymphoma types, more patients were upstaged versus downstaged when PET/CT was incorporated into the initial staging process. Changes in stage varied widely (8% to 50%) as did actual or theoretical change in management (6% to 70%). Studies that include mixed populations (e.g. mixed HL and NLH) are described in the full report. Data across studies reporting on specific lymphoma types are briefly summarized in Table E below
- Survival outcomes based initial staging with PET/CT were reported in two studies of adults (SOE INSUFFICIENT).
    - One study in patients with early-stage HL found that pretreatment PET imaging was associated with a greater probability of relapse-free survival (RFS) compared to no pretreatment PET imaging: 4-year RFS, 97% versus 67%, p=0.001.[28]
    - One study in patients with early-stage follicular lymphoma found no significant difference in overall survival (data not provided) or the probability of progression-free survival (PFS) over a median follow-up of 4.8 years when initial staging was done with PET/CT compared with conventional methods (hazard ratio [HR] for PFS for conventional vs. PET/CT: 0.87, 95% CI 0.47 to 1.62).[29]

Table E. Evidence summary: Change in clinical stage comparing PET/CT vs. CT alone or conventional staging in HL, aNHL and iNHL*

Lymphoma type	Number of Studies (comparator)	Total Change (Range)	Upstage (Range)	Downstage (Range)	Change in management due to PET/CT*
HL (Adults); Early-, late-stage	3 (CT alone) 1 (conventional)	32%-34% 22%	24%-31% 16%	3%–8% 6%	11%–21% 14% (theoretical) 6% (actual change)
HL (Pediatric)	2 (CT alone) 1 (conventional)	*23%–50%* *9%*	*23%–27%* *6%*	*0%–23%* *4%*	70% 21%
aNHL (Adult)	1 (CT alone) 1 (conventional)	5% 11%	2% 11%	3.2% 0%	0% 11%
iNHL (Adults)	3 (CT alone) 1 (conventional)	*29%–42%* *NR*	*29%–37%* *NR*	*0%–5%* *NR*	39% (1 study)† *NR‡*
iNHL (Pediatric)	1 (CT alone)	*15%*	*15%*	*0%*	*0%*

*See full report for summaries in mixed populations.

†A second study reported an increased likelihood of receiving early therapy if staged with PET/CT (adjusted OR 1.87 (95% CI 1.31-2.66), p=0.0006) and that PET/CT was associated with significant differences in choice of initial therapy; p=0.02 (more anthracycline based chemotherapy ± immunochemotherapy and less treatment with radiation only).

‡Authors only report that there was no significant difference (p=0.47) in treatments received following staging with PET/CT vs. conventional.

Interim assessment

Early-Stage Hodgkin Lymphoma (HL)

PET-adapted treatment for early-stage HL: RCTs

- No trials compared treatment strategies based on use of PET with strategies that did not use PET. No studies in children met the inclusion criteria.
- Two randomized trials of PET-adapted therapy in persons with early-stage HL (stage I, IA, II, IIA) were identified. One trial randomized patients with positive interim PET after two cycles of initial therapy (ABVD) to escalated (BEACOPP) versus standard treatment[4,55] (1-2 additional ABVD cycles plus involved node radiotherapy (INRT). The same trial randomized persons with negative interim PET to different de-escalation regimens based on favorable or unfavorable prognostic status at initial staging. The de-escalation component (omission of INRT) was stopped early and a safety amendment issued based on futility analyses for this non-inferiority trial. Better PFS in those in the standard treatment group (those with ABVD + INRT) vs. those receiving the ABVD only led to closing the arms which omitted INRT. The second trial randomized persons with a negative interim PET to no further treatment versus one additional cycle of ABVD and INRT.[54]
- Overall, in persons with early-stage HL, interim PET may be of most value for identifying those who may benefit from treatment escalation based on positive interim PET findings in one trial. For results across the two trials evaluating treatment de-escalation (elimination of consolidative RT) persons with negative interim PET the impact of PET-directed treatment is less clear.
    - **Escalated therapy in PET positive groups**: In the one trial[4,55] evaluating escalation of treatment versus standard treatment in persons with a positive interim PET, the probability of both OS and PFS was greater in those who had the escalated therapy. There was, however, substantially greater treatment-related toxicity reported in those with escalated therapy. (SOE MODERATE for all outcomes)
    - **De-escalation of therapy in PET negative groups:**
        - Results across the two trials that randomized persons with a negative interim[4,54,55] PET differed for OS, but were generally consistent for PFS; the overall benefit of PET-adapted treatment is less clear compared with escalation strategies described above.
        - At 3 years the difference in OS was small between those who received no further treatment vs. those who had INRT in the RAPID trial. In the H10 trial at 5 years, the differences in OS between groups randomized to de-escalated therapy omitting INRT and standard therapy with it were small regardless of initial staging as favorable (0.4/100) or unfavorable (1.6/100) and were not statistically different. Across the favorable and unfavorable groups the difference between de-escalated (99%) and standard treatment with INRT (98%) was estimated to be small, 1/100 and of unknown clinical significance (SOE MODERATE)
            - PFS in both trials was less common in groups receiving de-escalated therapy where INRT was omitted. In the RAPID trial an absolute

difference of -3.8% at 3 years was seen but was not statistically significant. Similarly in the H10 trial PFS was lower in those with an initially favorable prognosis (difference -11.9/100) and for those with an initially favorable prognosis (difference - 2.5/100). Across prognostic groups the overall difference was -7/100. Based on PFS, results suggest that exclusion of INRT may have a clinically important impact, particularly in those with an initially favorable prognosis and should be retained in the treatment regimen. (SOE MODERATE)

- Evidence on treatment-related toxicity was sparse and no clear conclusions regarding the impact of interim PET-adapted treatment to reduce such toxicities are possible. Follow-up time may have been insufficient to capture some RT-related toxicities. Power to detect rare events and differences between strategies was limited. (SOE LOW)
- Difference in therapeutic regimens, study design, patient populations and length of follow-up may contribute to differences across trials.

- Heterogeneity of patients within and across studies is likely and the impact of false-positives, false-negatives and other limitations of PET are unclear. The consistency and generalizability of these findings for other treatment regimens, different clinical settings and patient populations is unknown.
- Data through 2014 from the National Cancer Institute's SEER database indicate that 5-year relative survival for stages I and II HL are 92.3% and 93.4% respectively.

PET-Adapted Treatment: Observational studies (SOE not assessed)

- Across two observational studies[21,63] in which patients with HL stage I or II (without bulky tumor) with positive interim PET received escalated treatment and those with negative interim PET received de-escalated treatment after two ABVD cycles, PFS was substantially worse in PET-positive individuals. Prognosis in terms of OS, was also slightly worse. It is unclear how much improvement in survival might be expected from the escalated therapies and the ability of PET to guide effective treatment are unclear from the data provided.

Advanced-Stage Hodgkin Lymphoma (HL)

PET-Adapted treatment for advanced-stage HL: RCTs (SOE MODERATE for all outcomes)

- No trials compared treatment strategies based on use of PET with strategies that did not use PET. All trials were in adult populations
- A total of three randomized PET-adapted treatment trials in persons with advanced-stage HL were identified. Randomization in the trials was to treatment regimens based on interim PET results. Two trials compared randomized escalated treatment regimens in persons with positive interim PET results.[11,12,31] One of these also compared persons with negative interim PET results randomized to de-escalation of treatment versus continuation of an escalated treatment.[11,12]

The third trial randomized those with PET negative results to de-escalated therapy[37] versus continuation of standard treatment. There is substantial heterogeneity across trials with regard to initial treatments, PET-adapted treatment regimens, patient populations and differences in criteria for determining positivity of PET results.

- Overall, in persons with advanced-stage HL, interim PET appears to be of most value for the identification of persons who may be candidates for de-escalation of therapy based on negative PET scan results. Limited data across two trials suggest that the probability of OS and PFS in those receiving scaled back treatment are similar to those continuing more intensive treatment while many toxicities related to treatment were substantially reduced in those whose treatment was de-escalated.
- On the other hand, the value of PET to direct to escalation treatments when interim PET results are positive is less clear in the identified trials. Across trials of two different types of escalated therapy estimated differences per 100 patients between treatments for OS or PFS were small at 3 years and in one trial at 5 years and not beyond what may be expected by chance. Toxicities based on treatments received were also generally similar between treatment groups. The role of false positives and other limitations of PET are unclear.
- With regard to prognosis based on interim PET results alone, in the HD 18trial no difference in OS or PFS for PET positive and negative groups was observed. This may be interpreted in several ways. One interpretation is that PET may not adequately discriminate between individuals who would benefit most from escalated therapy perhaps due to population heterogeneity and/or limitations of PET. Another explanation may be that the PET + groups that receive escalated therapy in fact had better survival following escalated therapy versus survival had they not received it. There are differences in patient characteristics and disease severity between PET + and PET – groups. The other trial did not report results of statistical testing for these groups.
- The consistency and generalizability of these findings to other treatment regimens, clinical settings and patient population is unknown. Each trial used different induction therapies. PET positivity did not appear to identify high risk patients who may benefit most from escalate treatment in the HD18 trial.
- Data from 2008 through 2014 from the National Cancer Institute's SEER database indicate that 5-year relative survival for stage II is 93% stage III is 83% and 73% for Stage IV HL.[2] Across stages, 5-year relative survival is about 87% and 10 year survival 78%. Death rates have been falling on average 2.8% each year over 2006-2015. Authors of included trials indicate that standard ABVD treatment for advanced-stage classic HL results in a 3-5 year PFS that ranges from 61% to 76% with cure rates of 70% to 80%, and that 5-year freedom from treatment failure may be achieved in up to 90% of patients receiving more intensive regimens (i.e., escalated BEACOPP).

PET-Adapted treatment for advanced-stage HL: Observational studies (SOE not assessed)

- The benefit of interim PET across eight observational studies reporting PET-adapted treatment is unclear. Patient characteristics, treatments and methods across studies varied substantially making conclusions across studies difficult. Detailed data on the impact of specific PET-adapted

treatments were limited, including information on impact of escalated or de-escalated treatment on survival or treatment-related toxicity. Studies were at moderately high or high risk of bias.

- In all but two studies,[13,23] prognosis in terms of OS or PFS at various time frames was worse in those with a positive interim PET. Only a limited number of patients with positive interim PET were available in the two studies in which prognosis was similar for those with a positive versus a negative interim PET.

Relapsed or Refractory Hodgkin Lymphoma (HL)

Pre-Autologous Stem Cell Transplantation (ASCT) PET (SOE INSUFFICIENT)

- No RCTs comparing treatment strategies based on use of PET with strategies that did not use PET were identified and no randomized PET-adapted studies were identified.
- Three Observational studies of[42-44] at moderately high risk of bias tused PET results to allocate PET negative patients directly to high dose radiochemotherapy (HDT) and ASCT versus using extended salvage chemotherapy in those with positive PET were identified. Limited information from these studies suggests that pre-ASCT PET may facilitate identification of those who may be able to progress to HDT/ASCT if results are negative from those who may require extended savage therapy prior to HDT/ASCT; pre-transplant PET results may provide valuable prognostic information as achievement of PET negativity prior to HDT/ASCT have better probability of survival.
    - The 2012 study reports that EFS for persons with pre-HDT/ASCT negative FDG-PET, after 1 or 2 cycles of salvage therapy was > 80%, versus 28.6% for patient who had a positive scan (P < .001); no data comparing the impact of PET-adapted strategies on toxicity were provided.
    - EFS at 3 years for those with pre-ASCT PET positive results was 60%, significantly less compared with those having a with pre-ASCT PET negative result (p = 0.05) in persons who had ASCT based on combined data from the 2015 and 2017 publications.[42,43] Across patients, 76 % (95% CI 62–89) achieved PET-negative status at the end of all treatments.
- The reports do not provide adequate data to evaluate the impact of PET-adapted therapy in terms of OS, PFS and toxicities relative to specific PET-adapted therapies.
- A recent meta-analysis of ASCT in patients with HL reported pooled estimates at 2 years for OS (58% and 50%) and for relapse free survival (37% and 31% respectively)[56].

Aggressive Non-Hodgkin Lymphoma (aNHL): Diffuse Large B-cell Lymphoma (DLBCL)

PET-Adapted treatment for aNHL (DLBCL): RCTs

- No trials randomizing patients to PET/CT versus non-PET/CT treatment strategies were identified in persons with aggressive NHL. One RCT of PET-adapted escalation of treatment for those with positive and negative interim PET and one RCT of induction treatments in patients with stage III or IV NLH were identified. A third trial in early-stage (I or II with no bulky mass) low risk patients was also identified. Conclusions across these trials regarding the value of interim

PET differ, due to differences in patient populations, study designs and effectiveness of treatment escalation strategies.

- The PETAL trial[24] randomized patients with aNHL to treatments based on interim PET/CT findings. PET/CT was done after two cycles of CHOP or R-CHOP (rituximab use was restricted to patients with CD20 positive lymphomas), persons with positive interim PET were randomly assigned to receive six blocks of an intensive Burkitt's lymphoma treatment protocol or six additional cycles of R-CHOP. Individuals with a negative interim PET with CD20 positive lymphomas were randomly assigned to receive four additional cycles of R-CHOP or the same treatment with two additional doses rituximab. Recruitment was stopped early because PET-positivity was rare and the Burkitt's protocol was more toxic and possibly less effective than R-CHOP. (SOE Moderate for all outcomes)
    - In persons with a positive interim PET, OS, PFS and EFS were lower in patients who received the Burkitt's protocol versus those receiving additional RCHOP; the differences were within what might be expected by chance. Higher risk of treatment-related toxicities was seen in those receiving the Burkitt's protocol.
    - In persons with a negative interim PET, additional doses of rituximab did not appear to substantially improve OS, PFS or ES. Risks for treatment-related adverse events did not consistently favor one or the other treatment and differences may not have been beyond what might be expected by chance.
    - The prognosis for those with a positive interim PET was substantially worse than for those with a negative interim PET.
    - Based on this trial, the value of interim PET to identify those who may benefit most from escalated therapy from this trial is unclear.
- The LNH-3B trial randomized high-risk patients with CD20 positive DLBCL to different induction treatments and used results of interim PET after 2 and 4 cycles of therapy to adapt additional treatment[15]. Individuals with negative PET results at both time frames, received standard immunochemotherapy (SIC) consolidation, where as those with positive PET after cycle 2 received escalated therapy including autologous stem cell transplantation. Treatment of participants with positive PET results after 4 cycles of therapy was at the discretion of the local investigator. Based on this trial, interim PET positive results may facilitate identification of the subgroup of patients for whom escalation of treatment may improve PFS. This should be interpreted with caution as it assumes that the prognosis of PET-positive patients would be worse in the absence of escalated treatment. No data on the impact of these strategies on PET-adapted treatment-related toxicities were reported, making it difficult to draw firm conclusion regarding the overall value of interim PET in aNHL. (SOE low for survival outcomes, insuffiicent for toxicity)
    - At 4 years, the probability of OS was similar between those who received escalated therapy and those who had standard immunochemotherapy (SIC) consolidation (difference 0.8/100), however PFS was better in those receiving escalated therapy (difference 10/100).
    - Information on treatment-related toxicities by PET-adapted treatments was not provided.

- The LYSA/GOELAMS trial[39] randomized patients at initial staging to receive RT or not after 4-6 cycles of R-CHOP. Patients with interim PET negative results (those in complete remission) patients continued with the treatments as randomized. OS at 5 years was similar between those who did and those who did not receive RT (92% vs. 90%) as was the frequency of relapse (RD 1%) suggesting that RT might be omitted in these patients. (SOE Moderate).
- Specific to DLBCL, clinical guidelines suggest that 3-year overall survival may range from 59% for high risk patients to 91% for low risk patients[68] and has improved in the last decade due to the addition of rituximab[60]. Untreated, median survival is estimated at less than 1 year.[60] Data through 2014 from the National Cancer Institute's SEER database indicate that 5-year relative survival across types of NHL is around 71%[45] and may be as high as 82% in those with stage I or as low as 62% in those with stage IV.
- The consistency across and generalizability of the findings from the trials to other treatment regimens and patient population is unknown.

PET-Adapted treatment for aNHL (DLBCL): observational studies

- Across three observational studies[47,62,64] evaluating three different treatment escalation strategies in those with positive interim PET, probabilities of OS, PFS and EFS, were lower than for those with negative interim PET who continued with standard therapy. Probabilities for survival outcomes were generally in the same range as those reported in the RCTS. The impact of early escalated therapy in these patient populations is unclear. No study compared escalated therapy to standard therapy for interim PET+ patients and therapy regimens differed substantially across studies as did patients characteristics.
- Two studies[62,64] provided comparative data related to toxicity. The incidence of toxicity-related adverse events was greater in patients whose treatment was presumably escalated PET+) versus not-escalated (PET–), as were treatment-related deaths.

End-of-treatment assessment

Advanced-Stage Hodgkin Lymphoma (HL)

PET-Adapted treatment for advanced-stage HL: RCT

- One RCT[31] in adults with advanced-stage HL evaluated end-of-treatment (EOT) PET for direction of further therapy. No studies compared treatment strategies based on use of PET with strategies that did not use PET. All studies were in adult populations.
- The PET-adapted treatment trial randomized patients who had large nodal masses at baseline but whose interim and post-treatment PET/CTs were negative to consolidation radiation therapy versus no further treatment[31] (no radiation therapy). Similar probabilities for OS and PFS were seen for the treatment groups. PFS for the 260 patients with no LNM at baseline and not randomly assigned to RT were only slightly lower (92%, 95% CI, 88% to 95%) with an estimated difference of 5/100 between those randomized to RT and those who were not randomized. In patients with large nodal masses at baseline, PET may help identify patients in whom

consolidation RT may safely be eliminated based on negative results for both interim and end of treatment PET.

- Treatment-related toxicity for this set of treatments was not reported.

PET-Adapted treatment for advanced-stage HL: Observational

- PFS from a poor quality retrospective study[25,26] was similar between patients in partial remission with a negative EOT PET and those considered to be in complete remission after induction treatment: authors concluded that radiotherapy could be omitted in those with a negative PET in partial remission.

Aggressive Non-Hodgkin Lymphoma (aNHL)

PET-Adapted treatment for advanced-stage HL: Observational

- In a subset of 50 patients in the PETAL trial[24] who had positive interim PET findings, those with persistent abnormalities on EOT PET who had no clinical evidence of disease were assigned to observation and those and those who had clinical evidence of disease as well received additional immediate treatment. Both OS and PFS were better in the asymptomatic observation group versus those with clinical evidence of disease. Prognosis in those with negative EOT PET was better compared with the full group of patients with persistent PET abnormalities. This may suggest the importance of considering clinical evidence of disease together with findings of persistent PET abnormalities at the end of treatment. The role of false positives and false negatives is unclear. Data on toxicity specific to EOT PET decisions were not provided.
- One retrospective cohort study[70] provided reported that disease-free survival at 10-years was similar for those with positive EOT PET (received RT) and negative EOT PET (observation) groups (91% versus 90%, respectively). Another observational study[47] reported that EOT PET was a strong predictor of outcome; compared with a PET+ result after completion of treatment, a PET- result was associated with a significantly greater probability of both OS (98% vs. 67%) and PFS (81% vs. 57%).

Surveillance imaging and assessment of relapse

- One RCT (moderately low risk of bias)[50] compared PET/CT with ultrasound and chest radiography for post-remission surveillance in apparently asymptomatic patients following treatment for advanced-stage HL was identified. Four retrospective cohort studies (moderately high to high risk of bias) compared use of PET/CT in patients in asymptomatic patients with its use only in those with clinical suspicion of disease.
- In the RCT, relapse was detected equally in both PET/CT and US/chest radiography groups (40 in each); clinical findings were found in 36% (29/80) who had relapse. False positive findings were more common in the PET/CT group (13.7%) versus the US/radiography group (3.7%)

- Across three of the cohort studies[22,27,51] in persons with different types of lymphoma, OS and PFS were similar for asymptomatic groups who received follow-up PET/CT versus who received routine CT or those who received imaging for clinical indication only. (SOE INSUFFICIENT)
- Across three studies, PET/CT may not be superior to clinical follow-up alone (i.e., symptomatic detection) at identifying relapses.[18,27,51]
- One cohort study of NHL PET/CT done in those with a clinical suspicion of disease yielded more true positive results that scans done in asymptomatic patients. False positive scans in asymptomatic patients ranged from 2% to 7.5%.[65,66] PET/CT led to changes in clinical management less than 10% of the time in asymptomatic patients compared with >30% in patients receiving scans based on clinical suspicion of disease.
- In general, authors of these studies and those included for the contextual question for diagnostic accuracy conclude that routine surveillance of asymptomatic persons should not be done.

Table F. Evidence summary: Relapse detection comparing PET/CT vs. other imaging or clinical follow-up alone.

Study, year Study design Lymphoma type	Overall relapse rate	Relapse detected: Routine Surveillance PET	Relapse detected: Clinical follow up
Pingali 2014 Retrospective cohort Hodgkin Lymphoma	4.6% (11/241)	3.4% (6/174)	7.5% (5/67) (PET if clinical suspicion)
Cheah 2014 Retrospective cohort Transformed iNHL	29.1% (16/55)	12.7% (7/55) FP 5%	16.4% (9/55) FP: 11%
Epperla 2016 Retrospective cohort Aggressive NHL-relapsed or refractory	28% (45/160)	15% (24/160) (PET or PET/CT)	8% (13/160) (clinical only)
		Relapse detected: PET/CT Surveillance	**Relapse detected: US/Chest X-ray Surveillance**
Picardi 2014 RCT Hodgkin Lymphoma	26.7% (80/300)	26.7% (40/150) FP: 13.7%	26.7% (40/150) FP: 3.7% (4/40

FP = false positive; iNHL = indolent non-Hodgkin lymphoma; NHL = non-Hodgkin lymphoma; PET = positron emission tomography; PET/CT = combination positron emission tomography and computed tomography; RCT = randomized controlled trial; US = ultrasound.

KQ2. Summary of adverse events and safety:

For safety, a total of nine studies met inclusion criteria: one RCT,[50] seven observational, [22,34,35,47,48,51,57] and one case series.[20] Of the studies that provided data for efficacy or effectiveness, only three comparative observational studies reported safety outcomes.[22,47,51]

Summary of results:

- Few studies were identified that reported safety outcomes related to the use of PET/CT. Most studies were considered moderately high to high risk of bias with the exception of one RCT that was at moderately low risk of bias. Strength of evidence for all safety outcomes is considered low.
- Mean effective radiation doses from PET/CT varied across studies based on number of scans performed, length of follow-up and other factors such as patient's age. Reported doses for adults ranged from from 14.5 mSv to 26.1 mSv for surveillance (3 studies) and 26.1 mSv during treatment (1 study); for children, doses were estimated to be 6.4 mSv and 8.6 mSv for a 10- and 15-year-old child, respectively from one study.
- Across two studies, high false positive rates with PET/CT reported during surveillance led to additional perhaps unnecessary diagnostic testing: 50% (vs. 15% for US/chest radiography) (1 RCT) and 14% (vs. 6% for CT alone) (1 observational).
- Three studies reported relapse identification during surveillance. One RCT found that US/radiography was as effective at identifying relapse as PET/CT (96% vs. 100%, respectively) with significantly lower radiation exposure (10.2 mSv vs. 449.5 mSv). Two observational studies found that fewer imaging tests were needed to identify a relapse when done as clinically indicated (5 and 15 scans) versus as a routine practice with PET/CT (48 and 127 scans).

KQ 3. Summary Differential Efficacy and Harms:

No studies were identified that met the inclusion criteria.

KQ 4. Summary of Economic Studies:

One poor quality cost-effectiveness analysis (CEA)[17] for initial staging with PET/CT in adult patients with HL and one moderate-quality cost-utility analysis (CUA)[36] for use of PET/CT for surveillance in adults with diffuse large B-cell lymphoma (DLBCL)were identified. No full economic studies evaluating the cost-effectiveness of interim PET/CT (including restaging, treatment response or post-treatment evaluation) or for other types of lymphoma were identified.

- One poor-quality CEA (QHES 17/100) evaluating PET/CT for initial staging in patients with HL was conducted in Brazil and took a payer perspective It concluded that both combined PET/CT and PET with CT separately are cost effective compared with conventional staging.
    - ICER for PET/CT: $16,215 per modified treatment case. ICER for PET with separate CT: $35,490 per modified treatment case. Authors concluded that both strategies are cost-effective in Brazil, using a threshold of 3 times GDP per capita.
    - Limitations:
        - No model specification, unclear methodology, short time horizon, no sensitivity analyses
        - Physician fees not included, source of cost data was vague

  - Denominator of ICER derived using author-created "reference standard" which was poorly defined and not well validated
  - Because this is not a standardized way of reporting cost-effectiveness, results are not comparable across studies or treatments
  - Results not generalizable to US

- One moderate quality (QHES 78/100) CUA was conducted in the US. Although authors state that they report from a societal perspective indirect costs were not included. They concluded that neither PET/CT nor CT alone is cost effective compared with routine surveillance without imaging.
  - ICER for PET/CT: $168,750 per QALY. ICER for CT alone: $164,960 per QALY. Authors concluded that neither strategy is cost effective using a threshold of $150,000.
  - Probabilistic sensitivity analysis suggests that the PET/CT is not cost-effective. The likelihood of the intervention being cost-effective is 9% at the willingness-to-pay threshold of $150,000 per QALY over a lifetime time horizon based modeling of a hypothetical cohort.
  - Limitations:
    - Used a lifetime horizon but RCT data are limited to 6 years after autologous SCT
    - Not generalizable to patients at high risk of relapse
    - Source of utility weights not specified

Strength of Evidence Summaries

The following summaries of evidence for primary outcomes have been based on the highest quality of studies available. ***Detailed SoE tables, including reasons for downgrading are found in section 5 of the report***. Additional information on lower quality studies and secondary outcomes is available in the report. Summaries for each key question are provided in the tables below and are sorted by type of lymphoma and timing of PET (e.g., interim). Details of other outcomes are available in the report.

Key Question 1: Strength of Evidence Summary: Efficacy Results

Initial Staging PET in adults (Survival outcomes)

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion	Quality (SOE)
Hodgkin Lymphoma (early-stage)					
Probability of Relapse-Free Survival	4 years	1 Pro cohort (N=37) *Figura 2017*	Risk of Bias[1] Inconsistency (Unknown[1], single study) Imprecision[1]	**Pre-treatment PET vs. no PET** 97% (NR) vs. 67% (NR) p = 0.001 **Conclusions:** 4-year RFS was greater in patients who underwent a PET prior to starting	⊕◯◯ INSUFFICIENT

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion	Quality (SOE)
				chemotherapy compared with those who did not have a PET.	
Indolent NHL (follicular lymphoma)					
Probability of Overall and Progression Free Survival	5 years	1 Retro cohort (N=206) *Friedberg 2012*	Risk of Bias[1] Inconsistency (Unknown[2] single study)	**Initial staging using PET vs. conventional methods** OS: data NR, p=NS PFS: aHR 0.87 (95% CI 0.47 to 1.62) **Conclusion**: Authors report no statistical difference in OS or PFS at 5 years when initial staging was done with PET compared with conventional methods.	⊕◯◯ INSUFFICIENT

CI = Confidence Interval; aHR = adjusted Hazard Ratio; NR = Not Reported; NS = Not significant; OS = Overall Survival; PET = Positron Emission Tomography; PFS = Progression Free Survival; Pro = prospective; Retro = retrospective; RFS = Relapse Free Survival.

Reasons for downgrading:

1. Serious risk of bias: the majority of studies violated one or more of the criteria for good quality RCT (or cohort study) related to the outcome reported (see Appendix for details).
2. Inconsistency: differing estimates of effects across trials; If point estimates across trials are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies may be downgraded. Consistency is also unknown across PET-adapted studies given the differences between studies in initial as well as randomized treatment protocols based on PET results.
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

Interim PET in adults with early-stage Hodgkin Lymphoma (Randomized PET-adapted studies)

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion Effect Estimate (95% CI)	Quality (SOE)
Escalation of therapy (Randomization of PET + to escalated versus standard treatment)					
Probability of Overall Survival	5 years	1 (N=361 PET + persons) *H10 Trial (*Andre 2017,	Consistency (Unknown[2] single study)	**Escalated treatment vs. standard*, PET +** 96.0% (91.1%–98.2%) vs. 89.3% (83.4%–93.2%); Estimated difference 6.7/100 HR 0.45 (95% CI 0.19–1.07), p=0.062 **Conclusions:** OS was slightly higher in those receiving escalated treatment compared with standard therapy; results may be of clinical	⊕⊕⊕◯ MODERATE

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion Effect Estimate (95% CI)	Quality (SOE)
		Raemaekers 2014*)*		significance but did not reach statistical significance.	
Probability of Progression Free Survival			Consistency (Unknown[2] single study)	**<u>Escalated treatment vs. standard*, PET +</u>** 90.6% (84.7%–94.3%) vs. 77.4% (70.4%–82.9%); Estimated difference 12.9/100 HR 0.42 (95% CI 0.23–0.74), p=0.002 **<u>Conclusion</u>**: PFS was better in patients receiving escalated therapy compare with those who continued with standard therapy.	⊕⊕⊕◯ MODERATE
Treatment Toxicity (GRADE 3 or 4)§			Consistency (Unknown[2] single study) Imprecision[3] (No for all except death)	**<u>Escalated treatment vs. standard*, PET +</u>** **Hematologic grade 3/4** *Neutropenia:* 53.5% vs. 30.3%; RR 1.8 (1.4–2.3); RD 23% (13% to 33%); *Febrile neutropenia:* 23.9% vs. 1.1%; RR 22.7 (5.6–92.6); RD 23% (16% to 29%) *Thrombocytopenia:* 19.7% vs. 0%; RR NC; RD 20%; p<0.0001 *Anemia:* 4.9% vs. 0%; RR NC; RD 5%; p=0.002 **Infections, grade 3/4 (without neutropenia)** 5.6% vs. 1.1%; RR 5.1 (1.1–23.3); RD 4% (1% to 8%); p=0.02 **Toxicity-related death:** 1st line protocol: 0.6% vs. 0.5% ; RD 0.1% (-1% to 2%); NS 2nd line treatment: 0% vs. 1.1% ; RD -1%;NS **<u>Conclusion</u>:** Patients receiving escalated treatment had significantly higher risk of grade 3 or4 hematologic toxicities and infections compared with those receiving standard treatment. There may have not been sufficient statistical power to compare toxicity related deaths in the two groups.	⊕⊕⊕◯ MODERATE
De-escalation of therapy (Randomization of PET (-) to de-escalated treatments vs. continuation)					
Probability of Overall Survival	3 years	1 (N=420) *RAPID trial* (Radford, 2015)	Consistency (Unknown[2] single study)	**<u>No further treatment vs. INRT†, PET *negative*-</u>** 99.0% (95% CI, 97.6%–100%) vs. 97.1% (95% CI, 94.8%–99.4%); Estimated difference 1.9/100 HR 0.51 (95% CI, 0.15%–1.68%), p=0.27 **<u>Conclusion</u>:** In persons whose interim PET was negative, the difference in OS between groups was small (1.9/100) and did not reach statistical	⊕⊕⊕◯ MODERATE

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion Effect Estimate (95% CI)	Quality (SOE)
				significance. The clinical significance of this difference is unclear.	
	5 years	1 (N=515, PET - individuals) *H10 Trial (*Andre 2017, Raemaekers 2014*)*	Consistency (Unknown[2] single study)	**<u>De-escalated (INRT omission) vs. standard†, PET *negative*-</u>** Initial status Favorable ‡ 99.6% (97.0%–99.9%) vs. 100% Estimated difference -0.4/100 Unfavorable‡ 98.3% (96.0% –99.3%) vs. 96.7% (93.7% – 98.3%); Estimated difference 1.6/100 Across favorable and unfavorable (estimated): 99% vs. 98%, difference 1/100 **<u>Conclusion</u>:** In persons whose interim PET was negative, for 5-year OS there wereno statistical differences between those in whom INRT was omitted vs. those who continued with standard treatment. The clinical significance of the small differences is unclear.	
Probability of Progression Free Survival	3 years	1 (N=420) *RAPID trial* (Radford, 2015)	Consistency (Unknown[2] single study)	**<u>No further treatment vs. INRT†, PET *negative*</u>** 90.8% (86.9%–94.8%) vs. 94.6% (91.5%–97.7%); HR 1.57 (95% CI, 0.84%–2.97%) p=0.16; Absolute RD 3.8% (95% CI -1.3 to 8.8) **<u>Conclusion</u>:** In persons whose interim PET was negative, 3.8/100 fewer patients who received no further treatment had PFS vs. those who had had the standard treatment with INRT. The difference may be clinically important.	⊕⊕⊕◯ MODERATE
	5years	1 (N=515, PET - individuals) *H10 Trial (*Andre 2017, Raemaekers 2014*)*	Consistency (Unknown[2] single study) Imprecision (-1)[3] (for favorable)	**<u>De-escalated (INRT omission) vs. standard†, PET *negative*</u>** Favorable‡ 87.1% (82.1%–90.8%) vs. 99.0% (95.9% –99.7%); HR, 15.8; 95% CI, 3.8 to 66.1) in favor of ABVD + INRT; Estimated difference : -11.9/100 Unfavorable‡ 89.6% (85.5% –92.6%) vs. 92.1% (88.0% –94.8%); HR 1.45 (95% CI, 0.8 to 2.5); estimated difference: - 2.5/100 Across favorable and unfavorable (estimated): 88% vs. 95%; difference -7/100 **<u>Conclusion</u>:** In persons with initially favorable prognosis whose interim PET was negative, PFS at 5 years was better in those who received standard therapy (ABVD + INRT) compared with the those receiving the de-escalated regimen of ABVD alone, (difference 11.9/100) however,	⊕⊕⊕◯ MODERATE

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion Effect Estimate (95% CI)	Quality (SOE)
				wide confidence interval suggests a lack of precision for the estimate. In the unfavorable group, the difference between treatment strategies was smaller (2.5/100). Across prognostic groups, PFS was better with standard versus de-escalated therapy (difference -7/100). The differences may be clinically important.	
Treatment Toxicity (Grade 3 or 4)§		2 trials (N=935, PET - individuals) *H10 Trial (*Andre 2017, Raemaekers 2014*)* (N=420) *RAPID trial* (Radford, 2015)	Consistency (Unknown[2]) Imprecision (-2)[3]	**RAPID trial: No further treatment vs. INRT†; PET negative** *Death due to secondary malignancy:* 1.4% vs. 0.5%; RR 3.0 (0.3–28.3); RD 1% (-1% to 3%) **H10 trial: INRT omission vs. standard†, PET negative** Initial status Favorable‡ *Frequency of second malignancy:* 2.9% vs. 1.3% RR 2.2 (0.6–8.5); RD 2% (-1% to 4%), Unfavorable‡ *Frequency of second malignancy:* 3.0% vs. 3.4%; RR 0.9 (0.4–2.1); RD -0.4% (-3% to 2%); **Conclusion:** Evidence on the impact of PET-adapted de-escalation strategies on treatment-related toxicities is limited and power to detect rare events and differences between strategies was likely limited. Overall, there were no statistical differences between de-escalated strategies and standard treatments. Follow-up time may have been insufficient to capture some RT-related toxicities.	⊕⊕◯◯ LOW

CI = Confidence Interval; HR = Hazard Ratio; INRT = Involved Node Radiation Therapy; NC = Not calculable; NR = Not Reported; NS = Not significant; OS = Overall Survival; PET = Positron Emission Tomography; PFS = Progression Free Survival; R-ACVBP = rituximab, doxorubicin, cyclophosphamide, vindesine, bleomycin, and prednisone; R-CHOP = rituxan, cyclophosphamide, doxorubicin, vincristine, and prednisone; RD = Risk Difference; RR = Risk Ratio

*Escalated treatments: H10 trial: escalated treatment, 2 eBEACOPP cycles + involved-node radiotherapy (INRT) vs. standard arm consisted of ABVD followed by INRT)

†De-escalation of treatment: RAPID trial: Intervention was no further treatment, comparator was involved-field radiotherapy; H10 trial: Treatment strategies for "favorable" were omission of INRT (i.e. 2 additional cycles of ABVD alone) vs. standard treatment (1 additional ABVD cycle followed by INRT; for "unfavorable" were 4 additional ABVD cycles vs. 2 additional ABVD cycles followed by INRT

‡ Refers to prognostic status at initial staging: Favorable status indicates age <50 years with ≤3 involved nodal areas, absence of mediastinal bulk (mediastinum-to-thorax ratio <0.35), and erythrocyte sedimentation rate (ESR) <50 mm without B symptoms or ESR <30 mm with B symptoms; Unfavorable status indicates age ≥50 years, >4 involved nodal areas, presence of mediastinal bulk (mediastinum-to- thorax ratio ≥0.35), or ESR ≥50 mm without B symptoms or ESR ≥30 mm with B symptoms.

§Grading of Adverse Events (AE) is according to the Common Terminology Criteria for Adverse Events (CTCAE) v3.0. Grades of 1-5 refer to the severity of the AE. Grade 1 Mild AE, Grade 2 Moderate AE, Grade 3 Severe AE, Grade 4 Life-threatening or disabling AE, Grade 5 Death related to AE.

Reasons for downgrading:

1. Serious risk of bias: the majority of studies violated one or more of the criteria for good quality RCT (or cohort study) related to the outcome reported (see Appendix for details) None of the included randomized controlled trials for this report randomized patients to receive PET/CT versus a strategy that did not include use of PET/CT, but rather, patients were randomized to various treatment options based on results of their PET/CT. There was no downgrade of these studies based on lack of comparison to a no PET/CT strategy. Risk of bias was assessed based on criteria in Appendix D and any down grade for risk of bias was based on these criteria.
2. Inconsistency: differing estimates of effects across trials; If point estimates across trials are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies may be downgraded. Consistency is also unknown across PET-adapted studies given the differences between studies in initial as well as randomized treatment protocols based on PET results.
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

Interim PET in adults with advanced-stage Hodgkin Lymphoma (Randomized PET-adapted studies)

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion Effect Estimate (95% CI)	Quality (SOE)
Escalation of therapy (Randomization of PET + to different escalated treatments)*					
Probability of Overall Survival	3 years	2 (N=2232) *GITIL/FIL HD 0607 trial* (Gallamini 2018, N=788) *GHSG HD18 trial* (Borchmann 2017b, N=1444)†	Consistency (Unknown[2])	**Escalated treatments*, PET + (n= 582)** HD0607 trial: 89% (79%–95%) vs. 90% (78%–95%); Estimated difference: -1/100 HD18 trial: 95.6 (92.8%–98.4%) vs. 97.1% (94.8%–99.4%); Absolute RD –1.4% (95% CI –5.1% to 2.2%) **PET – (n= 1640)** HD0607 trial (n=630): 99% HD18 trial (n = 1010): 97.3% **PET + (overall) vs. PET -** HD0607 trial: 89% vs. 99%; p<0.001 HD18 trial: 95.6% and 97.1% vs. 97.3%;p= NR **Conclusions:** The probability of OS was similar for escalated treatments in each trial for those with a positive interim PET with a small difference between treatments (1/100 to 1.4/100). Prognosis: The HD0607 trial reported lower OS in PET+ vs. PET – groups, less difference between PET + and PET – was seen in HD18 suggesting somewhat similar prognosis for PET + and PET – in that trial. Somewhat higher OS was seen in the HD18 trial. Differences in treatments and patient populations may contribute to this.	⊕⊕⊕◯ MODERATE
	5 years	1 (N=1444)	Consistency	**Escalated treatments*, PET + (n=434)**	⊕⊕⊕◯

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion Effect Estimate (95% CI)	Quality (SOE)
		GHSG HD18 trial (Borchmann 2017b)†	(Unknown[2] single study)	HD18 trial: 93.9% vs. 96.4%;HR 1.62 (95% CI, 0.70–3.75, eBEACOPP vs. escalation with R-eBEACOPP; Absolute RD –2.5% (95% CI –6.8% to 1.7%) **PET – (n=1010)** HD18 trial: 96.3% **PET + vs. PET -** HD18 trial: 95.5% vs. 96.3%; p=0.49 Estimated difference: -0.8/100 **Conclusion**: OS was not significantly lower for the R-BEACOPP group versus the BEACOPP group. OS was similar across escalated treatment groups in those who were PET + and when compared with those with negative interim PET.	MODERATE
Probability of Progression Free Survival	3 years	2 (N=2232) *GITIL/FIL HD 0607 trial*(Gallamini 2018, N=788) *HD18 trial* (Borchmann 2017b, N=1444)†	Consistency (Unknown[2])	**Escalated treatments*, PET + (n = 582)** HD0607 trial: 63% (50%–74%) vs. 57% (45%–68%); p 0.534; estimated difference 6/100 HD18 trial: 93.2% (89.8%–96.7%) vs. 92.7% (89.1%–96.2%); Absolute RD 0.6% (95% CI– 4.4% to 5.5%) **PET – (n= 1640)** HD0607 trial:87%; HD18 trial: 93.5% **PET + vs. PET -** HD0607 trial: 60% vs. 87%; p<0.001 HD18 trial: 93.2% and 92.7%; vs. 93.5%; p=NR **Conclusion:** PFS was similar across treatment group for PET + group in the HD 18 trial. Although PFS was higher in for escalated treatment with R-BEACOPP in the HD0607 trial vs. BEACOPP, the difference may not be beyond what may be expected by chance. While the HD0607 trial reported lower PFS in PET + vs. PET – groups, no difference between PET + and PET – was seen in HD18.	⊕⊕⊕◯ MODERATE
	5 years	1 (N=1444) *GHSG HD18 trial* (Borchmann 2017b)†	Consistency (Unknown[2] single study)	**Escalated treatments*, PET + (n= 434)** HD18 trial: 88.1% (83.5%–92.7%) vs. 89.7% 89.7% (85.4%–94.0%); HR 1.25 (95% CI, 0.69–2.26 , eBEACOPP vs. escalation with R-eBEACOPP); Absolute RD –1.6% (95% CI –7.9% to 4.7%) **PET – (n=1010)** HD18 trial: 91.4% **PET + vs. PET -** HD18 trial: 88.3 vs. 91.4%; p=0.30 **Conclusion:** PFS was similar across treatment groups in those who were PET +.	⊕⊕⊕◯ MODERATE
Treatment Toxicity	NR	2 (N=2232, total)	Consistency (Unknown[2])	**Escalated treatments*, PET + (n= 434)** **HD18 trial:** R-eBEACOPP vs. eBEACOPP	⊕⊕⊕◯ MODERATE

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion Effect Estimate (95% CI)	Quality (SOE)
(GRADE 3 or 4)††		*FIL HD 0607 trial* (Gallamini 2018, N=788) *HD18 trial* (Borchmann 2017a and b, N=1444)†		• *Any Mortality:* 6.5% (14/217) vs. 4.1% *(9/217); RD 2% (-2% to 7%)* • *Acute toxicity:* 97% (213/220) vs. 98% (213/218); RD -1% (-4% to 2%) • *Second malignancy (any):* 3.7% (8/217) vs. 4.6% (10/217); 5-year cumulative incidence estimate‡: 3.5% (95% 0.7%-6.3%) vs. 4.0% (95% CI 1.3%-6.7%) **HD 0607 trial**: No adapted-therapy specific data; • *Death (PET+ vs. PET-): 11% (16/150) vs. 2% (12/630)* • *Toxicity (PET+ vs. PET-)* (Grade 3 or4) ▪ Hematologic toxicity: 76% vs. 30% ▪ Infections: 10% vs. NR ▪ Pulmonary toxicity: NR vs. 2% **Conclusion:** The HD18 trial reported similar proportions of patients experiencing grade 3 and 4 treatment-related toxicities for the escalated treatment groups. The HD 0607 trial reports that fewer toxic events occurred in persons with PET negative versus PET positive only.	
De-escalation of therapy (Randomization of PET - to de-escalated treatments vs. continuation)‡					
Probability of Overall Survival	3 years	2 (N=2551, total) *HD18 trial* (Borchmann 2017b, N=1444)† *RATHL trial* (Johnson 2016, N=1107)	Consistency (Unknown[2])	**De-escalated treatments‡, PET *negative* (n = 1940)** HD18 trial: 98.8% (97.8%–99.9%) vs. 95.7 (93.7%–97.6%); HR 0.32 (0.14 to 0.72, 6-8 cycles vs. 4 cycles); p= 0.0037; Absolute RD: 3.1% (95% CI 0.9% to 5.3%) RATHL trial: 97.6% (95.6%– 98.7%) vs. 97.2% (95.1%– 98.4%); HR 0.09 (95% CI, 0.47–1.74); Estimated difference: 0.4/100 **PET + (n = 1120)** HD18 trial: 97.1% (95.9%–98.2%) RATHL trial: 87.8% (81.5%–92.1%) **PET + vs. PET -** HD18 trial: 97.1% vs. [98.8% and 95.7]; p=NR RATHL trial: 87.8% vs. [84.45 and 85.7%];p=NR **Conclusion:** In persons whose interim PET was negative, the absolute difference in OS between the de-escalated group versus those	⊕⊕⊕◯ MODERATE

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion Effect Estimate (95% CI)	Quality (SOE)
				who received additional treatment cycles in the HD18 trial was 3.1%, which may be clinically important; OS was similar between those receiving de-escalated treatment and those who continued with standard therapy in the RATHL trial. De-escalation of treatment did not adversely impact OS.	
	5 years	1 (N=1444, total) *GHSG HD18 trial* (Borchmann 2017b)†	Consistency (Unknown[2])	**De-escalated treatments‡, PET *negative* (n=1005)** HD18 trial: 97.7% (4 cycles) vs. 95.4% (6-8 cycles); Absolute RD: 2.3% (95% CI –0.2% to 4.9%); **PET + (n=948)** HD18 trial: 95.5% (93.9% to 97.1%) **PET + vs. PET -** HD18 trial: 95.5% vs. 96.3%p =0.49 **Conclusion:** OS was similar at 5 years for those whose treatment was de-escalated and those who continued on more intense treatment. OS between those with PET + and PET – at the interim time point were also similar.	⊕⊕⊕◯ MODERATE
Probability of Progression Free Survival	3 years	2 (N=2551) *HD18 trial* (Borchmann 2017b, N=1444)† *RATHL trial* (Johnson 2016, N=1107)	Consistency (Unknown[2])	**De-escalated treatments‡, PET *negative* (n= 1940)** **HD18 trial**: 95.3% (93.3% to 97.3%) vs. 91.7% (89.0% to 94.4%); Absolute RD: 3.6% (95% CI –0.2% to 6.9%) **RATHL trial:** 84.4% (80.97.5%) vs. 85.7% (82.1%–88.6%), HR 1.13 (0.81, 1.5, ABVD vs. AVD); Absolute RD 1.6% (95% CI –3.2 to 5.3); **PET + (n= 1120)** HD18 trial: 92.5% (90.7%–94.3%) RATHL trial: 67.5% (59.7%–74.2%) **PET + vs. PET -** HD18 trial: 92.5% vs. [95.3% and91/7%] RATHL trial: 67.5% vs. [84.4% and 85.7%] **Conclusion:** In persons whose interim PET was negative, PFS was statistically similar in groups receiving de-escalated treatment versus those with more intensive therapy in both trials. In the HD 18 trial, PFS was slightly greater in those receiving de-escalated treatment. Thisindicates that PFS was not compromised with de-escalation of treatment.	⊕⊕⊕◯ MODERATE
	5years	1 (N=1444) *GHSG HD18 trial*	Consistency (Unknown[2])	**De-escalated treatments‡, PET *negative* (n=1005)** HD18 trial: 92.2% (4 cycles) vs. 90.8% (6-8cycles)	⊕⊕⊕◯ MODERATE

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion Effect Estimate (95% CI)	Quality (SOE)
		(Borchmann† 2017b)		HR 0.79 (0.50 to 1.24, 6-8 cycles vs. 4 cycles); Absolute RD:–1.4% (95% CI -2.7% to 5.4%) (6-8 cycles vs. 4 cycles) **PET + (n=948)** HD18 trial: 88.3% (85.8%–90.8%) **PET + vs. PET -** HD18 trial: 88.3% vs. [92.2% and 90.8%]; p=0.30 **Conclusion:** In persons whose interim PET was negative, the difference in PFS between groups was small (1.4/100), not statistically significant and of unclear clinical significance. No difference in PFS was seen between those whose interim PET was positive and those whose results were negative.	
Treatment Toxicity (GRADE 3 or 4) ††		2 (N=2551) *HD18 trial* (Borchmann 2017b, N=1444) RATHL Trial (Johnson 2016, N=1107)	Consistency (Unknown[2])	**De-escalated treatments‡, PET *negative*** **HD18 trial (n=1005):** 4 vs. 6 to 8 eBEACOPP cycles • *Any second malignancy:* 2.6% vs. 3.6%; 5-year cumulative incidence estimate§: 3.3% (95% CI 1.4%-5.3%) vs. 3.8% (95% CI 1.9% to 5.7%) • *Discontinuation due to toxicity:* 0.4% vs. 2.2%; RD -2% (-3% to -0.4%) • *Any acute toxicity*: 90.8% vs. 96.2%; RD -5% (-8% to -2%); • *Treatment-related morbidity*: 40.7% vs. 63.7%; RR 0.6 (0.56–0.72); RD -23% (-29% to -17%) • *Any organ toxicity*: 7.6% vs. 18.1%; RD -10% (-15% to -6%); • *Anemia:* 38.9% vs. 54.4%; RD-15% (95%CI -22% to -9%) • *Thrombocytopenia*: 57.1% vs. 71.8%; RD -15% (95%CI -21% to -9%) • *Leukopenia:* 87.4% vs. 92.7%; RD -5% (95%CI -9% to -1.5% • *Infection:* 8.0% vs. 14.9%; RD -7% (-11% to -3%);	⊕⊕⊕◯ MODERATE

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion Effect Estimate (95% CI)	Quality (SOE)
				• *Febrile neutropenia*: 21.8% vs. 28.8% RR 0.8 (0.6–0.9); RD -7% (-12% to -2%); • *Need for supportive measures*: Platelet infusion 24.4% vs. 40.5% (RD -16% 95% CI -22% to -10%); RBC infusion 46.5% vs. 65.9% (RD -19% (95%CI -25% to -13%) **RATHL trial ((n = 935)** • *Mortality from second malignancy: 1.3% vs. 0.9%; RD 0.4% (95%CI -1% to 2%)* • *Any second malignancy: 2.4% vs. 2.8%; RD -0.4% (95%CI -2% to 2%);* • *Any clinical adverse event**: 21% vs. 30.6%; RD -10% (95%CI -15% to -4%);* • *Any acute toxicity*: 65.4% vs. 68.8; RD -3% (-9% to 3%); • *Hematologic toxicities*: 59.7% vs. 59.8%; RD -0.1% (-6% to 6%) • *Thrombocytopenia*: 3.3% vs. 1.3%; RD 2% (0.1%–4%); • *Infection*: 10.3% vs. 14.5%; RD -4% (-8% to 0%); • *Febrile neutropenia:* 2.2% vs. 4.7%; RD -3% (-5% to -0.2%); • *Neurological event*: 3.1% vs. 4.9%; RD -2% (-4% to 1%) • *Pulmonary/respiratory event*: 0.7% vs. 3.2%; (RD -3% (-4% to -1%); • *Vascular event*: 2.6% vs. 4.9%; RD -2% (-5% to 0.2%) • *Fatigue*: 1.1% vs. 3.0%; RD -2% (-4% to -0.1%) **Conclusion**: Overall, both trials report that most treatment-related toxicities were less common in patients who received de-escalated therapy compared with those who	

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion Effect Estimate (95% CI)	Quality (SOE)
				continued on more intense regimens; some were substantially less common.	

CI = Confidence Interval; eBEACOPP = escalated Bleomycin, Etoposide, Doxorubicin, Cyclophosphamide, Vincristine, Procarbazine, Prednisolone; HR = Hazard Ratio; NR = Not Reported; OS = Overall Survival; PET = Positron Emission Tomography; PFS = Progression Free Survival; RD = Risk Difference

* Escalated treatments: HD 0607 trial R-BEACOPP vs. BEACOPP after 2 ABVD cycles; HD18Trial R-eBEACOPP vs. eBEACOPP after 2 eBEACOPP cyles

†Borchmann: Only OS and PFS data from randomized patients is reported here; this includes patients randomized prior to the protocol amendment (i.e., who received 8 total cycles) which introduced a reduction of standard therapy from 8 to 6 cycles of eBEACOPP in total. Because the required sample size for the treatment group comparison in the PET(+) cohort had already been reached, randomization between eBEACOPP and R-eBEACOPP after PET(+) was stopped at this point, and all patients with a PET(+) were subsequently assigned to receive the new standard treatment of 6 × eBEACOPP.

‡De-escalation of treatment: RAHTL trial bleomycin was omitted vs. included per standard ABVD after 2 ABVD cycles; HD18Trial 2 cycles of eBEACOPP vs.6- 8 cycles of eBEACOPP after initial 2 eBEACOPP cycles

§ Accounting for death as competing risk.

**excluding blood or bone marrow events and laboratory events

††Grading of Adverse Events (AE) is according to the Common Terminology Criteria for Adverse Events (CTCAE) v3.0. Grades of 1-5 refer to the severity of the AE. Grade 1 Mild AE, Grade 2 Moderate AE, Grade 3 Severe AE, Grade 4 Life-threatening or disabling AE, Grade 5 Death related to AE.

Reasons for downgrading:

1. Serious risk of bias: the majority of studies violated one or more of the criteria for good quality RCT (or cohort study) related to the outcome reported (see Appendix for details) None of the included randomized controlled trials for this report randomized patients to receive PET/CT versus a strategy that did not include use of PET/CT, but rather, patients were randomized to various treatment options based on results of their PET/CT. There was no downgrade of these studies based on lack of comparison to a no PET/CT strategy. Risk of bias was assessed based on criteria in Appendix D and any down grade for risk of bias was based on these criteria.
2. Inconsistency: differing estimates of effects across trials; if point estimates across trials are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies were downgraded. Consistency is also unknown across PET-adapted studies given the differences between studies in initial as well as randomized treatment protocols based on PET results.
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from “mild” to “substantial”. If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for “mild/small” effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

Interim PET in adults with relapsed or refractory Hodgkin Lymphoma (Observational PET-adapted studies)

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion	Quality (SOE)
Overall and (based on pre-ASCT PET)	Various	3 Observational *Moskowitz 2012 (N = 97) 2015 and 2017 (N = 65*)	Risk of Bias (-1)[1] Consistency (Unknown[2])	**2012 study (4.25 years)** OS: all patients 80%, ASCT 88%; data by PET/CT results are not provided **2015 (3.3 years)/2017 (2 years) cohorts;** OS all patients: 95% **Conclusions:** No firm conclusions regarding the impact of PET-adapted therapy are drawn.	⊕◯◯◯ Insufficient
Event Free Survival (based on pre-ASCT PET)				**2012 study (4.25 years) patients with ASCT** EFS: All patients 70%, ASCT, 70% EFS: PET + 28.6% vs. PET - >80% **2015 (3.3 years)/2017 (2 years) cohorts;** EFS all patients: 82% EFS: PET + 60% vs. PET – 85% (estimated)† **Conclusions:** Pre-ASCT PET positive status was associated with worse prognosis compared with PET negative status.	⊕◯◯◯ Insufficient
Treatment-related Toxicity‡	NR	2 Observational *Moskowitz 2015/2017 (N = 65)*	Consistency (Unknown[2])	Conclusions: Treatment-related toxicities were not reported based on PET-adapted treatments received. Overall, the reported frequency of grade 3 or 4 toxicities was ≤2%. No conclusions regarding the impact of PET on toxicities are possible.	⊕◯◯◯ Insufficient

ASCT = Autologous Stem Cell Transplantation; EFS= event free survival; NR = Not Reported; OS = Overall Survival; PET = Positron Emission Tomography

* PET Positive: extended salvage chemotherapy of GVD or augICE followed later by HDT, RT and ASCT; PET negative directly to consolidation with high dose chemotherapy (HDT) and ASCT, with the potential for RT

†Estimated from author figures for patients receiving ASCT.

‡ Grading of Adverse Events (AE) is according to the Common Terminology Criteria for Adverse Events (CTCAE) v3.0. Grades of 1-5 refer to the severity of the AE. Grade 1 Mild AE, Grade 2 Moderate AE, Grade 3 Severe AE, Grade 4 Life-threatening or disabling AE, Grade 5 Death related to AE.

Reasons for downgrading:

1. Serious risk of bias: the majority of studies violated one or more of the criteria for good quality RCT (or cohort study) related to the outcome reported (see Appendix for details) None of the included randomized controlled trials for this report randomized patients to receive PET/CT versus a strategy that did not include use of PET/CT, but rather, patients were randomized to various treatment options based on results of their PET/CT.
2. Inconsistency: differing estimates of effects across trials; If point estimates across trials are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was downgraded. Consistency is also unknown because of overlap of study populations (2015 and 2017) and use of the similar/the same treatment protocols.
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

Interim PET in adults with aggressive Non-Hodgkin Lymphoma (PET-adapted studies)

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion	Quality (SOE)
Escalated therapy					
PET-Adapted treatment RCT (Dührsen 2018)					
Probability of Overall Survival	2 years	1 (N=363, 60% stage III-IV, 86.4% DLBCL) *PETAL trial* (Dührshen 2018) (ΔSUV_{max} method)	Consistency (Unknown[2] single study)	**PET +: Burkitt protocol) vs. R-CHOP (n=108)** 55.4% vs. 63.6%; HR 1.3 (95% CI 0.8 to 2.4); Estimated difference: 7.9/100 **PET – : 4 R-CHOPP cycles vs. 4 R-CHOPP + 2 additional rituximab (n=255)** 87.2% vs. 88.2%; HR 0.9 (95% CI 0.5–1.5); Estimated difference: 1/100 **Conclusion:** OS probability was lower in those with positive interim PET receiving the Burkitt's protocol versus additional R-CHOP. Results are not statistically significant but may be clinically important. Failure to reach the recruitment goal for PET+ patients may have impacted statistical power to detect survival differences. In those with negative interim PET, OS was similar between treatments; the additional doses of rituximab did not substantially improve OS.	⊕⊕⊕◯ MODERATE
Probability of Progression Free Survival and EFS			Consistency (Unknown[2] single study)	**PET +: Burkitt protocol)vs. R-CHOP (n=108)** **PFS:** 43.1% vs. 49.4% Estimated difference:- 6.3/100 **EFS:** 31.5% vs. 42.0%; Adj. HR 1.5 (95% CI 0.9 to 2.5); Estimated Difference: 10.5/100 **PET – 4 R-CHOPP cycles vs. 4 R-CHOPP + 2 additional rituximab (n=255)** **PFS**: 77.5% vs. 82.0%; Estimated difference: -4.5/100 **EFS**: 73.5 vs. 76.4; Adj. HR 1.0 (95% CI 0.7 to 1.6), Estimated difference: -2.9/100 **Conclusion:** PFS and EFS were lower in PETpositive patients receiving the escalated treatment of Burkitt's protocol compared with those receiving additional R-CHOP; these were not statistically different but may be clinically important. In PET-negative patients, PFS and EFS were somewhat lower in those who received versus those who did not	⊕⊕⊕◯ MODERATE

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion	Quality (SOE)
				receive the additional rituximab doses. These differences may be clinically important;	
Treatment Toxicity (Grade 3 or4)†			Consistency (Unknown[2] single study) Imprecision (-1)[3]	**PET + (escalated) vs. R-CHOP (n = 108** *Treatment-related mortality*: 5.4% vs. 3.9%; (RD 2% (-6% to 9%) *Anemia:* 44.6% vs. 25.0%; RD 20% (2% to 37%); *Leukopenia: 80.4% vs. 59.6%;* RD 21% (4% to 38%); *Neutropenia: 33.9% vs. 23.1%;* RD 11% (-6% to 28%); *Thrombocytopenia: 58.9% vs. 21.2%;* RD 38% (21% to 55%) *Infection: 50% vs. 21.2%:* RD 29% (12% to 46%); *Mucositis: 37.5% vs. 11.5%;* RD 26% (11% to 41%) *Diarrhea:* 5.4% vs. 9.6%; RD -4% (-14% to 6%) *Creatinine:* 1.8% vs. 5.8%; RD -4% (-11% to 3%) **PET – 4 R-CHOPP cycles vs. 4 R-CHOPP + 2 additional rituximab (n = 255)** *Treatment-related mortality:* 3.9% vs. 1.6%; RD 2% (-2% to 6%); *Anemia:* 16.3% vs. 11.1%; RD 5% (-3% to 14%) *Leukopenia:* 54.3% vs. 60.3%; RD -6% (-18% to 6%) *Neutropenia:* 15.5% vs. 23.8%; RD -8% (-18% to 1%) *Thrombocytopenia:* 14.7% vs. 7.9%; RD 7% (-1% to 15%) *Infection: 14.7% vs. 11.1%;* RD 4% (-5% to 12%) *Mucositis:* 1.6% vs. 2.4%; RD -1% (-4% to 3%); *Diarrhea:* 3.1% vs. 0.8%; RD 2% (-1% to 6%) *Creatinine:* 3.1% vs. 0.8%; RD 2% (-1% to 6%) **Conclusions:** **PET +:** Escalated treatment with the Burkitt protocol was associated with an increased risk of a number of treatment related toxicities. This arm was discontinued early and there may not have been sufficient power to detect differences between treatments for rare events (e.g. treatment-related death).	⊕⊕◯◯ LOW

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion	Quality (SOE)
				PET –: Compared with those who had R-CHOP only, those receiving additional doses of rituximab to R-CHOP may have had clinically important higher risk of some treatment-related toxicities (leukopenia, neutropenia), but lower risk of others (thrombocytopenia, infection). Associations may not be beyond what might be expected by chance.	
RCT of induction treatment (Casasnovas 2017)					
Probability of Overall Survival	4 years	1 (N=211, 97% stage III-IV) *LNH2007-3B trial* (Casanovas 2017) *Initial therapy randomized:* R-ACVBP vs. R-CHOP14 **Interim PET at the end of 2 (PET 2) and 4 cycles (PET4);** IHP criteria for positivity	Risk of Bias (-1)[1] Consistency (Unknown[2] single study)	**Escalated (PET 2+, PET4 -) vs. SIC (PET 2- and 4 -)** (n=48): 90.4% (81%-95.1%) vs. (n=52): 89.6% (85%-92.2%); p = 0.21 Estimated difference: 0.8/100 **PET 4 + (overall) vs. PET 4 - (overall)** (n=100): 80% (69%-87.5%) vs, (n=100): 88.9% (82.1%-94.4%); p=0.08 Estimated difference: 8.9/100 **Conclusions:** OS was similar between those who received escalated therapy and those who had standard immunochemotherapy (SIC) consolidation suggesting that SCI may be sufficient in patients with a negative interim PET after 4 cycles, even if PET after 2 cycles was positive. With regard to prognosis, lower OS (estimated difference of ~9/100) in patients with a positive PET 4 scan compared with those whose PET 4 scan was negative was seen.	⊕⊕◯◯ LOW
Probability of Progression Free Survival			Risk of Bias (-1)[1] Consistency (Unknown[2] single study) Imprecision (-1)[3]	**Escalated (PET 2+, PET4 -) vs. SIC (PET 2- and 4 -)** (n=48): 85% (71.1%-92.6%) vs. (n=52) 75% (60.9%-84.6%); Estimated difference: 10/100 **PET 4 + (overall) vs. PET 4 - (overall)** (n=100): 72.9% (63.1%-80.6%) vs. 79.8 (79.4-86.4); Estimated difference 6.9/100 **Conclusion**: The probability of PFS was higher in those who received escalated therapy versus those who had standard immunochemotherapy (SIC) consolidation may suggest improved PFS with escalated treatment.	⊕⊕◯◯ LOW

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion	Quality (SOE)
				Prognosis for those who had PET 4 + results was slightly lower vs. PET 4 -; the estimated difference was 6.9/100.	
Treatment Toxicity	NR		Risk of Bias (-1)[1] Consistency (Unknown[2] single study) Indirectness (-1)[4] (outcomes by PET-adapted strategy not reported) Imprecision (-1)[3]	No toxicity data relevant to impact of PET on treatment decisions was provided in the randomized trial of initial treatment. Authors compare toxicities for R-ACVBP and R-CHOP noting that the former was associated with more frequent infections (64% vs. 27%) and need for transfusions versus R-CHOP. Given that patients in both PET + and PET – groups received each of these treatments initially, specific treatment regimens for those with a positive PET 4 scan were not provide and no toxicity data on use of escalated treatment strategies were provided, conclusions related to the impact of PET on reducing toxicity from this trial are not possible.	⊕◯◯◯ Insufficient
RCT Pet-adapted treatment; initial randomization continuation (Lamy 2018)					
Probability of Overall Survival	5 years	1 (N=281, negative interim PET) *LYSA/ GOELAMS trial (Lamy 2018)* (low risk, Stage I or II w/o bulky mass)	Consistency (Unknown[2] single study)	**R-CHOP vs. R-CHOP + RT** **PET negative (n = 281)** 90% vs. 92% (estimated difference 2/100) (NS, no data provided) **Conclusion:** In persons with a negative interim PET, OS was similar between groups; it is unclear if the estimated 2/100 difference is clinically important. Deletion of RT may not adversely impact OS.	⊕⊕⊕◯ **Moderate**

CI = Confidence Interval; DLBCL = Diffuse large B-cell lymphomaeBEACOPP = escalated Bleomycin, Etoposide, Doxorubicin, Cyclophosphamide, Vincristine, Procarbazine, Prednisolone; EFS= event free survival; HR = Hazard Ratio; IHP = International Harmonization Project; NR = Not Reported; OS = Overall Survival; PET = Positron Emission Tomography; PFS = Progression Free Survival; R-ACVBP = rituximab, doxorubicin, cyclophosphamide, vindesine, bleomycin, and prednisone; R-CHOP = rituxan, cyclophosphamide, doxorubicin, vincristine, and prednisone; RD = Risk Difference; SIC = Standard Immunochemotherapy

*PETAL Randomized treatments based in interim PET: PET + persons six blocks of an intensive Burkitt's lymphoma or six additional cycles of R-CHOP; PET – persons with CD20 positive lymphomas, four additional cycles of R-CHOP or the same treatment with two additional doses rituximab.

LNH2007-3B trial: Treatment escalated in all patients with positive PET at 2 and/or 4 cycles. PET (+) cycle 2/PET (–) cycle 4, received 2 cycles of high-dose methotrexate, then BEAM or Z-BEAM followed by ASCT; for patients with positive PET at cycle 4, final treatment decision left to local investigator. Individuals with negative PET results at both time frames, received standard immunochemotherapy (SIC) consolidation.

–

LYSA/GOELAMS Trial: Patients were randomized after initial staging to receive R-CHOP or R-CHOP with RT; those with negative interim PET (complete remission) continued in these random assignments; Authors ITT analysis includes all patients as randomized regardless of interim PET results (38 patients had positive interim PET) are included in the full report.

†Grading of Adverse Events (AE) is according to the Common Terminology Criteria for Adverse Events (CTCAE) v3.0. Grades of 1-5 refer to the severity of the AE. Grade 1 Mild AE, Grade 2 Moderate AE, Grade 3 Severe AE, Grade 4 Life-threatening or disabling AE, Grade 5 Death related to AE.

Reasons for downgrading:

1. Serious risk of bias: the majority of studies violated one or more of the criteria for good quality RCT (or cohort study) related to the outcome reported (see Appendix for details) None of the included randomized controlled trials for this report randomized patients to receive PET/CT versus a strategy that did not include use of PET/CT. There was no downgrade of these studies based on lack of comparison to a no PET/CT strategy. Risk of bias was assessed based on criteria in Appendix D and any down grade for risk of bias was based on these criteria. Casasnovas and Lamy did not randomize patients based on PET findings.
2. Inconsistency: differing estimates of effects across trials; If point estimates across trials are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies were downgraded. Consistency is also unknown across PET-adapted studies given the differences between studies in initial as well as randomized treatment protocols based on PET results.
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

<u>End of Treatment</u> PET in adults with <u>advanced-stage HL and aggressive NHL</u> (Randomized and Observational PET-adapted studies)

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion	Quality (SOE)
Advanced-stage HL					
Probability of Overall Survival	3 years	1 (N=296)* *GITIL/FIL HD 0607 trial* (Gallamini 2018)	Risk of Bias (-1)[1] Consistency (Unknown[2] single study)	**<u>Consolidation RT or no RT (PET -, with baseline large nodal mass)*</u>** HD0607: 100% vs. 99% (95%-100%), Estimated difference: 1/100 **<u>Conclusions</u>:** Consolidation RT did not provide substantial clinical benefit in terms of OS.	⊕⊕◯◯ LOW
Probability of Progression Free Survival			Risk of Bias (-1)[1] Consistency (Unknown[2] single study)	**<u>Consolidation RT or no RT (PET -, with baseline large nodal mass, n=263)*</u>** HD0607: 97% (92%-99%) vs.93% (87%-96%) Estimated difference: 4/100 Patients without LNM at baseline not randomized to RT (n = 260): 92% (95%CI 88 to 85%) **<u>Conclusion</u>:** Although PFS was slight better in those receiving RT, the difference was not	⊕⊕◯◯ LOW

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion	Quality (SOE)
				beyond what might be expected by chance. PFS was not substantially different in those has no baseline LNM.	
Aggressive NHL (DBLCL)					
Probability of Overall Survival	2 years	1 (N=50 from positive interim PET group)† *PETAL trial* (Duhrsen 2018 (IHP criteria used)	Risk of Bias (-1)[1] Consistency (Unknown[2]) Imprecision (-1)[3]	**EOT PET positive (n =31):** Observation (n=18 asymptomatic vs. additional treatment (n=13 clinical evidence of disease)† 88.9% (95% CI, 62.4% to 97.1) vs. 33.3% (95% CI, 10.3% to 58.8%) **EOT PET (-) n = 18 vs. PET(+, persistent abnormalities, n=31)** 100% vs. 65.9% (95% CI, 45.8% to 80.0%) **Conclusion:** Prognosis with respect to OS was better in patients with a negative EOT PET versus those with persistent abnormalities; Among those with positive EOT PET, asymptomatic patients who were observed had better prognosis than those who had persistent PET findings in addition to clinical evidence of disease.	⊕◯◯◯ INSUFFICIENT
Probability of Progression Free Survival	2 years	1 (N=50 from positive interim PET group)† *PETAL trial* (Duhrsen 2018 (IHP criteria used)	Risk of Bias (-1)[1] Consistency (Unknown[2]) Imprecision (-1)[3]	**EOT PET positive (n =31):** Observation (n = 18 asymptomatic vs. additional treatment (n=13 clinical evidence of disease)† 66.7% (95% CI, 40.4% to 83.4%) vs. 26.4% (95% CI, 0.7% to 52.2%) **EOT PET (-) n = 19 vs. PET(+, persistent abnormalities, n=31)** 84.2% (95% CI, 58.7% to 94.6%) vs. 50.5% (95% CI, 31.8% to 66.6%) **Conclusion:** PFS was better in patients with a negative EOT PET versus those with persistent abnormalities; Among those with positive EOT PET, asymptomatic patients who were observed had better prognosis than those who had persistent PET findings in addition to clinical evidence of disease.	⊕◯◯◯ INSUFFICIENT

LNM = large nodal mass; RT = radiation therapy; EOT= end-of-treatment;

*HD 0607 Treatments: 296 patients who had large nodal masses at baseline and whose interim and end of treatment PET were both negative were randomized to radiotherapy or no further treatment base on PET/CT done after completion of first line therapy (6 cycles of ABVD)

† PETAL Trial treatments were not randomly assigned: End-of-treatment PET was restricted to patients with a positive interim scan who did **not** experience progression on therapy; 19 were PET(-) and 31 were PET(+) and had persistent abnormalities on EOT PET; Residual masses (CT) or residual FDG uptake were NOT considered lymphoma unless there was clinical evidence of disease. Of the 31 PET(+) patients, 18 (with no other evidence of active disease) received observation while 13 (with active disease) received additional treatment

Reasons for downgrading:

1. Serious risk of bias: the majority of studies violated one or more of the criteria for good quality RCT (or cohort study) related to the outcome reported (see Appendix for details) None of the included randomized controlled trials for this report randomized patients to receive PET/CT versus a strategy that did not include use of PET/CT, but rather, patients were randomized to various treatment options based on results of their PET/CT. There was no downgrade of these studies based on lack of comparison to a no PET/CT strategy. Risk of bias was assessed based on criteria in Appendix D and any down grade for risk of bias was based on these criteria. In the PETAL trial EOT PET-adapted treatments were not randomized. Risk of bias for a given study may differ by outcome.
2. Inconsistency: differing estimates of effects across trials; If point estimates across trials are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies were downgraded. Consistency is also unknown across PET-adapted studies given the differences between studies in initial as well as randomized treatment protocols based on PET results.
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded

Surveillance and Evaluation of Relapse PET in adults (Survival outcomes)

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion	Quality (SOE)
Hodgkin Lymphoma					
Probability of Overall Survival	5 years	1 retro cohort (N=241) *Pingali 2014*	Risk of Bias[1] Inconsistency (Unknown[2] single study)	**Routine/surveillance vs. imaging if indicated** 97% (95% CI 92% to 99%) vs. 96% (95% CI 86% to 99%), p = 0.41 **Conclusions:** OS was similar for asymptomatic patients in remission who received surveillance PET/CT and those who received it only if clinically indicated	⊕◯◯◯ INSUFFICIENT
Probability of Progression Free Survival	3, 5 years	1 retro cohort (N=368) *Dann 2014*	Risk of Bias[1] Inconsistency (Unknown[2] single study)	**Routine/surveillance vs. imaging if indicated** 3 year PFS: 93% (NR) vs. 86% (NR), p = NS 5 year PFS: 92% (NR) vs. 85% (NR), p= NS **Conclusion**: Authors report no statistical difference in PFS at 3 or 6 years when imaging was routinely done vs. for clinical suspicion of relapse	⊕◯◯◯ INSUFFICIENT

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion	Quality (SOE)
Relapsed, Refractory Aggressive NHL (DLBCL)					
Probability of Overall Survival From time of ASCT	Time from ASCT	1 retro cohort (N=45) *Epperla 2016*	Risk of Bias[1] Inconsistency (Unknown[2] single study) Imprecision[3]	**Routine/surveillance* vs. clinical follow-up/lab** Median years: from time of transplantation† 1.8 (range 0.4–12.4) vs. 1.6 (range 0.4–8.4) Median years: from time of relapse post ASCT† 1.0 (range 0.02–11.6) vs. 0.34 (range 0.01–2.6) **Conclusion**: Authors report no statistical difference in median survival; sample size is small.	⊕◯◯◯ INSUFFICIENT

ASCT = autologous stem cell transfusion; CI = confidence interval; OS = overall survival, Retro = retrospective study design; PET/CT = positron emission tomography/computed tomography * Routine CT and/or PET
†Overall survival (OS) was defined as the time from autologous hematopoietic cell transplantation to date of death or last follow-up. Post-relapse survival was defined as time from relapse (after auto-HCT) to date of death or last follow-up.

Reasons for downgrading:

1. Serious risk of bias: the majority of studies violated one or more of the criteria for good quality RCT (or cohort study) related to the outcome reported (see Appendix for details).
2. Inconsistency: differing estimates of effects across trials; If point estimates across trials are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies were downgraded. Consistency is also unknown across PET-adapted studies given the differences between studies in initial as well as randomized treatment protocols based on PET results.
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded

Key Question 2: Strength of Evidence Summary: Safety and Adverse Events Results

Safety and Adverse Events Results for PET/CT

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion	Quality (SOE)
Effective and Cumulative Effective Radiation Dose	Adults *Surveillance*: range, median 2–5 years *Treatment*: 8 months Pediatric 2 years	5 studies total N=1071 Adults (4 studies), N=972 Pediatric (1 study), N=99 Adults 1 RCT, N=300 (Picardi 2014) 2 comparative observational studies, N=186 (Guttikonda 2014, Patel 2013) 1 case series N=486 (Crowley 2016) Pediatric 1 comparative observational, N=99 (Rathore 2012)	Risk of Bias (-1)[1]	**Adults** *Effective radiation dose** Range of mean effective radiation doses for PET/CT during: • Surveillance (1 RCT, 2 observational): 14.5 mSv to 26.1 mSv • Treatment (1 observational): 26.1 mSv *Cumulative effective radiation dose* Median 69 mSv (IQR 42-118) per patient from all ionizing radiation imaging (1 case series): • PET/CT: 8% of total (~5.5 mSv) • CT with contrast: 89% of total (~61.4 mSv) **Pediatric** One study reports effective radiation doses from FDG alone are approximately 6.4 and 8.6 mSv for a 10- and 15-year-old child, respectively. Conclusion: Mean effective radiation doses from PET/CT varied across studies based on number of scans performed, length of follow-up and other factors such as patient's age. RSNA and ACR estimate the effective dose for a PET/CT is 25 mSv in adults.	⊕⊕◯◯ LOW
Additional Testing during surveillance	3–5 years	2 studies total, N=406 (adults) 1 RCT, N=300 (Picardi 2014)	Risk of Bias (-1)[1]	*Additional testing due to false positives* • *1 RCT*: High number of false positives (data not provided) obtained with PET/CT compared with US/chest radiography resulted in greater need	⊕⊕◯◯ LOW

Outcome	Follow-up	Studies	Reason for Downgrading	Conclusion	Quality (SOE)
		1 comparative observational, N=106 (Hong 2014)		for mediastinotomies: 50% (20/40) vs. 15% (6/40), p=0.04 • *1 observational*: False positives for PET/CT vs. CT: 20% (21/106) vs. 7% (7/106); leading to additional testing in 14% (15/106) vs. 6% (6/106) Conclusion: High false positive rates reported led to additional perhaps unnecessary diagnostic testing.	
Relapse identification during surveillance	4–5 years	3 studies total, N=909 (adults) 1 RCT, N=300 (Picardi 2014) 2 comparative observational, N=609 (Dann 2014, Pingali 2014)	Risk of Bias (-1)[1]	PET/CT vs. US/radiography (1 RCT) • Effectiveness at detecting relapse: 100% vs. 95.7% • Estimated radiation dose to detect one relapse: 449.5 mSv vs. 10.2 mSv (p=0.0002) Routine PET/CT imaging versus clinically-indicated imaging† (2 observational studies): • Number of imaging tests needed to diagnose/detect a single relapse: 48 and 127 vs. 5 and 15 Conclusion: US/radiography was as effective at identifying relapse with significantly lower radiation exposure. Fewer imaging tests are needed to identify a relapse when done as clinically indicated vs. as a routine practice.	⊕⊕◯◯ LOW

ACR = American College of Radiology; CHOP = cyclophosphamide, doxorubicin, vincristine, prednisolone; FDG = Fludeoxyglucose; PET/CT = positron emission tomography/computed tomography; R = rituximab; RSNA = Radiological Society of North America

*Mean effective radiation dose for FDG PET alone versus combined PET/CT for surveillance was 14.1 mSv vs. 26.1 mSv in one observational study (Patel 2013).

†Imaging done only when indicated by signs or symptoms; imaging was usually contrast enhanced CT.

Reasons for downgrading:

1. Serious risk of bias: the majority of studies violated one or more of the criteria for good quality RCT (or cohort study) related to the outcome reported (see Appendix for details) One RCT was identified that compared PET/CT with ultrasound/chest radiography.
2. Inconsistency: differing estimates of effects across trials; If point estimates across trials are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was downgraded. Consistency is also unknown because of overlap of study populations (2015 and 2017) and use of the similar/the same treatment protocols.

3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

Key Question 4: Strength of Evidence Summary: Cost-Effectiveness

One poor quality cost-effectiveness analysis (CEA)[17] for initial staging with PET/CT in adult patients with HL and one moderate-quality cost-utility analysis (CUA)[36] for use of PET/CT for surveillance in adults with diffuse large B-cell lymphoma (DLBCL)were identified. There currently is no method for evaluating and reporting strength of evidence across economic studies. Major finding are reported in the previous results section.

REFERENCES

1. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality. January 2014. Chapters available at: www.effectivehealthcare.ahrq.gov.
2. Surveillance, Epidemiology, and End Results (SEER) Program, National Cancer Institute. Cancer Stat Facts: Hodgkin Lymphoma. at https://seer.cancer.gov/statfacts/html/hodg.html.)
3. Abou-Nassar KE, Vanderplas A, Friedberg JW, et al. Patterns of use of 18-fluoro-2-deoxy-D-glucose positron emission tomography for initial staging of grade 1-2 follicular lymphoma and its impact on initial treatment strategy in the National Comprehensive Cancer Network Non-Hodgkin Lymphoma Outcomes database. Leukemia & lymphoma 2013;54:2155-62.
4. Andre MPE, Girinsky T, Federico M, et al. Early Positron Emission Tomography Response-Adapted Treatment in Stage I and II Hodgkin Lymphoma: Final Results of the Randomized EORTC/LYSA/FIL H10 Trial. J Clin Oncol 2017;35:1786-94. doi: 10.200/JCO.2016.68.6394. Epub 2017 Mar 14.
5. Angelopoulou MK, Mosa E, Pangalis GA, et al. The Significance of PET/CT in the Initial Staging of Hodgkin Lymphoma: Experience Outside Clinical Trials. Anticancer Res 2017;37:5727-36.
6. Awan UE, Siddiqui N, SaadUllah M, et al. FDG-PET scan in assessing lymphomas and the application of Deauville Criteria. JPMA The Journal of the Pakistan Medical Association 2013;63:725-30.
7. Bakhshi S, Radhakrishnan V, Sharma P, et al. Pediatric nonlymphoblastic non-Hodgkin lymphoma: baseline, interim, and posttreatment PET/CT versus contrast-enhanced CT for evaluation--a prospective study. Radiology 2012;262:956-68.
8. Barrington SF, Mikhaeel NG, Kostakoglu L, et al. Role of imaging in the staging and response assessment of lymphoma: consensus of the International Conference on Malignant Lymphomas Imaging Working Group. Journal of clinical oncology 2014;32:3048.
9. Bednaruk-Mlynski E, Pienkowska J, Skorzak A, et al. Comparison of positron emission tomography/computed tomography with classical contrast-enhanced computed tomography in the initial staging of Hodgkin lymphoma. Leukemia & lymphoma 2015;56:377-82.
10. Berkman ND, Lohr KN, Ansari M, et al. Grading the Strength of a Body of Evidence When Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville (MD)2013.
11. Borchmann P, Goergen H, Kobe C, et al. PET-guided treatment in patients with advanced-stage Hodgkin's lymphoma (HD18): final results of an open-label, international, randomised phase 3 trial by the German Hodgkin Study Group. The Lancet 2017b;390:2790-802.
12. Borchmann P, Haverkamp H, Lohri A, et al. Progression-free survival of early interim PET-positive patients with advanced stage Hodgkin's lymphoma treated with BEACOPPescalated alone or in combination with rituximab (HD18): an open-label, international, randomised phase 3 study by the German Hodgkin Study Group. Lancet Oncol 2017a;18:454-63. doi: 10.1016/S470-2045(17)30103-1. Epub 2017 Feb 22.
13. Carras S, Dubois B, Senecal D, et al. Interim PET Response-adapted Strategy in Untreated Advanced Stage Hodgkin Lymphoma: Results of GOELAMS LH 2007 Phase 2 Multicentric Trial. Clin Lymphoma Myeloma Leuk 2018;18:191-8. doi: 10.1016/j.clml.2018.01.003. Epub Jan 31.

14. Carrillo-Cruz E, Marin-Oyaga VA, de la Cruz Vicente F, et al. Role of 18F-FDG-PET/CT in the management of marginal zone B cell lymphoma. Hematol Oncol 2015;33:151-8. doi: 10.1002/hon.2181. Epub 014 Nov 19.

15. Casasnovas RO, Ysebaert L, Thieblemont C, et al. FDG-PET-driven consolidation strategy in diffuse large B-cell lymphoma: final results of a randomized phase 2 study. Blood 2017;130:1315-26. doi: 10.182/blood-2017-02-766691. Epub 2017 Jul 12.

16. Casulo C, Schoder H, Feeney J, et al. 18F-fluorodeoxyglucose positron emission tomography in the staging and prognosis of T cell lymphoma. Leukemia & lymphoma 2013;54:2163-7.

17. Cerci JJ, Trindade E, Buccheri V, et al. Consistency of FDG-PET accuracy and cost-effectiveness in initial staging of patients with Hodgkin lymphoma across jurisdictions. Clinical lymphoma, myeloma & leukemia 2011;11:314-20.

18. Cheah CY, Dickinson M, Hofman MS, et al. Limited clinical benefit for surveillance PET-CT scanning in patients with histologically transformed lymphoma in complete metabolic remission following primary therapy. Annals of hematology 2014;93:1193-200.

19. Cheng G, Servaes S, Zhuang H. Value of (18)F-fluoro-2-deoxy-D-glucose positron emission tomography/computed tomography scan versus diagnostic contrast computed tomography in initial staging of pediatric patients with lymphoma. Leukemia & lymphoma 2013;54:737-42.

20. Crowley MP, O'Neill SB, Kevane B, et al. Ionizing radiation exposure as a result of diagnostic imaging in patients with lymphoma. Clin Transl Oncol 2016;18:533-6. doi: 10.1007/s12094-015-1394-8. Epub 2015 Aug 26.

21. Dann EJ, Bairey O, Bar-Shalom R, et al. Modification of initial therapy in early and advanced Hodgkin lymphoma, based on interim PET/CT is beneficial: a prospective multicentre trial of 355 patients. Br J Haematol 2017;178:709-18. doi: 10.1111/bjh.14734. Epub 2017 Jun 7.

22. Dann EJ, Berkahn L, Mashiach T, et al. Hodgkin lymphoma patients in first remission: routine positron emission tomography/computerized tomography imaging is not superior to clinical follow-up for patients with no residual mass. Br J Haematol 2014;164:694-700.

23. Deau B, Franchi P, Briere J, et al. PET2-driven de-escalation therapy in 64 high-risk Hodgkin Lymphoma patients treated with escalated BEACOPP. British journal of haematology 2015;170:278-81.

24. Duhrsen U, Muller S, Hertenstein B, et al. Positron Emission Tomography-Guided Therapy of Aggressive Non-Hodgkin Lymphomas (PETAL): A Multicenter, Randomized Phase III Trial. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2018;36:2024-34.

25. Engert A, Goergen H, Markova J, et al. Reduced-Intensity Chemotherapy in Patients With Advanced-Stage Hodgkin Lymphoma: Updated Results of the Open-Label, International, Randomised Phase 3 HD15 Trial by the German Hodgkin Study Group. HemaSphere 2017;1:e5.

26. Engert A, Haverkamp H, Kobe C, et al. Reduced-intensity chemotherapy and PET-guided radiotherapy in patients with advanced stage Hodgkin's lymphoma (HD15 trial): a randomised, open-label, phase 3 non-inferiority trial. Lancet (London, England) 2012;379:1791-9.

27. Epperla N, Shah N, Hamadani M, et al. Impact of Routine Surveillance Imaging on Outcomes of Patients With Diffuse Large B-Cell Lymphoma After Autologous Hematopoietic Cell Transplantation. Clin Lymphoma Myeloma Leuk 2016;16:672-8. doi: 10.1016/j.clml.2016.08.018. Epub Aug 29.

28. Figura N, Flampouri S, Mendenhall NP, et al. Importance of baseline PET/CT imaging on radiation field design and relapse rates in patients with Hodgkin lymphoma. Adv Radiat Oncol 2017;2:197-203. doi: 10.1016/j.adro.2017.01.006. eCollection Apr-Jun.

29. Friedberg JW, Byrtek M, Link BK, et al. Effectiveness of first-line management strategies for stage I follicular lymphoma: analysis of the National LymphoCare Study. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2012;30:3368-75.

30. Fujiwara H, Maeda Y, Nawa Y, et al. The utility of positron emission tomography/computed tomography in the staging of extranodal natural killer/T-cell lymphoma. European journal of haematology 2011;87:123-9.

31. Gallamini A, Tarella C, Viviani S, et al. Early Chemotherapy Intensification With Escalated BEACOPP in Patients With Advanced-Stage Hodgkin Lymphoma With a Positive Interim Positron Emission Tomography/Computed Tomography Scan After Two ABVD Cycles: Long-Term Results of the GITIL/FIL HD 0607 Trial. J Clin Oncol 2018;36:454-62. doi: 10.1200/JCO.2017.75.543. Epub 8 Jan 23.

32. Ganesan P, Rajendranath R, Kannan K, et al. Phase II study of interim PET-CT-guided response-adapted therapy in advanced Hodgkin's lymphoma. Ann Oncol 2015;26:1170-4. doi: 10.093/annonc/mdv077. Epub 2015 Feb 20.

33. Grellier JF, Vercellino L, Leblanc T, et al. Performance of FDG PET/CT at initial diagnosis in a rare lymphoma: nodular lymphocyte-predominant Hodgkin lymphoma. European journal of nuclear medicine and molecular imaging 2014;41:2023-30.

34. Guttikonda R, Herts BR, Dong F, Baker ME, Fenner KB, Pohlman B. Estimated radiation exposure and cancer risk from CT and PET/CT scans in patients with lymphoma. European journal of radiology 2014;83:1011-5.

35. Hong J, Kim JH, Lee KH, et al. Symptom-oriented clinical detection versus routine imaging as a monitoring policy of relapse in patients with diffuse large B-cell lymphoma. Leukemia & lymphoma 2014;55:2312-8.

36. Huntington SF, Svoboda J, Doshi JA. Cost-effectiveness analysis of routine surveillance imaging of patients with diffuse large B-cell lymphoma in first remission. J Clin Oncol 2015;33:1467-74. doi: 10.200/JCO.2014.58.5729. Epub 2015 Mar 30.

37. Johnson PW. Response-adapted frontline therapy for Hodgkin lymphoma: are we there yet? Hematology Am Soc Hematol Educ Program 2016;2016:316-22. doi: 10.1182/asheducation-2016.1.316.

38. Kedmi M, Apel A, Davidson T, et al. High-Risk, Advanced-Stage Hodgkin Lymphoma: The Impact of Combined Escalated BEACOPP and ABVD Treatment in Patients Who Rapidly Achieve Metabolic Complete Remission on Interim FDG-PET/CT Scan. Acta haematologica 2016;135:156-61.

39. Lamy T, Damaj G, Soubeyran P, et al. R-CHOP 14 with or without radiotherapy in nonbulky limited-stage diffuse large B-cell lymphoma. Blood 2018;131:174-81.

40. Metser U, Dudebout J, Baetz T, et al. [(18) F]-FDG PET/CT in the staging and management of indolent lymphoma: A prospective multicenter PET registry study. Cancer 2017;123:2860-6. doi: 10.1002/cncr.30672. Epub 2017 Mar 13.

41. Moghbel MC, Mittra E, Gallamini A, et al. Response Assessment Criteria and Their Applications in Lymphoma: Part 2. Journal of nuclear medicine : official publication, Society of Nuclear Medicine 2017;58:13-22.

42. Moskowitz AJ, Schoder H, Gavane S, et al. Prognostic significance of baseline metabolic tumor volume in relapsed and refractory Hodgkin lymphoma. Blood 2017;130:2196-203. doi: 10.1182/blood-2017-06-788877. Epub 2017 Sep 5.

43. Moskowitz AJ, Schoder H, Yahalom J, et al. PET-adapted sequential salvage therapy with brentuximab vedotin followed by augmented ifosamide, carboplatin, and etoposide for patients with relapsed and refractory Hodgkin's lymphoma: a non-randomised, open-label, single-centre, phase 2 study. Lancet Oncol 2015;16:284-92. doi: 10.1016/S470-2045(15)70013-6. Epub 2015 Feb 13.

44. Moskowitz CH, Matasar MJ, Zelenetz AD, et al. Normalization of pre-ASCT, FDG-PET imaging with second-line, non-cross-resistant, chemotherapy programs improves event-free survival in patients with Hodgkin lymphoma. Blood 2012;119:1665-70.

45. Cancer Stat Facts: Hodgkin Lymphoma. (Accessed Aug., 2018, at https://seer.cancer.gov/statfacts/html/hodg.html.)

46. Papajik T, Myslivecek M, Skopalova M, et al. Determining the extent and stage of disease in patients with newly diagnosed non-Hodgkin's lymphoma using 18F-FDG-PET/CT. Neoplasma 2011;58:291-7.

47. Pardal E, Coronado M, Martin A, et al. Intensification treatment based on early FDG-PET in patients with high-risk diffuse large B-cell lymphoma: a phase II GELTAMO trial. Br J Haematol 2014;167:327-36.

48. Patel V, Buckstein M, Perini R, Hill-Kayser C, Svoboda J, Plastaras JP. Computed tomography and positron emission tomography/computed tomography surveillance after combined modality treatment of supradiaphragmatic Hodgkin lymphoma: a clinical and economic perspective. Leukemia & lymphoma 2013;54:2168-76.

49. Paulino AC, Margolin J, Dreyer Z, Teh BS, Chiang S. Impact of PET-CT on involved field radiotherapy design for pediatric Hodgkin lymphoma. Pediatric blood & cancer 2012;58:860-4.

50. Picardi M, Pugliese N, Cirillo M, et al. Advanced-stage Hodgkin lymphoma: US/chest radiography for detection of relapse in patients in first complete remission--a randomized trial of routine surveillance imaging procedures. Radiology 2014;272:262-74.

51. Pingali SR, Jewell SW, Havlat L, et al. Limited utility of routine surveillance imaging for classical Hodgkin lymphoma patients in first complete remission. Cancer 2014;120:2122-9.

52. Pommier P, Dussart S, Girinsky T, et al. Impact of 18F-fluoro-2-deoxyglucose positron emission tomography on treatment strategy and radiotherapy planning for stage I-II Hodgkin disease: a prospective multicenter study. International journal of radiation oncology, biology, physics 2011;79:823-8.

53. Press OW, Li H, Schoder H, et al. US Intergroup Trial of Response-Adapted Therapy for Stage III to IV Hodgkin Lymphoma Using Early Interim Fluorodeoxyglucose-Positron Emission Tomography Imaging: Southwest Oncology Group S0816. J Clin Oncol 2016;34:2020-7. doi: 10.1200/JCO.2015.63.1119. Epub 2016 Apr 11.

54. Radford J, Illidge T, Counsell N, et al. Results of a trial of PET-directed therapy for early-stage Hodgkin's lymphoma. N Engl J Med 2015;372:1598-607. doi: 10.056/NEJMoa1408648.

55. Raemaekers JM, André MP, Federico M, et al. Omitting radiotherapy in early positron emission tomography-negative stage I/II Hodgkin lymphoma is associated with an increased risk of early

relapse: Clinical results of the preplanned interim analysis of the randomized EORTC/LYSA/FIL H10 trial. Journal of clinical oncology 2014;32:1188-94.

56. Rashidi A, Ebadi M, Cashen AF. Allogeneic hematopoietic stem cell transplantation in Hodgkin lymphoma: a systematic review and meta-analysis. Bone marrow transplantation 2016;51:521.
57. Rathore N, Eissa HM, Margolin JF, et al. Pediatric Hodgkin lymphoma: are we over-scanning our patients? Pediatric hematology and oncology 2012;29:415-23.
58. Robertson VL, Anderson CS, Keller FG, et al. Role of FDG-PET in the definition of involved-field radiation therapy and management for pediatric Hodgkin's lymphoma. International journal of radiation oncology, biology, physics 2011;80:324-32.
59. Romano A, Parrinello NL, Vetro C, et al. Prognostic meaning of neutrophil to lymphocyte ratio (NLR) and lymphocyte to monocyte ration (LMR) in newly diagnosed Hodgkin lymphoma patients treated upfront with a PET-2 based strategy. Ann Hematol 2018;97:1009-18. doi: 10.7/s00277-018-3276-y. Epub 2018 Feb 14.
60. Rovira J, Valera A, Colomo L, et al. Prognosis of patients with diffuse large B cell lymphoma not reaching complete response or relapsing after frontline chemotherapy or immunochemotherapy. Annals of hematology 2015;94:803-12.
61. Sohn BS, Yoon DH, Kim KP, et al. The role of (1)(8)F-fluorodeoxyglucose positron emission tomography at response assessment after autologous stem cell transplantation in T-cell non-Hodgkin's lymphoma patients. Annals of hematology 2013;92:1369-77.
62. Stewart DA, Kloiber R, Owen C, et al. Results of a prospective phase II trial evaluating interim positron emission tomography-guided high dose therapy for poor prognosis diffuse large B-cell lymphoma. Leukemia & lymphoma 2014;55:2064-70.
63. Straus DJ, Jung S-H, Pitcher B, et al. CALBG 50604: risk-adapted treatment of non-bulky early stage Hodgkin lymphoma based on interim PET. Blood 2018:blood-2018-01-827246.
64. Swinnen LJ, Li H, Quon A, et al. Response-adapted therapy for aggressive non-Hodgkin's lymphomas based on early [18F] FDG-PET scanning: ECOG-ACRIN Cancer Research Group study (E3404). Br J Haematol 2015;170:56-65. doi: 10.1111/bjh.13389. Epub 2015 Mar 30.
65. Taghipour M, Marcus C, Jones S, Sarangi R, Trahan TJ, Subramaniam RM. The value of fourth and subsequent post-treatment 18F-FDG PET/CT scans in the management of patients with non-Hodgkin's lymphoma. Nuclear medicine communications 2016b;37:699-704.
66. Taghipour M, Marcus C, Nunna P, Subramaniam RM. Follow-up FDG PET/CT in Patients With Non-Hodgkin Lymphoma: Value to Clinical Assessment and Patient Management. Clin Nucl Med 2016a;41:e93-7.
67. Terezakis SA, Schoder H, Kowalski A, et al. The utility and limitations of (18)F-fluorodeoxyglucose positron emission tomography with computed tomography in patients with primary mediastinal B-cell lymphoma: single institution experience and literature review. Int J Radiat Oncol Biol Phys 2014;89:376-83. doi: 10.1016/j.ijrobp.2014.02.006. Epub Apr 11.
68. Tilly H, Gomes da Silva M, Vitolo U, et al. Diffuse large B-cell lymphoma (DLBCL): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of oncology : official journal of the European Society for Medical Oncology 2015;26 Suppl 5:v116-25.

69. van Hamersvelt HP, Kwee TC, Fijnheer R, Beek FJ, de Klerk JM, Nievelstein RA. Can full-dose contrast-enhanced CT be omitted from an FDG-PET/CT staging examination in newly diagnosed FDG-avid lymphoma? Journal of computer assisted tomography 2014;38:620-5.

70. Zinzani PL, Broccoli A, Casadei B, et al. The role of rituximab and positron emission tomography in the treatment of primary mediastinal large B-cell lymphoma: experience on 74 patients. Hematological oncology 2015;33:145-50.

71. Zinzani PL, Broccoli A, Gioia DM, et al. Interim Positron Emission Tomography Response-Adapted Therapy in Advanced-Stage Hodgkin Lymphoma: Final Results of the Phase II Part of the HD0801 Study. J Clin Oncol 2016;34:1376-85. doi: 10.200/JCO.2015.63.0699. Epub 2016 Feb 16.

1. Appraisal

1.1. Background and Rationale

Positron emission tomography (PET) is a type of nuclear medicine imaging that utilizes small amounts of radioactive materials called radiotracers to examine and measure physiological functions in the body. The more energy a group of cells needs, the more the radiotracer will build up in that location. For lymphoma, the radioactive particle most commonly used for PET is ^{18}fluorine, which binds to glucose to form fluorodeoxy-D-glucose (FDG). Most types of lymphoma are metabolically active and use more glucose compared with normal structures (i.e. are termed FDG avid). This results in greater uptake of the radioactive FDG creating a "hot" spot on the PET image. In general, ^{18}FDG-PET is not routinely used for initial diagnosis of lymphoma as histologic samples obtained via biopsy are required; however, ^{18}FDG-PET may assist in identification of the best place for biopsy (e.g. most metabolically active site, locations where biopsy may be difficult, previous non-diagnostic needle biopsy). ^{18}FDG-PET has become widely used as an imaging tool for staging of lymphoma after histological diagnosis and for interim evaluation (e.g. restaging and evaluation of treatment response) although there is some debate regarding the value of interim PET based on the quality of evidence available and inconsistency regarding criteria for PET interpretation in the literature. Interim PET is not generally performed for indolent lymphoma. There are data suggesting that ^{18}FDG-PET may be a predictor of prognosis when performed early during treatment. PET is generally not used for routine surveillance following treatment completion, largely due to concerns regarding false-positive findings and lack of impact on patient outcomes.

Fryback and Thornbury's hierarchical model for the evaluation of efficacy of medical tests describes six levels of efficacy: 1) technical efficacy, 2) diagnostic accuracy efficacy, 3) diagnostic thinking efficacy, 4) therapeutic efficacy, 5) patient outcome efficacy and 6) societal efficacy. Efficacy at lower levels is needed as a basis for efficacy at the higher levels, however increases in efficacy at the lower levels (e.g. diagnostic accuracy) may not have a corresponding improvement at the higher levels (e.g. patient outcomes).[72] The AHRQ's adaptation of this model for systematic reviews was used to conceptualize the framework for this report.[105] The 2011 report provided information on diagnostic accuracy, i.e. the ability to correctly classify patients into a disease/pheontotype or prognosis category (e.g. test performance measures such as sensitivity, specificity and predictive values) which has been augmented and updated via the contextual question that is part of this report. Subsequent to the 2011 report, additional studies of the impact of PET/CT on clinical decision making may further address the clinical utility of PET/CT in lymphoma. Thus, the focus of this re-review is on evidence for levels 3-6, to the extent that it is available, with a primary focus on the ability of PET/CT to direct clinical management and improve patient outcomes in a cost-effective manner.

The value of a medical test relates to its ability identify persons for whom there are appropriate and effective treatments. To evaluate the value of PET/CT, studies comparing treatment given based on PET/CT results with treatment that would have been given without PET/CT information are of primary interest. In the absence of studies randomizing patients to receive PET/CT or no PET/CT and following patients through treatments to final clinical outcomes, the impact of PET on clinical decision making leading to effective treatment and outcomes can only be assessed indirectly. The outcomes observed are a function of PET/CT results in combination with the treatments received on the basis of those results. Thus, our approach considers the clinical outcomes of OS and PFS in relation to the modifications of treatment received based on PET results, along with treatment related-toxicity. In other

words, for this report the value of PET for assisting with decisions to escalate, de-escalate or continue therapy is indirectly assessed based upon the impact of the PET-adapted treatment.

Policy Context

This topic was originally reviewed in 2011. It is being re-reviewed in 2018 due to newly available published evidence.

Objectives

The aim of this report is to update the 2011 HTA on positron emission tomography (PET) for lymphoma by summarizing information on diagnostic accuracy (e.g. sensitivity, specificity, predictive values) for context and systematically reviewing, critically appraising and analyzing new research evidence evaluating the clinical effectiveness (i.e. the ability of PET to stage, and influence therapeutic decisions, clinical management and clinical outcomes), safety, differential efficacy and safety in subpopulations, and cost-effectiveness of PET for lymphoma in adult and pediatric patients. The combination of PET with diffusion weighted MRI is an emerging technology for the evaluation of lymphoma. Currently this combination is not widely used, so the focus of this report will be on PET/CT as it is the current standard of care. Evidence on PET/MRI will be included as appropriate.

1.2. Contextual Questions

In patients with histologically proven HL and NHL included in the report, what are the accuracy and reliability of ^{18}FDG-PET alone or in combination with CT for initial staging, interim evaluation (including monitoring during treatment, evaluation of treatment response and prognosis), re-staging (e.g. at end of treatment or other times), end of treatment assessment, evaluation of disease recurrence and surveillance of patients in remission? Specifically, provide a summary of the:

- Sensitivity and specificity and prognostic value (positive and negative predictive values)
- Inter- and intra-rater reliability (reproducibility)

In addition, a brief summary of the diagnostic accuracy and use of PET/CT for initial diagnosis will be provided. Summaries of accuracy will be based on highest quality systematic reviews which critically appraise included studies. Contextual information on the combination of PET with MRI will be presented. Contextual questions are not systematically reviewed, and use a “best evidence” approach.

1.3. Key Questions

The focus of this portion of the report is on the clinical impact of ^{18}FDG PET/CT as this is the current standard of care. Except where noted the term PET will be construed to mean PET/CT. Information related to the PET/MRI combination will be included if relevant. In patients with histologically proven HL or NHL undergoing PET/CT for initial staging, interim evaluation (including monitoring during treatment, evaluation of treatment response and prognosis), re-staging (e.g. at end of treatment or other times), end of treatment assessment, evaluation of disease recurrence and surveillance of patients in remission:

1. What is the evidence of clinical effectiveness of ^{18}FDG imaging in combination with CT (PET/CT) results?

   a. How do ^{18}FDG PET/CT results impact therapeutic decisions or clinical management? Do test results lead to use of effective treatment strategies (e.g. including initial treatment following staging or treatment acceleration, deceleration or termination at interim imaging) compared with treatment strategies not using such test results?
   b. How do clinical outcomes (e.g. overall survival) differ based on PET/CT-related treatment decisions compared with decisions made in the absence of such test results?
   c. Does the use of ^{18}FDG PET for treatment decisions lead to reduction in treatment-related adverse events/sequelae in general compared with treatment decisions that do not involve PET/CT?
   d. Is there a reduction in the need for other tests?
   e. How do end of treatment ^{18}FDG PET/CT results impact clinical decision making compared with clinical decisions that do not involve PET/CT??
   f. How does surveillance ^{18}FDG PET/CT results impact clinical decision making compared with clinical decisions that do not involve PET/CT??

2. What is the safety profile of ^{18}FDG PET/CT for lymphoma?
   a. What adverse events are reported: type and frequency directly attributable to ^{18}PET/CT (mortality, major morbidity, radiation exposure, other)?

3. What is the evidence that ^{18}FDG PET/CT imaging in patients with known lymphoma has differential efficacy or safety in subpopulations? Including consideration of:
   a. Patient age, sex, characteristics or evidence-based patient selection criteria
   b. Type of scanning machine and software, reader training, and other operational factors
   c. Provider type, setting or other provider characteristics
   d. Health care system type, including worker's compensation, Medicaid, state employees

4. What is the evidence of short and long-term cost-effectiveness of ^{18}FDG PET/CT for patients with lymphoma compared with other imaging or clinical care not involving ^{18}FDG PET/CT?

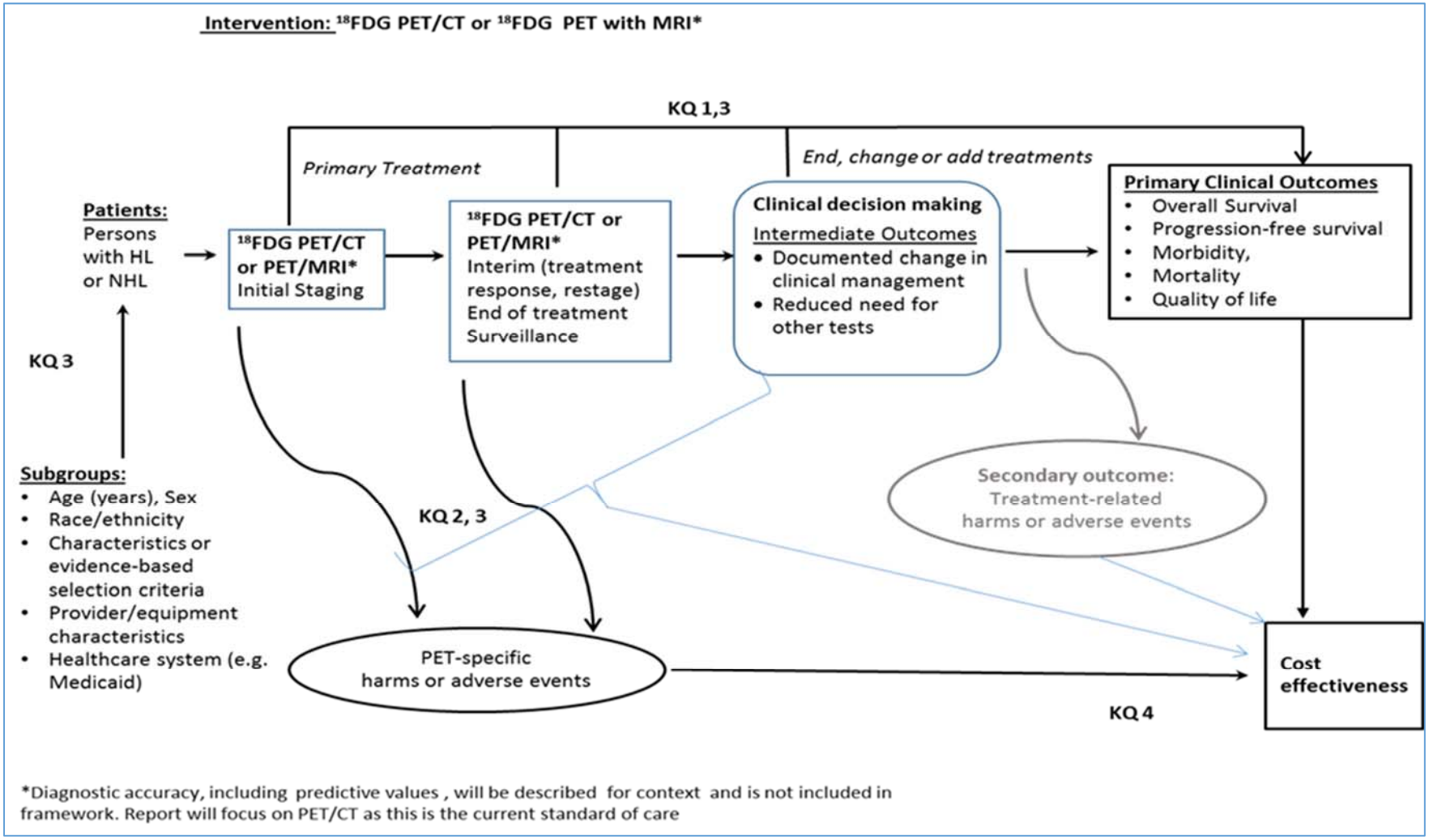
Inclusion and exclusion criteria are summarized as follows and detailed in table 8. Briefly, included studies met the following requirements with respect to participants, intervention, comparators, outcomes, and study design:

- **Population**: Adults, adolescents or children with biopsy-proven HL or NHL. Diagnoses of interest (to be evaluated separately) include HL or aggressive NHL, indolent NHL and other NHL with focus on most common types.
- **Interventions:** Focus is on positron emission tomography (PET) to measure glucose metabolism (^{18}FDG-PET) in addition to computed tomography (CT) (including combined PET/CT equipment). Combination of PET with MRI (including diffusion weighted MRI) will be included as appropriate.
- **Comparators:** Other imaging (CT alone, MRI, including diffusion weighted MRI); standard clinical protocols or standard prior tests/evaluations (including history and physical examination, laboratory studies, biopsy) that do not involve ^{18}FDG PET.
- **Outcomes:** Primary outcomes are 1) improvement in clinical outcomes based on PET/CT-derived clinical decision making (focus is on overall survival, progression- or event-free survival, morbidity and mortality, and 2) quality of life. Safety outcomes are type and frequency of adverse events directly attributable to PET/CT (e.g., incidental findings, repeat/additional

procedures, radiation exposure, other). Economic outcomes are cost-effectiveness (e.g., cost per improved outcome), cost-utility (e.g., cost per quality adjusted life year (QALY), incremental cost effectiveness ratio (ICER) outcomes.

- **Timing:** PET for initial staging, interim evaluation, end-of-treatment evaluation, and surveillance.
- **Studies:** For efficacy and effectiveness, the focus was on comparative PET/CT studies with the least potential for bias (i.e. RCTs). For safety, the focus will be on studies characterizing direct PET/CT harms (including incidental findings, repeat biopsy, radiation safety). For differential efficacy or effectiveness, only studies which stratify on patient or other characteristics and formally evaluate statistical interaction (effect modification) were considered. For cost-effectiveness, only full, formal economic studies (i.e., cost-effectiveness, cost-utility, cost-minimization, and cost-benefit studies) were included. Given that this is re-review, only studies published subsequent to the 2011 report (Search dates February 2011 to May 2018) were considered.

Figure 1. Analytic framework



1.4. Outcomes Assessed

The primary outcomes of interest for this report are listed below.

- Overall survival (OS)
- Progression-free survival (PFS) or Event-free survival (EFS)
  - Disease-free survival (DFS)
  - Relapse-free survival (RFS)
- Treatment-related toxicity (as reported specific to PET-adapted therapy)

OS and PFS/EFS were stated *a priori* as primary outcomes of interest. Some of the included studies also reported DFS and RFS. Excluding OS, definitions of these outcomes varied slightly between the studies. Details can be found in the data abstraction in Appendices F-I.

Other outcomes reported included proportion of patients upstaged or down staged, changes in clinical management as a results of change in stage, and detection of relapse.

Strength of evidence was assessed for the primary clinical outcomes only.

1.5. Washington State Utilization and Cost Data

Populations

The **Positron Emission Tomography (PET) Scans for Lymphoma** analysis includes member utilization and cost data from the following agencies: PEBB/UMP (Public Employees Benefit Board Uniform Medical Plan); PEBB Medicare, the Department of Labor and Industries (L&I) workers' compensation plan; and the Managed Care Medicaid and the Medicaid Fee-for-Service. All agency utilization and costs information for PET Lymphoma are presented in aggregate.[1]

The analysis period was three (3) calendar years, 2015 - 2017. Primary population inclusion criteria included incurring a paid claim(s) comprised of: at least one of the targeted CPT/HCPCS codes from Table I; a modifier code PI or PS (see Table II); and a targeted diagnosis from Table III. Denied claims were excluded from the analysis.

Methods

First, patients with a paid claim for a targeted CPT procedure (Table I) were identified; that patient's claim was matched with a targeted CPT modifier PI or PS (see Table II); finally, the patient's claim had to include a primary targeted diagnoses from Table III. Data evaluation included examining utilization by member; analysis of individual and aggregate CPT codes and by total claims' costs incurred by a member on the date of their service.

Table I
Targeted CPT Descriptions

TYPE	CPT	Procedure Code Description
PET/CT	78814	Positron emission tomography (PET) with concurrently acquired computed tomography (CT) for attenuation correction and anatomical localization imaging; limited area (e.g., chest, head/neck)
PET/CT	78815	Positron emission tomography (PET) with concurrently acquired computed tomography (CT) for attenuation correction and anatomical localization imaging; skull base to mid-thigh
PET/CT	78816	Positron emission tomography (PET) with concurrently acquired computed tomography (CT) for attenuation correction and anatomical localization imaging; whole body
PET	A9597	Positron emission tomography radiopharmaceutical, diagnostic, for tumor identification, not otherwise classified

[1] Specific agencies reported less than the minimum number of patents (greater than 30) necessary to safely release agency-by-agency findings and still protect patient confidentiality.

Table II

Targeted Modifier Code Descriptions

The analysis excluded claims with a Q0 modifier code. Q0 modifiers identify services classified as investigational clinical service provided in a clinical research study that is in an approved clinical research study

Modifier	Modifier Code Description
PI	PET or Pet/CT to inform the initial treatment strategy of tumors that are biopsy proven or strongly suspected of being cancerous based on other diagnostic testing.
PS	PET or Pet/CT to inform the subsequent treatment strategy of cancerous tumors when the beneficiary's treating physician determines that the PET study is needed to inform subsequent antitumor strategy.

Table III

Target Diagnosis Codes

ICD-9/10	Diagnosis Code Description
C81 - C88.9	
C81	Hodgkin lymphoma
C82	Follicular lymphoma
C83	Non-follicular lymphoma
C84	Mature T/NK-cell lymphomas
C85	Other specified and unspecified types of non-Hodgkin lymphoma
C86	Other specified types of T/NK-cell lymphoma
C88	Malignant immunoproliferative diseases and certain other B-cell lymphomas
200.0-202.98	
200	Lymphosarcoma and reticulosarcoma and other specified malignant tumors of lymphatic tissue
201	Hodgkin's disease
202	Other malignant neoplasm of lymphoid and histiocytic tissue

Table IV
Aggregate Utilization 2015 – 2017

PEBB/UMP (Public Employees Benefit Board Uniform Medical Plan); PEBB Medicare, the Department of Labor and Industries (L&I) workers' compensation plan; and the Managed Care Medicaid and the Medicaid Fee-for-Service

Year	Unique Patients	SCANS*	Total Paid Amt -Technical and Professional
2015	46	49	$39,237
2016	227	273	$232,448
2017	224	255	$205,509

* Range of scans/patient 1–6

Table V
Definitions for Utilization Tables

Unique members	Unique, non-duplicated member, reported by agency
Count of Scans (CPT)	Count of unduplicated scans with a CPT code and targeted modifier on a single date.
Total dollars on date of service	Total paid dollars for all services received on the date of the scan procedure – Technical and Professional

2. Background

2.1. Epidemiology and Burden of Disease

Lymphoma is a heterogeneous group of blood cancers that affect lymphocytes of the immune system within the lymphatics, a complex network of vessels, tissues and organs which carry a fluid called lymph throughout the body. Lymphocytes are a type of white blood cell that account for 20% to 40% of the total number of white blood cells in adults, and play an integral role in the human immune response by recognizing and destroying infectious organisms and abnormal cells. Lymphoma most often starts at a single site within the lymphatic system, such as a lymph node, and then progresses to adjacent lymph nodes via lymphatic channels before disseminating to distant nonadjacent sites and organs. Between 10% and 35% of patients will have primary lymphoma at extranodal sites (e.g. beyond the lymph nodes) at initial diagnosis, with the most common sites being the GI tract and skin.

There are two main types of lymphoma, Hodgkin (HL) and non-Hodgkin (NHL) (see sections below for more details), with the latter being the most common. In 2015, there were an estimated 686,042 people living with NHL and 208,805 people living with HL in the United States.[2,3] For both men and women in the United States, the lifetime risk of getting NHL is 2.1 percent[3]; the risk of getting HL is around 0.2 percent.[2]

Lymphoma can occur at any age, but it is one of the most common cause of cancer in children and young adults. Lymphomas are the 4th most common cancer in children aged 1 to 14 years, comprising 10% of childhood cancers; in adolescents (ages 15 to 19 years), lymphomas account for 22% of cancers in this age group, with approximately twice as many cases of HL as NHL.[157]

2.1.1. Hodgkin (HL) and non-Hodgkin lymphoma (NHL)

Lymphomas are divided into two major categories: HL (previously called Hodgkin's disease) and NHL (all other lymphomas). While HL and NHL occur in the same locations, may be associated with the same symptoms (e.g. swollen lymph nodes), and often have similar appearance on physical examination, they can be differentiated histologically by the specific type of lymphocyte each involves. Further, immunohistochemistry and flow cytometry are important for the diagnosis and subclassification of most forms of lymphoma. Due to confusion with classifying lymphomas and other blood cancers (e.g. leukemias), the World Health Organization in 2001 devised a classification of lymphoid malignancies through a process of consensus development, taking into account morphologic, clinical, immunologic and genetic information. The system was most recently revised in 2016[142] and generally classifies lymphoid malignancies into five categories: mature B cell origin neoplasms, mature T and natural killer (NK) cell neoplasms, Hodgkin lymphoma, posttransplant lymphoproliferative disorders, and histiocytic and dendritic cell neoplasms.

Hodgkin Lymphoma (HL)

HL arises from a malignant transformation of B-lymphocytes and accounts for approximately 10% of all lymphomas in the US. There are two main subtypes of HL: classic HL (further divided into four distinct subtypes) which accounts for 95% of all HLs in adults, and nodular lymphocyte-predominant HL which accounts for the remaining 5%. Classic HL is characterized by the presence of Reed-Sternberg cells (a

specific abnormal B lymphocyte); lymphocyte-predominant HL is characterized by the presence of lymphocyte-predominate cells, also called "popcorn cells," and rare Reed-Sternberg cells. HL is rare, accounting for 0.6% of all new cancers diagnosed (approximately 8500 cases were diagnosed in 2016) and is more prevalent in younger patients (32% are age 20 to 34 years, with a median age of 39 years). There is a bimodal distribution of age at diagnosis,[137] with one peak incidence occurring in younger individuals (approximately 20 years) and the other in older adults (≥ age 65 years). HL accounts for about 7% of childhood cancers and 1% of childhood cancer deaths in the US. It is very rare in younger children (age 1 to 14 years) accounting for 4% of cancers in this age group, but is the most common cancer in the adolescent (age 15 to 19) age group,[157] comprising 15% of cancer in adolescents.

HL is a highly curable malignancy with 5-year survival rates as high 92% among early-stage adult patients and 98% among children.[2]

Non-Hodgkin Lymphoma (NHL)

NHL arises from a malignant transformation of either B-lymphocytes (primarily, about 85%),T-lymphocytes (15%), and natural killer lymphocytes (<1%). It is estimated that over 74,000 individuals are diagnosed with NHL yearly in the US, making it one of the most commonly diagnosed cancers. Patients diagnosed with NHL tend to be older with a median age of 66 years (75% are age 55 or older). There are upwards of 60 different subtypes of NHL and they are separated into two distinct categories: indolent NHL (iNHL) and aggressive NHL (aNHL).

iNHLs progress slowly and make up about 40% of all NHL in the U.S. Forms include cutaneous T-cell lymphoma, follicular lymphoma, lymphoplasmacytic lymphoma, Waldenström macroglobulinemia, small cell lymphocytic lymphoma, chronic lymphocytic leukemia and marginal zone lymphoma and mucosa-associated lymphoid tissue (MALT) lymphoma. Conversely, aNHLs are highly aggressive, present with a rapidly growing mass, and may result in mortality within weeks if left untreated. The most common type of aNHL is diffuse large B-cell lymphoma (DLBCL) which accounts for 30% to 40% of all cases. Examples of other subtypes of aNHL include Burkitt lymphoma, adult T-cell leukemia/lymphoma, and precursor B and T lymphoblastic leukemia/lymphoma. Burkitt's lymphoma is a rare disease in adults in the U.S. (<1% of adult NHLs) but comprises approximately 19% of childhood NHL. Other common childhood NHLs are DLBCL (~22% of childhood NHLs), B- or T-cell lymphoblastic lymphoma (~20% of childhood NHLs), and anaplastic large-cell lymphoma (~10% of childhood NHLs). Another issue for iNHL is late histological transformation from indolent to aggressive NHL. Histological transformation may be detected by change in symptoms, enlargement of tumor mass, change in chemical markers or changes in imaging studies. In the US, NHL accounts for 6% of cancer in children (age 1 to 14 years) and 8% of cancer in adolescents (age 15 to 19).[157]

For adults with NHL, 5-year survival rates are around 70%; for children and adolescents are they are higher at an estimated 85% or greater.

2.1.2. Diagnosis and Clinical Management of Lymphoma

Each type of lymphoma behaves, spreads, and responds to treatment differently, so an accurate diagnosis is essential to determining the appropriate treatment strategy, expected response to treatment, and monitoring for recurrence. Lymphoma diagnosis and initial staging is based on physical symptoms (i.e., systemic complaints of fever, weight loss, or night sweats), physical exam, lymph node

excisional biopsy (of enlarged lymph node), and imaging exams, blood tests, and bone marrow examination. The initial evaluation must establish the precise histologic subtype to determine the appropriate therapy for the patient. In general, FDG-PET is not routinely used for initial diagnosis of lymphoma as histologic samples obtained via biopsy are required; however, FDG-PET may assist in identification of the best place for biopsy (e.g. most metabolically active site, locations where biopsy may be difficult, of areas of previous non-diagnostic needle biopsy).

After diagnosis, patients undergo initial staging, which quantifies the severity of an individual's cancer based on the size of the primary tumor as well as on the extent of metastasis. Imaging studies are a key component of the staging evaluation to help determine the severity and sites of disease. The preferred modality for staging a patient depends upon the fluorodeoxyglucose (FDG) avidity of the histologic subtype. Integrated positron emission tomography with computed tomography (PET/CT) is preferred for staging FDG-avid nodal lymphomas and is generally considered the current standard of practice, while CT alone is preferred for FDG-nonavid and variably FDG-avid histologies.[51]

After initial staging, typical management of lymphoma entails induction or initial treatment, restaging after one or more cycles of initial treatment, end of treatment assessment, and either surveillance if complete response has been achieved at the end of treatment or additional treatment followed by restaging. Treatment options can include chemotherapy (either in combinations or as a mono-therapy), radiation therapy (either as an adjunctive treatment or a mono-therapy), targeted therapy (i.e., monoclonal antibodies, immune checkpoint inhibitors), stem cell transplantation and in rare cases, surgery. Treatments are based on lymphoma type, disease stage, age, co-existent medical conditions and prognostic factors (Table 1). All treatments have been associated with a broad range side effects depending on the therapies used, and may include serious effects such as fertility issues, damage to the thyroid, heart and lungs as well as increased risk for infection, stroke and secondary cancers. Appendix L provides a summary of guideline-recommended treatments and related abbreviations.

Chemotherapy Regimens

An expansive range of chemotherapy agents and numerous agent combinations are available for use in the treatment of lymphoma, generally with or without radiation therapy and/or subsequent high-dose therapy followed by stem cell transplantation. For common lymphoma types, standard chemotherapy regimens have emerged that balance high survival rates with lower treatment-related toxicities, especially in early treatment. For rarer subtypes, however, evidence definitively supporting specific regimens over others is often limited. Survival rates, too, vary distinctly across lymphoma types and subtypes, depending on key factors such as age, stage at diagnosis, prognostic indicators (such as B-symptoms or bulky disease) and the relative indolence or aggressiveness of the lymphoma type. Table 1 provides general information on survival rates from clinical guidelines and the SEER data base (see Appendix L for full summary of guidelines).[34].

Table 1. Summary of survival by lymphoma type described in guidelines and the SEER database.

Lymphoma	Survival Rates by Type/Subtype*†
Hodgkin Lymphoma	5-year: 86.6% (by stage): [64] - I: 92.3% -II: 93.4% -III or higher: 78.2%
Classical HL[64,92]	1 Year: 92% 10 Year: 78.1%
Nodular Lymphocyte Predominant HL	1 year- 98.1% 10 year - 85.7%
Non-Hodgkin Lymphoma	5-year: 71.4% (by stage): [64] - I: 83.3% -II: 75.2% -III or higher: 64.1%
DLBCL[151,163]	3-year OS: 59% (high-risk) to 91% (low-risk) 5-year OS: 95%
Follicular Lymphoma[62,163]	5-year OS: 52.5% to 91% 2-year OS: 90% 3-year OS: 89-95% 4-year OS: 83-87% 8-year OS: 79%
Other NHL subtypes	No summary rate due to subtype heterogeneity; survival rates vary greatly depending on subtype aggressiveness and staging

HL = Hodgkin Lymphoma; NHL = non-Hodgkin Lymphoma; OS = overall survival;
* Survival rates are Relative Survival unless otherwise noted. SEER defines Relative Survival as an "estimate of the percentage of patients who would be expected to survive the effects of their cancer" over the given time period.
† For a full list of definitions of lymphoma abbreviations and chemotherapy regimen acronyms and abbreviations see Appendix M.

2.2. Technologies/Interventions

2.2.1. Positron Emission Tomography (PET) and Timing of Evaluation

Positron emission tomography (PET) is a type of nuclear medicine imaging that utilizes small amounts of radioactive materials called radiotracers to examine and measure physiological functions in the body. The more energy a group of cells needs, the more the radiotracer will build up in that location. For lymphoma, the only FDA-approved radioactive particle used for PET is ^{18}fluorine, which binds to glucose to form fluorodeoxy-D-glucose (FDG). Most types of lymphoma are metabolically active and use more glucose compared with normal structures (i.e. are termed FDG avid). This results in greater uptake of the radioactive FDG creating a "hot spot" on the PET image.

Today, most PET scans are performed on combination PET-CT scanners. The combined PET-CT scans provide images that pinpoint the anatomic location of abnormal metabolic activity within the body and provide more accurate staging and evaluation than the two scans performed separately. Most clinical guidelines recommend the use of PET in conjunction with CT, whether each test is done separately or via an integrated PET-CT scanner for initial staging and re-staging at critical points after treatment. The combination of PET with diffusion weighted MRI is an emerging technology for the evaluation of lymphoma and is the modality of choice for suspected central nervous system (CNS) involvement.[34]

While PET/CT is generally considered the standard for staging, interim evaluation (e.g. restaging and evaluation of treatment response), and end of treatment assessment it is not without some potential drawbacks. The claim for PET compared to other imaging methods such as CT and MRI is that uptake of ^{18}FDG by cancer cells is both more sensitive and specific for cancer than alterations in local anatomy and tissue properties that might be detected by CT and MRI. However, false negative PET scans can result from areas of cancer that may be too small or too metabolically inactive to accumulate enough ^{18}FDG to be detected by the PET scan. A false negative result can cause a delay in diagnosis or treatment. Alternatively, false positive PET scans can result from other causes of increased glucose metabolism such as hyperemia, infection, inflammation or tissue healing that may lead to abnormal accumulation of ^{18}FDG and then appear as "hot spots" on PET scans. False-positive PET results at interim and end-of-treatment assessment are frequently encountered and may lead to additional, likely unnecessary, testing (including biopsy) and create worry in patients.[21,23]

Initial Staging

FDG-PET/CT has been formally incorporated into the standard staging process for lymphoma. For HL and FDG-avid NHL subtypes, PET improve the accuracy of staging compared with CT scans alone due to greater sensitivity for nodal and extranodal involvement[50,162] as well as its ability to indicate the overall level of metabolic activity of lymphoma, which correlates with level of aggressiveness and with serum lactate dehydrogenase (LDH) level (a prognostic factor).[55] Consequently, PET-CT is preferred for staging of FDG-avid, nodal lymphomas, and CT scan is preferred for staging of other lymphomas, including variably FDG-avid histologies.[51]

Interim Evaluation Response

Some guidelines recommend interim imaging to assess treatment response and disease progression. PET shows metabolic response earlier than anatomic response and thus is considered superior to CT alone to assess early response.[34] Interim PET-CT imaging may facilitate the ability to discriminate between those for whom additional or intensified treatment may be important and those for whom additional therapy may not be necessary or some forms of therapy (e.g. radiation) may not be needed. This may permit optimization of therapy and maintenance of treatment efficacy while decreasing or avoiding treatment-related side-effects or sequelae. Escalation of treatment in appropriate patients may improve survival. De-escalation of treatment may reduce treatment-related toxicity and result in cost-savings. Furthermore, there are data suggesting that ^{18}FDG-PET may be a predictor of prognosis when performed early during treatment. Across 13 studies summarized in a recent narrative review, interim PET (performed after 1 to 4 cycles of treatment) generally predicted better progression-free survival (PFS) in those with negative interim results (PFS range 71% to 95%) a lower PFS in those with positive results (range 13% to 100%)[108] depending on type and stage of lymphoma.

End-of-Treatment Response

When PET/CT is performed after completion of treatment, the primary purpose is to detect residual active lymphoma (i.e., to confirm or exclude complete remission according to the response criteria). Persistent disease is associated with early relapse and poor clinical outcome, and thus further therapy may be indicated.[55] As mentioned previously, one frequent difficulty in the interpretation of posttreatment PET/CT is differentiating residual FDG uptake due to lymphoma from FDG uptake due to posttreatment inflammation (which can last as long as 2 weeks after completion of chemotherapy and as long as 3 months after chemoradiation therapy), coexisting infection, and normal physiologic metabolic activity. New response evaluation criteria in lymphoma (RECIL)[162] were developed in 2017 which include updated response definitions as well as the additional of a new minor response category to address atypical response patterns related to several of the newer investigational treatments.

Surveillance
PET is generally not used for routine surveillance following treatment completion, largely due to concerns regarding false-positive findings and lack of impact on patient outcomes.[51,117]

PET-adapted trials

Recently, studies have emerged which explore the role of interim PET findings for response-adapted therapy to identify individuals with a high- or low-risk of progression or relapse. To evaluate the prognostic value of PET-CT, studies comparing treatment given based on PET-CT results with treatment that would have been given without PET-CT information are of primary interest. Currently, changing treatment on the basis of interim PET-CT is not recommended unless there is clear evidence of progression.[34] Interim PET is not generally performed for indolent lymphoma. There are data suggesting that FDG-PET may be a predictor of prognosis when performed early during treatment.

2.2.2. Response Criteria

Response criteria for PET/CT is evolving. The 5-point scale (5-PS) (previously called the Deauville scale) is a validated, flexible, treatment response scoring system adopted as the preferred reporting method by the First International Workshop on PET in Lymphoma (Deauville, France, 2009) and later recommended in the 2014 Consensus Statement of the International Conference on Malignant Lymphomas working group.[34] The scale is based on visual interpretation of FDG-uptake in two reference organs, the mediastinum and the liver; scores range from 1 (no uptake) to 5 (markedly increased uptake or any new lesions). The most recent classification system, the Lugano criteria,[50,162] was developed to simplify, update and standardize staging and response assessment criteria; it recommends the use of the Deauville 5-point scale (5-PS) for assessing treatment response, recommends modification to the Ann Arbor staging and makes other recommendations related to the interpretation of PET/CT scans and defining the role of PET/CT in interim and end of treatment assessment. A more recently proposed method of response assessment, standardized uptake value (SUV_{max}), attempts to provide a more quantitative means to evaluate response; it is a measure of glucose metabolized by a tumor as detected by an FDG-PET/CT system that assesses the tumor's aggressiveness (or FDG-avidity). It accounts for tumor activity through regional radioactive tracer uptake, and adjusts based on factors such as patient weight, variability from inconsistent image capturing and processing, and physical and biological sources of error.[96] The studies included in this report primarily report 5-PS but vary in their threshold for PET positivity; details of the scoring criteria used by the studies are provided whenever possible.

2.2.3. Comparators

Computed Tomography (CT)

Computed tomography (CT) is an imaging test that uses a combination of x-rays and a computer to create thin slice (1-10 millimeters) digital images of the region of the body studied. It is similar to a plain x-ray however it takes lots of pictures of a section (cross sections) of the body instead of just one; these image slices can either be displayed individually or digitally "stacked" together to generate a three-dimensional image. Computed tomographic images are always obtained in the axial plane (the plane perpendicular to the long axis of the body). Today most CT systems are capable of "spiral" (also known as "helical") scanning as well as scanning in the formerly more conventional "axial" mode. In addition, many CT systems are capable of imaging multiple slices simultaneously. Such advances allow relatively larger volumes of anatomy to be imaged in less time. Computed tomography can be performed from the brain to the extremities and can be performed without or with the administration of intravenous contrast material. Contrast material represents a non-radioactive, iodine containing chemical compound that increases x-ray absorption in the tissues in which it accumulates. Disease is detected on CT by alterations in the normal anatomy compared either to expected normal patients or to previous studies of the same patient. Other information obtained on CT images that help in detecting disease includes changes in tissue density and level of enhancement from intravenous contrast compared to normal. For lymphoma, CT scans are used for initial diagnosis and staging (to help detect enlarged lymph nodes or extralymphatic cancers) and are also used to monitor a tumor's response to therapy or detect a return of cancer after treatment.

Magnetic Resonance Imaging (MRI)

Magnetic resonance imaging (MRI) is an imaging test that uses radio waves and a computer in the presence of a strong magnetic field to create thin slice digital images of the region of the body studied. An MRI scan differs from CT scans and X-rays, as it does not use potentially harmful ionizing radiation. MRI images are particularly well suited to evaluate the non-bony parts or soft tissues of the body and can be obtained in any plane (axial, sagittal, coronal or off axis). Like CT, MRI images are digital and can be similarly manipulated on a computer, can be performed from the brain to the extremities and can be performed without or with intravenous contrast. The contrast agent for MRI is a non-radioactive, gadolinium containing chemical compound that alters the magnetic field of tissues in which it accumulates. Magnetic resonance imaging demonstrates cross sectional anatomy and changes in proton density. Disease detection relies on changes in anatomy, proton density and contrast enhancement. MRI techniques such as whole-body MRI and diffusion-weighted imaging may be good radiation-free alternatives to FDG-PET/CT for the staging and response assessment of HL and NHL, which may be particularly relevant for children.[155]

2.2.4. Potentional complications and harms of PET/CT

Overall, PET/CT is considered to be a safe diagnostic procedure. It does, however, involve radiation exposure, although the radiation risk is considered very low compared with the potential benefit of the procedure. According to the American Cancer Society, a PET/CT exposes an average-sized adult to an estimated 25 millisieverts (mSv) of radiation which is equal to about 8 years of average background radiation exposure. Other potential risks of the procedure include slight pain and redness at the

radiotracer injection site (which should rapidly dissipate) and allergic reactions to the radioactive material; the latter are extremely rare and are usually mild.

In contrast, harms related to the toxicities of the various treatments for lymphoma are common and may be severe and are considered more concerning.

2.3. Clinical Guidelines, Consensus Statements & Appropriateness Criteria

PubMed was searched for guidelines related to the use of PET imaging for lymphoma. A key word search was conducted utilizing the following terms (PET imaging [tiab] OR "Positron Emission Tomography Computed Tomography"[Mesh] AND lymphoma [TIAB] OR lymphoma [mesh]; Filters: Guidelines, Practice Guidelines, and English) and yielded 76 citations going back to 2011. Additionally, the National Guidelines Clearing House (prior to its defunding), Google, Google Scholar were all searched, and hand-searching was also conducted, which together yielded an additional 16 citations. In total, after review, 40 guidelines (including appropriateness criteria and consensus statements, further detailed below) were included and summarized, including key documents from the following organizations:

- **International Conference on Malignant Lymphomas (ICML) Imaging Working Group (Consensus Statement [34] and Summary of Recommendations [51])**
- **National Comprehensive Cancer Network (NCCN) Guidelines**
- **NICE UK Guidelines**
- **British Journal of Haematology (BJH)**
- **European Society for Medical Oncology (ESMO)**
- **American College of Radiology (ACR) Appropriateness Criteria**
- **Alberta Health Services (AHS)**
- **Australian Government Medical Services Advisory Committee (MSAC)**
- **Clinical Oncology Working Group (Herst 2017)**
- **AIM Specialty Health**

Additional guidelines from other organizations, including a full summary of each guideline, and a bibliography of all included guidelines can be found in Appendix L. For each guideline, key recommendations related to PET staging and response assessment in lymphoma were abstracted, the relevant evidence base was summarized, and the strength of recommendations was characterized. Data from across available guidelines was then arranged and condensed to evaluate consensus across treatment assessment stages within each lymphoma subtype. In the abridged summary table below (Table 2), the consensus of recommendations related to the types of lymphoma most commonly encountered in this report are presented alongside the overall strength of the recommendations and the level of evidence used to inform those recommendations, when available. Additional summaries of the ICML consensus statement and American College of Radiology Appropriateness Criteria follow. A complete table of consensus across all lymphoma subtypes can be found in Appendix L.

Table 2. Abridged summary of guideline recommendation consensus across lymphoma subtypes

Condition [number of guidelines]	Areas of Consensus	Recommendation Consensus (Strength of Recommendation)				
		Initial Staging	Interim	End of Treatment	Restaging/Relapse or Refractory	Surveillance/Follow-up
FDG-Avid Lymphomas						
FDG-Avid Lymphomas* [2] [34,51]	Recommendation	Yes (++)	Yes (++)	Yes (++)	Yes (++)	N (-)
	Quality of Evidence	Low	Low	NR	NR	NR
Hodgkin Lymphoma						
classical HL [6] [9,10,61,64,92,131,149,159]	Recommendation	Yes (++/-)	Mixed (++/+++)	Yes (-)	Yes (++/-)	Mixed (++/-)
	Quality of Evidence	Low-Moderate	Low-High	NR	Low or NR	Low or NR
Nodular Lymphocyte Predominant HL (NLPHL) [2] [92,106]	Recommendation	Yes (+)	No (+++)	Mixed (++/+++)	Yes (++)	No (++)
	Quality of Evidence	Moderate	Low	Low	Low	Low
Aggressive Non-Hodgkin Lymphoma (aNHL)						
aggressive/High-Grade Non-Hodgkin Lymphomas* [2] [5,9]	Recommendation	Yes (-)	Mixed (-)	Mixed (-)	Mixed (-)	No (-)
	Quality of Evidence	NR	NR	NR	NR	NR
Diffuse Large B-Cell Lymphomas (DLBCL) [5] [5,44,57,151,163]	Recommendation	Yes (++/+++)	Mixed (+++/-)	Yes (+++)	Mixed (+)	No (++/+++)
	Quality of Evidence	Low-Moderate	Low-High	High	Moderate	Low
Indolent NHLs						
Indolent NHLs* [4] [5,9,54,91]	Recommendation	Yes (-)	NR	Yes (-)	Mixed (-)	No (-)
	Quality of Evidence	NR	NR	NR	NR	NR
Follicular Lymphoma [4] [5,10,62,163]	Recommendation	M (+/++)	NR	U (++)	NR	N (-)
	Quality of Evidence	Low or NR	NR	Moderate	NR	NR

HL = Hodgkin Lymphoma; NHL = non-Hodgkin Lymphoma; NR = not reported

Strength of Recommendation: + = Low or weak; ++ = Moderate; +++ = High; - = No evidence or not reported; +/++ = ranges between the levels shown

* In cases where a guideline reports on an overarching category of lymphoma (such as all FDG-avid Lymphomas, which includes all HL subtypes and a variety of NHL types, or aggressive or high grade NHL, etc.), its recommendations are not necessarily repeated or tabulated in the summary of consensus for more specific subtypes within this table, even if their recommendations apply.

International Consensus Statements

In 2014, a consensus statement was released that summarized the recommendations reached by the International Conference on Malignant Lymphoma Imaging Working Group and updated guidance from the International Harmonization Project regarding the use of PET in the staging and response assessment for lymphoma.[34] A summary of recommendations presented by response assessment stage is provided in Table 3 below.

Table 3. Summary of international consensus statements on PET imaging for lymphoma

Guideline	Evidence Base	Recommendations*	Rating/Strength of Recommendation
Consensus Statement - International Conference on Malignant Lymphomas Imaging Working Group 2014 [34] And **Recommendations for Initial Evaluation, Staging, and Response Assessment of Hodgkin and Non-Hodgkin Lymphoma: The Lugano Classification (Cheson et al., 2014)**[51] FDG-Avid Lymphomas (Hodgkin's and Non-Hodgkin's Lymphomas) *Barrington et al., Role of imaging in the staging and response*	Literature review and Expert Consensus Staging: 19 key studies comparing PET or PET/CT with CT alone, type NR (1248 patients) Interim PET: 8 studies (type NR) for HL, 7 studies (type NR) for aggressive NHL, 1 clinical trial (type NR), 8 studies (type NR) for DLBCL End of Treatment PET: Previously cited studies and 8 additional	**FDG-avid, nodal lymphomas** <u>Staging</u> PET-CT is preferred for staging of FDG-avid lymphomas, and CT scan is preferred in the other lymphomas. 1. PET-CT should be used for staging in clinical practice and clinical trials but is not routinely recommended in lymphomas with low FDG avidity; PET-CT may be used to select best site to biopsy **(type 1)** 2. Contrast-enhanced CT when used at staging or restaging should ideally occur during single visit combined with PET-CT, if not already performed; baseline findings will determine whether contrast-enhanced PET-CT or lower-dose unenhanced PET-CT will suffice for additional imaging examinations **(type 2)** 3. Bulk remains an important prognostic factor in some lymphomas; volumetric measurement of tumor bulk and total tumor burden, including methods combining metabolic activity and anatomical size or volume, should be explored as potential prognosticators **(type 3)** <u>Interim PET</u> PET-CT should be used for interim (early) response assessment in FDG-avid histologies, using the 5-point scale; CT is preferred for low or variable FDG avidity. 1. If mid-therapy imaging is performed, PET-CT is superior to CT alone to assess early response; trials are evaluating role of PET response–adapted therapy; currently, it is not recommended to change treatment solely on basis of interim PET-CT unless there is clear evidence of progression **(type 1)** 2. Standardization of PET methods is mandatory for use of quantitative approaches and desirable for routine clinical practice **(type 1)**	Recommendations were formulated as follows: * **Type 1**: Based on established current knowledge * **Type 2**: To identify emerging applications * **Type 3**: To highlight key areas requiring further research

Guideline	Evidence Base	Recommendations*	Rating/Strength of Recommendation
assessment of lymphoma: consensus of the International Conference on Malignant Lymphomas Imaging Working Group. Journal of clinical oncology. 2014 Sep 20;32(27):3048.	studies (type NR) Additional citations for each stage are reported in Cheson et al., 2014	3. Data suggest that quantitative measures (eg, SUVmax) could be used to improve on visual analysis for response assessment in DLBCL, but this requires further validation in clinical trials **(type 2)** <u>End of Treatment</u> PET-CT should be used for end of treatment response assessment in FDG-avid histologies, using the 5-point scale; CT is preferred for low or variable FDG avidity. 1. PET-CT is standard of care for remission assessment in FDG-avid lymphoma; in presence of residual metabolically active tissue, where salvage treatment is being considered, biopsy is recommended **(type 1)** 2. Investigation of significance of PET-negative residual masses should be collected prospectively in clinical trials; residual mass size and location should be recorded on end-of-treatment PET-CT reports where possible (**type 3**) 3. Emerging data support use of PET-CT after rituximab-containing chemotherapy in high–tumor burden FL; studies are warranted to confirm this finding in patients receiving maintenance therapy **(type 2)** 4. Assessment with PET-CT could be used to guide decisions before high-dose chemotherapy and ASCT, but additional studies are warranted **(type 3)** <u>Surveillance</u> Surveillance scans after remission are discouraged, especially for DLBCL and HL, although a repeat study may be considered after an equivocal finding after treatment. Judicious use of follow-up scans may be considered in indolent lymphomas with residual intra-abdominal or retroperitoneal disease.	

ASCT = autologous stem cell transplant; DLBCL, diffuse large b-cell lymphoma; FDG = fluorodeoxyglucose; HL = Hodgkin lymphoma; NHL = non-Hodgkin Lymphoma; NR = not reported; PET/CT = positron emission tomography / computerized tomography

American College of Radiology Appropriateness Criteria

Appropriateness criteria from the American College of Radiology are considered alongside clinical guidelines in Appendix L, however, a summary is included here for reference (Table 4). In Hodgkin Lymphoma:

Interim PET

- As of 2016, interim PET response assessment is considered potentially appropriate for favorable prognosis early-stage Hodgkin Lymphoma, but preferably conducted in the context of a clinical trial.[61]
- As of 2015, the removal of radiation therapy is considered appropriate when there is a rapid early response after 2 cycles of ABVE-PC in pediatric, adolescent, and young adult patients with unfavorable stage I and II prognosis, especially when there is[131]:
    - a negative PET scan and
    - complete response by CT criteria at end of treatment

Surveillance

- Based on 2014 criteria for follow-up of Hodgkin Lymphoma, routine use of PET for disease surveillance is not recommended.[86]
- No other PET-specific recommendations are made for advanced-stage, recurrent or pediatric Hodgkin Lymphoma.[122,149,159]

Regarding non-Hodgkin Lymphomas, ACR criteria from 2013 and 2014 related to diffuse large B-cell lymphomas (DLBCL) and localized nodal indolent lymphomas determine that:

Initial Staging and End of Treatment Response Assessment

- PET/CT is well established with appropriate routine use for the initial staging and end of treatment assessment (using 5-point Deauville scale) for patients with DLBCL; PET/CT results can eliminate the need for bone-marrow biopsy if bone-marrow involvement is indicated on scans. Conversely, if bone marrow involvement is not indicated by PET/CT, then bone-marrow biopsy should still be performed if relevant to clinical trial enrollment or management of the patient. PET/CT is also considered appropriate for initial staging of localized nodal indolent lymphomas based on its use for demonstrating more disseminated disease, upstaging patients thought to have localized involvement, and in RT treatment planning. No recommendations on end of treatment PET were made for localized nodal indolent lymphomas.

Interim PET

- Despite some initial evidence of the use of interim PET/CT in DLBCL, interim PET is not recommended to influence clinical decisions until further confirmatory studies become available. No recommendations on interim PET were made for localized nodal indolent lymphomas.

Table 4. Summary of American College of Radiology Appropriateness Criteria

Guideline	Evidence Base	Recommendations	Rating/ Strength of Recommendation
American College of Radiology Appropriateness Criteria: Hodgkin Lymphoma-Favorable Prognosis Stage I and II (2016) [61]	Systematic Literature review (67 references total) and Expert Consensus Response Adapted Therapy: 4 studies, 1 RCT, 3 ongoing trials (EORTC H10F, RAPID trial, and GHSG HD16 trial)	**Favorable Early-Stage Hodgkin Lymphoma** Interim For favorable prognosis stage I and II HL, changing chemotherapy or omitting RT based on PET response for early-stage patients may be appropriate with caution after careful consideration and preferably as part of a clinical trial, at least until preliminary data mature.	NR
American College of Radiology Appropriateness Criteria: Hodgkin Lymphoma-Unfavorable Clinical Stage I and II (2015) [131]	Literature Review (83 references total) and Expert Consensus 3 RCTs, international consensus panel (Barrington et al 2014), 1 study (type NR)	**Unfavorable Early-Stage Hodgkin Lymphoma** End of Treatment The pediatric regimen ABVE-PC can be used in pediatric, adolescent, and young adult patients without RT when there is a rapid early response after 2 cycles, especially if defined by a negative PET scan and if at the end of chemotherapy there is a complete response by CT criteria.	NR
American College of Radiology Appropriateness Criteria: Follow-up of Hodgkin Lymphoma (2014) [86]	Literature review and Expert Consensus 4 studies (type NR), 1 cost-effectiveness analysis, 1 prospective study, 3 retrospective studies	**Hodgkin Lymphoma** Surveillance Though PET scan is a useful tool in defining a subset of patients who require additional therapy after completion of the initial therapy, the routine use of PET for surveillance is not recommended in general due to low positive predictive value, high false-positive rate, and unfavorable cost-effectiveness.	NR
American College of Radiology Appropriateness Criteria: Hodgkin Lymphoma-Stage III and IV (2016) [122]	NA	**Advanced-Stage Hodgkin Lymphoma** No PET-specific Recommendations	NR
American College of Radiology Appropriateness Criteria: Pediatric Hodgkin Lymphoma (2012) [149]	NA	**Pediatric Hodgkin Lymphoma** No PET-specific Recommendations	NR

Guideline	Evidence Base	Recommendations	Rating/ Strength of Recommendation
American College of Radiology Appropriateness Criteria: Recurrent Hodgkin Lymphoma (2016) [159]	NA	**Recurrent Hodgkin Lymphoma** No PET-specific Recommendations	NR
American College of Radiology Appropriateness Criteria: Diffuse Large B-Cell Lymphoma (2014) [57]	Systematic Literature review (90 references) and Expert Consensus Diagnosis, Staging and Response Evaluation: 7 prospective and retrospective studies, type NR Interim: 2 Case series, 3 studies type NR, 1 retrospective study	**Diffuse Large B-Cell Lymphoma (DLBCL)** Initial Staging: Review of data suggests that PET/CT is well established and should be routinely used for the evaluation of patients with DLBCL. Bone-marrow biopsy is not performed if bone-marrow involvement is indicated on PET/CT conversely, if bone marrow involvement is not indicated by PET/CT, then bone-marrow biopsy should be performed if relevant to clinical trial enrollment or management of the patient Interim (mid-treatment) Some studies suggest that interim PET/CT can predict outcome and thus allow for modifications of therapy (de-escalation or escalation); however, until further confirmatory studies become available, interim PET/CT should not influence clinical decisions. End of Treatment End-of-therapy assessment is to be performed using PET/CT, and the current recommendation is to grade the PET/CT using the 5-point Deauville scale.	NR
American College of Radiology Appropriateness Criteria: Localized Nodal Indolent Lymphoma (2013) [91]	Literature Review (46 references total) and Expert Consensus At least Two studies (type NR), other guidelines	**Localized Nodal Indolent Lymphoma** Initial Staging FDG-PET/CT scans can demonstrate more disseminated disease and upstage patients thought to have localized involvement. Additionally, they can be useful in RT treatment planning.	NR

NR = not reported;

2.4. Previous Systematic Review/Technology Assessments

The 2016 Australian Medical Services Advisory Committee (MSAC)[54] report on the efficacy of FDG-PET for indolent non-Hodgkin's lymphoma, an update to a previous MSAC report (2010) [53] on FGD-PET for all lymphomas, included five studies investigating the clinical effectiveness of PET/CT and one study evaluating the prognostic value of PET/CT (Table 5). Due to the lack of available direct evidence, a linked evidence approach was used, focusing on evidence of comparative diagnostic, prognostic, and therapeutic efficacy and effectiveness for data on initial staging, response to treatment, and restaging. Initial staging was informed by three studies reporting on the effectiveness of identifying additional disease sites compared to bone marrow biopsy. Response to therapy at end of first line treatment was informed by one study where PET showed higher specificity than CT, whereas a similar specificity and higher sensitivity than CT was found in another study informing restaging. Although no studies of clinical effectiveness for prognosis were found, PET/CT did show superior prognostic value in one comparative pooled analysis prognostic study. Results of economic model suggest that while PET/CT may be modestly cost-saving, it may modestly increase costs due to increased use of immunochemotherapy. Overall, the evidence base was considered weak and unlikely to improve due to the rareness of indolent non-Hodgkin's Lymphoma. The final coverage determination supported FDG-PET/CT in place of CT for indolent non-Hodgkin's lymphoma for initial staging, end of treatment assessment of response and restaging following confirmation of recurrence.

Table 5. Summary of MSAC 2016 Health Technology Assessment

Assessment (year) and search dates	Reference standard for PET and critical appraisal	Evidence base and Outcomes	Conclusions
FDG-PET			
MSAC (2016) [54] Search dates NR	Reference standard: pathology or clinical follow-up Critical appraisal:Study-level RoB not reported; overall SoE assessed with GRADE	Overall <u>Comparative Safety</u> PET was determined to be a safe procedure and no new safety concerns since the 2010 MSAC report were raised by studies included in assessment <u>Clinical Effectiveness</u> No direct evidence, only linked evidence in the form of: -Comparative diagnostic performance (5 studies, rated as VERY LOW quality evidence) -Comparative prognostic evidence (1 study: Trotman et al., 2014) <u>Clinical Management</u> Two studies reported a change in patient management based on PET/CT at initial staging (Scott 2009, Fulham 2006), but none for restaging or interim PET.	Overall <u>Diagnostic Performance</u> PET/CT detects more disease sites than CT; PET/CT detects more true responders to treatment than CT but also results in more misclassification of responders to treatment <u>Overall Coverage Conclusion</u> Positron emission tomography (18F-FDG PET) is supported in place of CT for indolent non-Hodgkin's lymphoma for initial staging, end of treatment assessment of response and restaging following confirmation of recurrence. <u>Clinical Effectiveness</u>

Assessment (year) and search dates	Reference standard for PET and critical appraisal	Evidence base and Outcomes	Conclusions
		Cost Effectiveness One cost-consequence analysis (Luminari et al., 2013) indicated less cost per patient if PET/CT replaced CT for indolent NHL, however MSAC noted that there were multiple areas of uncertainty in the estimates, and noted that the impact of this was likely to be small given that indolent NHL is a uncommon cancer. If the economic model were accepted, PET/CT may be modestly cost-saving to the MBS, but may modestly increase costs to the PBS due to increased use of immunochemotherapy. The majority of assumptions included in the model appeared to be based almost entirely on expert opinion rather than being evidence based. Initial Staging Evidence: 3 studies (Adams 2013, Lee 2015, Perry 2016); SoE: Very Low Outcomes: Identification of additional lesions; End of Treatment Evidence base: 1 study (Le Dortz 2010), SoE: Very Low Outcomes: Sensitivity: PET/CT (72%, 95% CI: 0.47-0.90) vs CT (83%, 95% CI: 0.59-0.96) Specificity: PET/CT (100%, 95% CI: 0.87-1.00) vs CT (52%, 95% CI 0.32-0.71) Restaging (Relapsed/Refractory) 1 study, (Fueger 2009), SoE: Very low Outcomes: Sensitivity: PET/CT 0.77, 95% CI 0.69-0.84 versus CT 0.54, 95% CI: 0.45-0.63 Specificity: Same as CT (0.98, 95% CI: 0.97-0.99) .	While the evidence base for clinical effectiveness is weak, it is unlikely to improve as indolent NHL is an uncommon cancer Initial staging PET/CT has the same specificity as CT and higher sensitivity than CT for detecting nodal regions or a per-lesion basis Coverage Conclusion FDG PET/CT is supported in place of CT for staging of all newly diagnosed and relapsed indolent non-Hodgkin's lymphoma. End of Treatment PET PET/CT has higher specificity than CT Although evidence was not directly comparable, End of treatment PET/CT was weakly predictive of PFS but not OS, and was more predictive than CT based response in multivariate analysis Coverage Conclusion FDG PET/CT is supported in place of CT for indolent non-Hodgkin's lymphoma for assessment of response to therapy Restaging (Relapsed/Refractory) PET/CT had the same specificity and higher sensitivity than CT There was insufficient evidence to determine the effectiveness of restaging PET/CT in modifying therapy Coverage Conclusion FDG PET/CT is supported in place of CT for indolent non-Hodgkin's

Assessment (year) and search dates	Reference standard for PET and critical appraisal	Evidence base and Outcomes	Conclusions
		Prognosis Evidence: 1 study, (Trotman 2014) SoE: NR Outcomes: PFS: HR 3.9 (95% CI 2.5 to 5.9), p<0.0001 OS: HR 6.7 (95% CI 2.4 to 18.5), p=0.001 Therapeutic Efficacy (Change in Clinical Management) Initial Staging Evidence: 2 studies, (Scott 2009, Fulham 2006) SoE: NR Outcomes: Change in management plan: 34% of patients (95% CI 24 to 45) Change from RT to chemotherapy: 8% Change from chemotherapy to observation: 3% Change from RT to observation: 5% Information on the impact of these therapeutic changes was not reported. No other evidence provided for therapeutic efficacy or effectiveness (changes in clinical management). Surveillance No evidence reported	lymphoma after confirmation of recurrence. Prognosis Diagnostic Accuracy There were no studies of clinical effectiveness although PET/CT did show superior prognostic value. Coverage Conclusion The value of prognostic information to patients was difficult to measure and to cost. Nevertheless, it may have a significant impact on quality of life (patient reassurance, making important life decisions etc.) and on societal costs (patients more likely to remain in the workforce). This information may be of particular value in a disease which may have a long clinical course, which affects people in late middle age, and which patients may die with rather than from. Surveillance Coverage Conclusion PET/CT should not be used for surveillance of patients, which was seen as not clinically justified and would significantly increase the frequency of utilization per patient, and recommended monitoring this item for inappropriate or excessive use and also for co-claiming with other imaging services.

CI = confidence interval; CT = Computerized Tomography; FDG = fluorodeoxyglucose; HR = hazard ratio; MSAC = Medical Services Advisory Committee; NR = not reported; PET/CT = Positron Emission Tomography / Computerized Tomography; RT = radiation therapy; SoE = Strength of Evidence;

Summary of Systematic Reviews of PET-adapted therapy:

Two systematic reviews evaluated the effects of PET-adapted therapy on long-term outcomes for patients with advanced-stage HL (Table 6). [25,136] Studies contained in these reviews that met the inclusion criteria for this HTA were included (excluding those published during the span of search dates reviewed and considered by the prior report). Sickinger et al. (2015), pooled results from 3 RCTs (encompassing 4 comparisons) to evaluate the impact of PET-adapted therapy in persons with negative interim PET. A total of 1,480 patients (out of 1,999 total patients) were included in the analysis. Overall, although they found moderate quality evidence that PET-adapted therapy (chemotherapy without radiation therapy) resulted in shorter PFS in interim PET negative patients (HR 2.39, 95% CI 1.62 to 3.50) compared to standard care (with radiation therapy), they concluded that it was still uncertain whether PET positive individuals benefit from treatment adaptation in those with advanced-stage HL. They determined that there was insufficient data to make conclusions about OS, adverse events, response to treatment, treatment-related mortality, or quality of life. A more recent systematic review (Amitai et al.) included thirteen studies (N=6856 across 17 publications, including 4 RCTs). In the four RCTs reviewed (n=4920), randomization was not conducted for PET vs. no PET (or another comparator), instead, patients were randomized to different treatments based on PET results with escalation of therapy done for PET positive and/or de-escalation for PET negative. Due to the heterogeneity of the included studies, the authors were unable to pool the results. They note that none of the RCTs of escalated therapy conducted a true randomization between PET-adapted escalation of treatment and a standard treatment regimen of ABVD and that baseline prognostic factors and stage affect prognosis. They concluded that, based primarily on RCT evidence, an early PET-adapted treatment approach should become the standard of care for patients with advanced-stage HL.

Table 6. Previous Systematic Reviews for PET-adapted therapy

Assessment (year), search details	Reference standard and critical appraisal	Objective Evidence base and Outcomes	Conclusion
FDG-PET			
Amitai (2018)[25] Search dates 1990-2017 Cochrane CENTRAL, and PubMed	Reference standard: clinical follow-up Critical appraisal: Cochrane Collaborations tool for RCTs (RCTs rated at unclear or low risk of bias); ROBINS-I tool for non-RCTs; No overall SoE described	Objective: Evaluate the effect of PET-adapted therapy on long-term outcomes (OS and PFS) in patients with advanced-stage HL. Overall <u>Evidence base:</u> 13 studies (17 publications, N=6856 patients) (4 RCTs , 7 prospective, 2 retrospective,) De-escalation of therapy if PET was negative <u>Evidence base:</u> 4 studies (1 RCT, 1 prospective, 1 retrospective, 1 prospective and retrospective) <u>Outcomes (1 RCT):</u>	Due to heterogeneity, no meta-analysis was conducted. A systematic review of each study is provided instead, with conclusions focusing on data from RCTs detailed below: <u>Overall</u> This systematic review suggests (based primarily on evidence from RCTs) that a change to the treatment paradigm, with interim PET-adapted approach as standard of care, is appropriate for advanced-stage HL. Present evidence shows that PET-guided escalation and de-escalation is warranted in certain cases, with implications on safety and treatment planning:

Assessment (year), search details	Reference standard and critical appraisal	Objective Evidence base and Outcomes	Conclusion
		2-year PFS (Casasnovas 2018) PET+ vs PET-: 72.9% vs 92.8%, p<0.0001 PET+ vs PET- (standard arm): 75.1% vs 94%, p<0.0001 PET+ vs PET- (experimental): 70.8% vs 91.6%, p<0.0001 <u>Outcomes (3 non-randomized studies):</u> 2-year PFS (range across studies) PET+: 47% to 73% PET- : 80% to 93% 2-year PFS (Deau 2015): PET+ vs PET- : 47% vs 87% 3-year OS (Pavlovsky 2016) PET+ & PET- : 96% 3-year EFS (Pavlovsky 2016): PET+ vs PET- : 57.1% vs 90% 5-year OS (Kedmi/Avigdor): PET+ vs PET - : 79% vs 98% 5-year PFS (Kedmi/Avigdor) PET+ vs PET - : 60% vs 80% Escalation of therapy if PET was positive <u>Evidence base:</u> 5 studies (1 RCT, 3 prospective, 1 retrospective) <u>Outcomes (1 RCT):</u> 4-year OS PET + vs PET-: 89% vs 97% 4-year PFS PET + vs PET- : 69% vs 87% <u>Outcomes (4 non-randomized studies):</u> 2-year OS (range across 4 studies) Entire cohort: 88% to 98%	-addition of rituximab did not result in better outcomes in studies reviewed -higher IPS and more advanced-stage lower the NPV of an iPET- after two ABVD cycles but not after 2 eBEACOPP cycles -Completing 4 eBEACOPP cycles in iPET- patients was as effective as completing 6 cycles and found to be less toxic. Further RCTs are needed. Novel combinations and chemical agents should be investigated further, as should adjunct methods of predicting treatment failure alongside PET (such as biological stratification of disease characteristics at diagnosis via immunohistochemistry, gene expression analysis, etc.). <u>Detailed Conclusions by Category</u> De-escalation therapy if PET was negative De-escalation in PET-adapted therapy was equally efficient as non-adapted therapy and better tolerated; PFS show similar between standard and experimental arms across 3 de-escalating RCTs. Escalation therapy if PET was positive Neither of the two PET+ escalation RCTs truly randomized between PET-adapted and standard control arm, however results indicated that: -adding rituximab didn't result in superior outcomes in 2 RCTs that escalated on PET+; -escalating to BEACOPP from ABVD was a feasible strategy based on present data; -PFS in iPET- groups that completed a full course of ABVD was substantially lower than historic norm (95%); this was consistent across different groups and either implies that iPET

Assessment (year), search details	Reference standard and critical appraisal	Objective Evidence base and Outcomes	Conclusion
		2-year PFS or EFS (range across studies) PET+: 50% to 76% PET- : 81% to 82% 2-year PFS PET+ vs PET- (Press): 64 vs 82% PET+ vs PET- (Zinzani): 76 vs 81% 2-year EFS PET+ vs PET- (Ganesan): 50 vs 82% 2-year FFS PET+ vs PET- (Gallamini): 65% vs 92% escalation and de-escalation according to PET results <u>Evidence base:</u> 4 studies (2 RCTs, 2 prospective) <u>Outcomes (2 RCTs):</u> 3-year OS (Johnson 2016) PET positive: 87.8% PET negative (standard arm): 97.2% PET negative (experimental de-escalation): 97.6% 5-year OS (Johnson 2016) PET+: 85.1% PET- (standard arm): 95.3% PET- (experimental de-escalation): 95% 5 year OS (Borchmann 2017) PET+ vs PET- (standard arms): 96.4% vs 95.4% PET+ (experimental escalation): 93.9% PET- (experimental de-escalation): 98% 3-year PFS (Johnson 2016) PET+: 67.5% PET- (standard arm): 85.4%	scans have inherent technical limitations or that iPET- patients are a heterogeneous group and that a proportion of iPET- patients might benefit from escalation. Although direct comparisons between studies is impossible due to heterogeneity, results seem to indicate that negative iPET has the highest NPV in patients with less extensive disease at presentation and in those treated with most intensive chemotherapy Escalation and de-escalation according to PET results The 2 RCTs of escalation and de-escalation suggest that Deuville score of 5, as opposed to 4, are at a particularly higher risk of relapse; this implies that PET+ patients comprise a heterogeneous group and that strategies other than escalation of therapy should be examined, such as early myeloablative therapy and novel biological combinations. Classical prognostic factors other than PET (IPS and stage) still affect prognosis as evidenced by lower NPV of iPET scans in patients more advanced-stage or higher IPS in RATHL study and lower likelihood of PET negativity in higher IPS patients (GITIL study)

Assessment (year), search details	Reference standard and critical appraisal	Objective Evidence base and Outcomes	Conclusion
		PET- (experimental de-escalation): 84% 5-year PFS (Johnson 2016) PET+: 65.7% PET- (standard arm): 82.7% PET- (experimental de-escalation): 80.6% 5 year PFS (Borchmann 2017) PET+ vs PET- (standard arms): 89.7% vs 90.8% PET+ (experimental escalation): 88.1% PET- (experimental de-escalation): 92.2% <u>Outcomes (2 non-randomized studies):</u> 5-year PFS (Dann 2017) PET+ vs PET- (IPS <3): 59% vs 80% PET+ vs PET - (IPS ≥3): 79% vs 81%	
Sickinger 2015 [136] Cochrane Review Search dates 1990-2014 Cochrane CENTRAL, and MEDLINE	Reference standard: Clinical diagnosis Critical appraisal: Risk of Bias tools and GRADE for overall SoE (Cochrane Handbook for Systematic Reviews of Interventions)	Objective: To assess the effects of interim PET treatment modification in individuals with HL <u>Evidence base</u>: 3 studies (3 RCTs, 1 with only PET-results available)) Data from 1,480 (out of 1,999) participants with PET negative results were analyzed based on participants randomized to receive either standard therapy (chemotherapy followed by radiotherapy) or PET-adapted therapy (chemotherapy only). <u>Outcomes:</u> PFS (moderate-quality evidence) standard therapy with RT vs chemotherapy without RT: HR	Overall From available data, it is unclear whether PET-positive individuals benefit from PET-adapted treatment modification in advanced-stage HL. No robust data on OS, Treatment-Related Mortality, Quality of Life, or short- and long-term adverse events was available (very low quality of evidence). Very low quality evidence of no difference in short-term adverse events between treatment arms. There was moderate quality evidence that PFS was shorter in individuals with early-stage HL and negative interim PET receiving PET-adapted (standard chemotherapy without radiotherapy) therapy compared to

Assessment (year), search details	Reference standard and critical appraisal	Objective Evidence base and Outcomes	Conclusion
		2.38 (95% CI 1.62 to 3.50), p<0.0001 standard RT vs no additional RT: HR 1.86 (95% CI 1.07 to 3.23) p=0.03 RT vs no RT + chemotherapy: HR 3.00 (95%CI 1.75 to 5.14), p< 0.0001 Adverse events (treatment toxicity) (1 study, very low quality evidence): RR 0.91 (95% CI 0.54 to 1.53), p=NS between groups Mortality (3, studies, very-low quality evidence) -PET-adapted vs standard therapy: 2 vs 2	standard therapy (including radiation), and confirmed with sensitivity analysis using random-effects model (HR 2.69, 95%CI 1.46 to 4.96; p=0.001). From available evidence, this indicates that early-stage, PET-negative individuals did not benefit from PET-adapted omission of RT, but data is too sparse to make a clear statement on the impact on OS. Short PFS from omitting RT may still be of benefit however, sparing patients secondary malignancies or organ toxicities, but further longer follow-up research is needed to explore this and impact on other outcomes.

CI = confidence interval; CT = Computerized Tomography; eBEACOPP = Bleomycin, Etoposide, Adriamycin, Cyclophosphamide, oncovin, procarbazine, prednisone; EFS = event free survival; FDG = fluorodeoxyglucose; HL = hodgkin's lymphoma; HR = hazard ratio; iPET = interim Positron Emission Tomography; IPS = International Prognostic Score; MSAC = Medical Services Advisory Committee; NPV = negative predictive value; NR = not reported; OS = overall survival; PET/CT = Positron Emission Tomography / Computerized Tomography; PFS = progression free survival; RCT = randomized controlled trial; RR = risk ratio/relative risk; RT = radiation therapy; SoE = Strength of Evidence;

2.5. Medicare and Representative Private Insurer Coverage Policies

Coverage decisions are summarized briefly below (Table 7). The most recent Medicare National Coverage Determination, as well as two national and two regional payer policies (requirements for this report are to provide information on at least two bell-weather payers) were summarized, including:

- **Centers for Medicare and Medicaid Services (CMS) NCD (2014)**
  In 2014, CMS released an update to its 2010 NCD. CMS considers PET to be medically necessary and covered for Initial treatment strategy (formerly referred to as 'diagnosis' and 'staging') and nationally covers one PET study for biopsy-proven cancers or those strongly suspected by other diagnostic testing. Up to three additional PET scans are nationally covered when used to guide subsequent management of anti-tumor treatment strategy (formerly 'restaging' and 'monitoring response to treatment') after completion of initial anti-cancer therapy. The NCD covers lymphomas broadly and does not distinguish between types.
- **Aetna (2018)**
  Aetna considers FDG-PET medically necessary for the Hodgkin's Lymphoma and Non-Hodgkin's Lymphoma, under certain general and disease-specific criteria for diagnosis, staging, restaging and monitoring , when FDG-PET scan is necessary to guide management.
- **United Healthcare (2017)**
  United Healthcare based their decision on the most recently updated CMS NCD (2014, see above). They refer specifically to NCD for coverage criteria
- **BCBS Minnesota (2017) and Louisiana (2018) Regional Policies**
  No national coverage policy was found for the BCBS network, two BCBS regional policies are summarized below. Both consider FDG-PET as medically necessary under similar circumstances as the CMS NCD.

Table 7. Overview of Medicare and Payer Policies

Payer (year)	Evidence Base Available	Policy	Rationale/Comments
Centers for Medicare and Medicaid Services (2014) [4]	NR	CMS released an initial NCD in 2010, which was updated in 2014. FDG PET for Lymphoma covered for Initial Treatment Strategy and Subsequent Treatment Strategy as outlined below. <u>Initial Treatment Strategy Guidelines</u> CMS continues to nationally cover one FDG PET study for beneficiaries who have cancers that are biopsy proven or strongly suspected based on other diagnostic testing when the beneficiary's treating physician determines that the FDG PET study is needed to determine the location and/or	<u>Rationale:</u> "CMS continues to believe that the evidence is adequate to determine that the results of FDG PET imaging are useful in determining the appropriate initial anti-tumor treatment strategy for beneficiaries with suspected cancer and improve health outcomes and thus are

Payer (year)	Evidence Base Available	Policy	Rationale/Comments
		extent of the tumor for the following therapeutic purposes related to the initial anti-tumor treatment strategy: • To determine whether or not the beneficiary is an appropriate candidate for an invasive diagnostic or therapeutic procedure; or • To determine the optimal anatomic location for an invasive procedure; or • To determine the anatomic extent of tumor when the recommended anti-tumor treatment reasonably depends on the extent of the tumor. <u>Subsequent Anti-Tumor Treatment Strategy</u> Three FDG PET scans are nationally covered when used to guide subsequent management of anti-tumor treatment strategy after completion of initial anti-cancer therapy. Coverage of more than three FDG PET scans to guide subsequent management of anti-tumor treatment strategy after completion of initial anti-cancer therapy shall be determined by the local Medicare Administrative Contractors.	reasonable and necessary under §1862(a)(1)(A) of the Social Security Act (the Act)."
National Policies			
Aetna (2018)[8]	• 2 studies: Murphy et al., 2008; Dickinson et al., 2010; • 2 case reports: Funauchi et al., 2008; Kawano et al., 2012 • 3 SRs: Chen et al., 2011, with 8 studies; Adams et al., 2016 with 11 studies; Adams & Kwee 2016, with 11 studies; • 2 HTAs: Institute for	Aetna considers FDG-PET medically necessary for the Hodgkin's Lymphoma and Non-Hodgkin's Lymphoma, when the following general and disease-specific criteria for diagnosis, staging, restaging and monitoring are met, and the FDG-PET scan is necessary to guide management <u>General Criteria:</u> 1. Diagnosis The PET results may assist in avoiding an invasive diagnostic procedure, or the PET results may assist in determining the optimal anatomic location to perform an invasive diagnostic procedure. In general, for most solid tumors, a tissue diagnosis is made prior to the performance of PET scanning. PET scans following a tissue diagnosis are performed for the purpose of staging, not diagnosis. Therefore, the use of PET in the diagnosis of lymphoma, esophageal carcinoma, colorectal cancers, and melanoma is rarely considered medically necessary.	<u>Rationale:</u> NR

Payer (year)	Evidence Base Available	Policy	Rationale/Comments
	Clinical Systems Improvement, 2001; Clark et al., 2011 • 1 Clinical Guideline: NCCN 2015	2. Staging: PET is considered medically necessary in situations in which clinical management of the member would differ depending on the stage of the cancer identified and either: ▪ The stage of the cancer remains in doubt after completion of a standard diagnostic work-up, including conventional imaging (computed tomography, magnetic resonance imaging, or ultrasound); or ▪ The use of PET would potentially replace one or more conventional imaging studies when it is expected that conventional study information is insufficient for the clinical management of the member. 3. Re-staging: PET is considered medically necessary for re-staging after completion of treatment for the purpose of detecting residual disease, for detecting suspected recurrence in persons with signs or symptoms of recurrence, or to determine the extent of recurrence. Use of PET is also considered medically necessary if it could potentially replace one or more conventional imaging studies when it is expected that conventional study information is insufficient for the clinical management of the member. PET for post-treatment surveillance is considered experimental and investigational, where surveillance is defined as use of PET beyond the completion of treatment, in the absence of signs or symptoms of cancer recurrence or progression, for the purpose of detecting recurrence or progression or predicting outcome. 4. Monitoring: PET for monitoring tumor response during the planned course of therapy is not considered medically necessary except for breast cancer. Re-staging occurs only after a course of treatment is completed. <u>Lymphoma Disease-Specific Criteria:</u>	

Payer (year)	Evidence Base Available	Policy	Rationale/Comments
		Hodgkin's lymphoma: FDG-PET scans are considered medically necessary for the diagnosis*, staging and re-staging of Hodgkin lymphoma when the general medical necessity criteria for oncologic indications (II. A. listed above) are met. *Note: A diagnostic tissue sample is usually obtainable without PET localization. Therefore, PET for diagnosis of Hodgkin lymphoma is rarely considered medically necessary. Non-Hodgkin's Lymphoma (including post-transplant lymphoproliferative disorder and Castleman's disease): FDG-PET scans are considered medically necessary for the diagnosis*, staging and re-staging of non-Hodgkin's lymphoma when the general medical necessity criteria for oncologic indications (II. A. listed above) are met. *Note: A diagnostic tissue sample is usually obtainable without PET localization. Therefore, PET for diagnosis of non-Hodgkin's lymphoma is rarely considered medically necessary.	
United Healthcare (2018) [7]	• NR	PET (FDG) and/or combined PET-CT scans are covered when determined to be medically necessary and specific criteria are met. Positron Emission Tomography (PET) is covered for specific indications when coverage criteria are met. Lymphoma is a covered condition for initial treatment strategy and subsequent treatment strategy. United Healthcare based their decision on the most recently updated CMS NCD. Refer to the following National Coverage Determinations (NCDs) for specific coverage criteria	Rationale: NR
Regional Policies			
Blue Cross Blue Shield **Minnesota 2017** [6] **and**	NR	No national coverage policy was found for the BCBS network, two BCBS regional policies are represented below: **Minnesota 2017**	Rationale: NR

Payer (year)	Evidence Base Available	Policy	Rationale/Comments
Louisiana 2018[11]		Initial Treatment Strategy: PET or PET/CT may be considered MEDICALLY NECESSARY AND APPROPRIATE as an Initial Treatment Strategy (Diagnosis and Staging) for known or suspected malignancy (including Lymphomas) when the following criteria are met: • One (1) PET or PET/CT for solitary pulmonary nodule, myeloma, and all solid malignant tumors (except those listed below as experimental/investigative) when the test is needed to determine the location and/or extent of the suspected or proven malignancy in order to make at least one of the following determinations: 1. Whether or not the patient is an appropriate candidate for an invasive diagnostic or therapeutic procedure; OR 2. The optimal anatomic location for an invasive procedure; OR 3. The anatomic extent of malignancy, when recommended therapy reasonably depends on the extent of malignancy; AND • Other standard imaging modalities (e.g., CT, MRI or ultrasound) are either not indicated or are unable to conclusively provide the required information. Subsequent Treatment Strategy: PET or PET/CT may be considered **MEDICALLY NECESSARY AND APPROPRIATE** as a Subsequent Treatment Strategy (Restaging and Monitoring) for known or suspected malignancies (including Lymphomas) when the following criteria are met: • PET or PET/CT for myeloma and all solid primary malignant tumors (except those listed below as experimental/investigative) when the test is performed after completion of initial therapy for malignancy and the imaging results are required to assess therapeutic success, in order to establish the need for any subsequent therapy, by determining at least one of the following: 1. Presence or extent of residual disease; or 2. Presence or extent of recurrent disease; or 3. Presence or extent of metastasis; or 4. Other assessment of tumor response; AND • Other standard imaging modalities (e.g., CT, MRI, or ultrasound) are either not indicated or unable to conclusively provide the required information.	

Payer (year)	Evidence Base Available	Policy	Rationale/Comments
		Early treatment response assessment, also referred to as interim PET (i.e., involving comparison of PET images before treatment and at some interval during the initial course of treatment), may be considered MEDICALLY NECESSARY AND APPROPRIATE for patients with Hodgkin lymphoma after completion of at least 2 cycles of chemotherapy, when the result of interim PET is needed to guide treatment decisions. **Louisiana 2018** PET is considered **Eligible for coverage** for Diagnosis, Staging and Restaging of Lymphoma, including Hodgkin's Lymphoma Diagnosis: Only in clinical situations where PET results may assist in avoiding invasive diagnostic procedure or may assist in determining optimal anatomical location for such procedures. PET scans follow tissue diagnosis, and are performed for staging, and thus diagnostic use of PET is considered rare. Staging: Use in staging is eligible in the following situations • For staging lymphoma during initial staging; or • In clinical situations in which the stage of the cancer remains in doubt after completion of a standard diagnostic workup. Restaging: For restaging at follow-up; or • For the purpose of detecting residual disease, detecting suspected recurrence, or • to determine the extent of a known recurrence; or • For restaging after the completion of treatment. PET is considered **investigational** when used for applications not discussed above.	

ceCT = contrast enhanced Computerized Tomography; CMS = Centers for Medicare and Medicaid Service; CT: Computerized Tomography; FDG-PET = Fluorodeoxyglucose Positron Emission Tomography; HTA = Health Technology Assessment; NCCN = National Comprehensive Cancer Network; NCDs = National coverage determination; NR = not reported; PET = Positron Emission Tomography; PET-CT = Positron Emission Tomography/Computerized Tomography; SR = systematic review.

3. The Evidence

3.1. Methods of the Systematic Literature Review

3.1.1. Objectives

The aim of this report is to update the 2011 HTA on positron emission tomography (PET) for lymphoma by summarizing information on diagnostic accuracy (e.g., sensitivity, specificity, predictive values) for context and systematically reviewing, critically appraising and analyzing new research evidence evaluating the clinical effectiveness (i.e. the ability of PET to stage, and influence therapeutic decisions, clinical management and clinical outcomes), safety, differential efficacy and safety in subpopulations, and cost-effectiveness of PET for lymphoma in adult and pediatric patients. The combination of PET with diffusion weighted MRI is an emerging technology for the evaluation of lymphoma. Currently this combination is not widely used, so the focus of this report will be on PET/CT as it is the current standard of care. Evidence on PET/MRI will be included as appropriate.

3.1.2. Contextual Questions

In patients with histologically proven HL and NHL included in the report, what are the accuracy and reliability of ^{18}FDG-PET alone or in combination with CT for initial staging, interim evaluation (including monitoring during treatment, evaluation of treatment response and prognosis), re-staging (e.g. at end of treatment or other times), end of treatment assessment, evaluation of disease recurrence and surveillance of patients in remission? Specifically, provide a summary of the:

- Sensitivity and specificity and prognostic value (positive and negative predictive values)
- Inter- and intra-rater reliability (reproducibility)

In addition, a brief summary of the diagnostic accuracy and use of PET/CT for initial diagnosis will be provided. Summaries of accuracy will be based on highest quality systematic reviews which critically appraise included studies. Contextual information on the combination of PET with MRI will be presented. Contextual questions are not systematically reviewed, and use a “best evidence” approach.

3.1.3. Key Questions

The focus of this portion of the report is on the clinical impact of ^{18}FDG PET/CT as this is the current standard of care. Except where noted the term PET will be construed to mean PET/CT. Information related to the PET/MRI combination will be included if relevant. In patients with histologically proven HL or NHL undergoing PET/CT for initial staging, interim evaluation (including monitoring during treatment, evaluation of treatment response and prognosis), re-staging (e.g., at end of treatment or other times), end of treatment assessment, evaluation of disease recurrence and surveillance of patients in remission:

1. What is the evidence of clinical effectiveness of ^{18}FDG imaging in combination with CT (PET/CT) results?
    a. How do ^{18}FDG PET/CT results impact therapeutic decisions or clinical management? Do test results lead to use of effective treatment strategies (e.g. including initial treatment following staging or treatment acceleration, deceleration or termination at interim imaging) compared with treatment strategies not using such test results?
    b. How do clinical outcomes (e.g. overall survival) differ based on PET/CT-related treatment decisions compared with decisions made in the absence of such test results?

c. Does the use of ^{18}FDG PET for treatment decisions lead to reduction in treatment-related adverse events/sequelae in general compared with treatment decisions that do not involve PET/CT?
d. Is there a reduction in the need for other tests?
e. How do end of treatment ^{18}FDG PET/CT results impact clinical decision making compared with clinical decisions that do not involve PET/CT??
f. How does surveillance ^{18}FDG PET/CT results impact clinical decision making compared with clinical decisions that do not involve PET/CT??

2. What is the safety profile of ^{18}FDG PET/CT for lymphoma?
    a. What adverse events are reported: type and frequency directly attributable to ^{18}PET/CT (mortality, major morbidity, radiation exposure, other)?

3. What is the evidence that ^{18}FDG PET/CT imaging in patients with known lymphoma has differential efficacy or safety in subpopulations? Including consideration of:
    a. Patient age, sex, characteristics or evidence-based patient selection criteria
    b. Type of scanning machine and software, reader training, and other operational factors
    c. Provider type, setting or other provider characteristics
    d. Health care system type, including worker's compensation, Medicaid, state employees

4. What is the evidence of short and long-term cost-effectiveness of ^{18}FDG PE/CT for patients with lymphoma compared with other imaging or clinical care not involving ^{18}FDG PET/CT?

3.1.4. Inclusion/exclusion criteria

Inclusion and exclusion criteria are summarized in Table 8. Briefly, included studies met the following requirements with respect to participants, intervention, comparators, outcomes, and study design:

- **Population**: Adults, adolescents or children with biopsy-proven HL or NHL. Diagnoses of interest (to be evaluated separately) include HL or aggressive NHL, indolent NHL and other NHL with focus on most common types.
- **Interventions:** Focus is on positron emission tomography (PET) to measure glucose metabolism (^{18}FDG-PET) in addition to computed tomography (CT) (including combined PET/CT equipment). Combination of PET with MRI (including diffusion weighted MRI) will be included as appropriate.
- **Comparators:** Other imaging (CT alone, MRI, including diffusion weighted MRI); standard clinical protocols or standard prior tests/evaluations (including history and physical examination, laboratory studies, biopsy) that do not involve ^{18}FDG PET.
- **Outcomes:** Primary outcomes are 1) improvement in clinical outcomes based on PET/CT-derived clinical decision making (focus is on overall survival, progression- or event-free survival, morbidity and mortality, and 2) quality of life. (See Table 8 below for secondary and indirect outcomes). Safety outcomes are type and frequency of adverse events directly attributable to PET/CT (e.g., incidental findings, repeat/additional procedures, radiation exposure, other). Economic outcomes are cost-effectiveness (e.g., cost per improved outcome), cost-utility (e.g., cost per quality adjusted life year (QALY), incremental cost effectiveness ratio (ICER) outcomes.
- **Timing:** PET performed at any time after initial diagnosis including for interim evaluation (including monitoring during treatment, evaluation of treatment response and prognosis), re-

staging (e.g., at end of treatment or other times), end of treatment assessment, evaluation of disease recurrence and surveillance of patients in remission.

- **Studies:** For efficacy and effectiveness, the focus was on comparative PET/CT studies with the least potential for bias (i.e. RCTs). For safety, the focus will be on studies characterizing direct PET/CT harms (including incidental findings, repeat biopsy, radiation safety). For differential efficacy or effectiveness, only studies which stratify on patient or other characteristics and formally evaluate statistical interaction (effect modification) were considered. For cost-effectiveness, only full, formal economic studies (i.e., cost-effectiveness, cost-utility, cost-minimization, and cost-benefit studies) were included. Given that this is re-review, only studies published subsequent to the 2011 report (Search dates February 2011 to May 2018) were considered.

Table 8. Summary of inclusion and exclusion criteria

Study Component	Inclusion	Exclusion
Population	Adults, adolescents or children with biopsy-proven HL or NHL. Diagnoses of interest include (to be evaluated separately): • Hodgkin lymphoma (HL) or aggressive non-Hodgkin lymphoma (NHL) • Indolent NHL • Other NHL with focus on most common types	• Studies with <80% of population with HL or NHL • Studies with <80% of population who got PET
Interventions	• **FOCUS:** Positron emission tomography (PET) to measure glucose metabolism (^{18}FDG-PET) in addition to computed tomography (CT) (including combined PET/CT equipment) • Combination of PET with MRI (including diffusion weighted MRI)	• 3'-deoxy-3'-^{18}F-fluorothymidine (i.e., ^{18}F-FLT-PET) or other uncommonly used tracers for lymphoma; Investigational or non-FDA approved tracers • Outdated PET technology or methods; gamma PET, CDET • Studies of PET alone (i.e. in the absence of either separate CT or done on combination PET/CT equipment.
Comparator	• Other imaging (CT alone, MRI, including diffusion weighted MRI) • Standard clinical protocols or standard prior tests/evaluations (including history and physical examination, laboratory studies, biopsy) that do not involve ^{18}FDG PET	• Indirect comparisons of imaging methods or protocols
Outcomes	• **Primary outcomes:** Improvement in clinical outcomes based on PET/CT –based clinical decision making (focus on overall survival, event-free or progression-free survival, morbidity and mortality), quality of life • **Secondary outcomes:** Summary of adverse treatment effects based on PET-related treatment decisions • **Indirect outcomes:** Documentation of impact on therapeutic decisions or clinical	• Technical efficacy (i.e., the ability of a diagnostic test to conform to technical specifications) • Impact on diagnosis, therapeutic decisions, and clinical outcomes of patients with diagnosis other than Hodgkin or non-Hodgkin lymphoma. Studies that focus on specific PET features

Study Component	Inclusion	Exclusion
	management (e.g., reduced need for other tests, change in patients' management [e.g., continuation or discontinuation of therapy], change in treatments planned or given, change in stage) • **Safety:** Adverse events directly attributable to PET/CT; type and frequency (e.g. incidental findings, repeat/additional procedures, radiation exposure, other) • **Economic:** Cost-effectiveness outcomes (e.g., cost per improved outcome) or cost-utility outcomes (e.g., cost per quality adjusted life year (QALY), incremental cost effectiveness ratio (ICER) or similar)	
Timing	• Initial staging, interim evaluation (including monitoring during treatment, evaluation of treatment response and prognosis), re-staging (e.g., at end of treatment or other times), end of treatment assessment, evaluation of disease recurrence and surveillance of patients in remission	• Initial diagnosis
Study Design	Focus will be on PET/CT studies with the least potential for bias. • KQ 1-2: Effectiveness: o Updated, high quality systematic reviews that include research published subsequent to the prior report will be considered. o RCTs and prospective, longitudinal observational studies will be sought that compare treatment strategies based on PET/CT results with strategies that do not involve PET/CT results are of primary interest. In the absence of such studies, studies comparing treatments given based on PET findings will be included with a focus on RCTs which randomize patients to treatment based on PET/CT findings.* o Treatment and response planning and clinical decision making studies must provide specifics regarding use of PET/CT results to inform treatment assessment or therapy planning; preference will be given to prospective comparative studies. Safety: Studies characterizing direct PET/CT harms (including incidental findings, repeat	• Studies of diagnostic accuracy, including prognosis (covered in the contextual questions) • Technical papers (e.g. SUVs, FDG uptake, phantom studies, quantitation papers) • Studies solely evaluating bone marrow involvement seen on PET (e.g. evaluation of bone marrow only, use of PET in lieu of bone marrow biopsy or comparing PET with bone marrow biopsy) • Indirect comparisons of imaging modalities or treatment strategies • Incomplete economic evaluations such as costing studies • Studies with fewer than 30 patients for HL and DBCL and fewer than 15 patients for more rare lymphoma • Case reports • Studies whose abstracts do not allow study characteristics to be determined. • SUV studies that only evaluate prognosis, not treatment response or planning with a view to evaluating patient management

Study Component	Inclusion	Exclusion
	biopsy, radiation safety); Included studies which describe the impact of PET/CT–related decision making on treatment-related adverse events. • KQ 3: Studies which stratify on patient or other characteristics and formally evaluate statistical interaction (effect modification) • KQ 4: Only full, formal economic studies (i.e., cost-effectiveness, cost-utility, cost-minimization, and cost-benefit studies) comparing PET/CT with other imaging or clinical care not involving FDG PET/CT will be considered.	
Publication	• Studies published in English in peer reviewed journals, technology assessments or publically available FDA reports • Studies published subsequent to the 2011 report (Search dates February 2011 to May 2018) • For question 5, full formal economic analyses (e.g., cost-effectiveness, cost-utility studies) published in English in a peer reviewed journal	• Abstracts, conference proceedings, posters, editorials, letters • Duplicate publications of the same study that do not report different outcomes or follow-up times • Single reports from multicenter trials • White papers • Narrative reviews • Articles identified as preliminary reports when full results are published in later versions • Incomplete economic evaluations such as costing studies

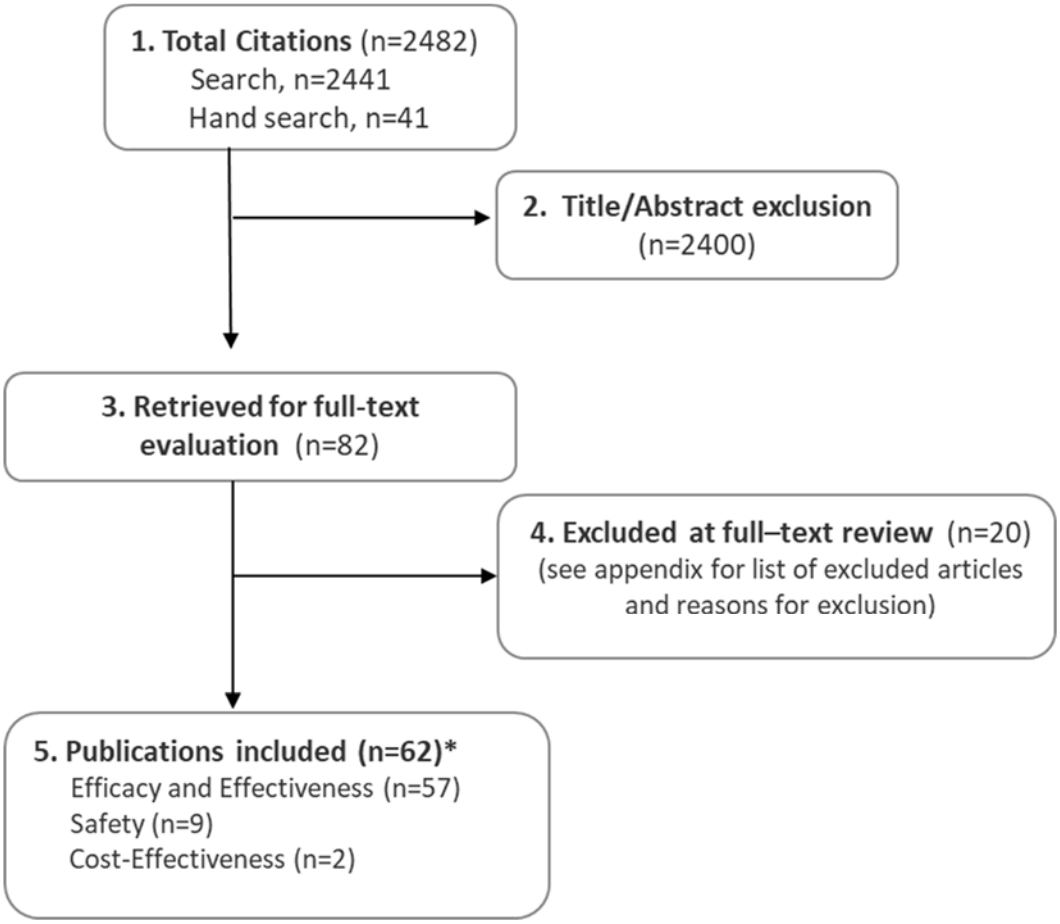
*In the absence of such studies, contextual information on treatments and outcomes in untested patients will be described.

3.1.5. Data sources and search strategy

We searched electronic databases from March 2011 to May 26, 2018 to identify publications assessing the use of PET for the treatment of lymphoma that had been published since the original report. The 2011 report used a rapid review methodology of selected literature sources and focused on the diagnostic accuracy of PET reported in previous health technology assessments and systematic reviews. Subsequent to the 2011 report, more recent evidence has evaluated the role of PET/CT for clinical decision making and adapting treatment based on PET findings. Thus, the focus of this re-review is on studies evaluating the clinical impact of PET/CT on the primary clinical outcomes of overall survival, progression-free survival, and treatment-related toxicity necessitating a broader search methodology. A formal, structured systematic search of the peer-reviewed literature was performed across a number of databases including PubMed, EMBASE, the Cochrane Central Register of Controlled Trials, the Cochrane Database of Systematic Reviews, and the National Guideline Clearinghouse (see Appendix B for full search strategy) to identify relevant peer reviewed literature as well as other sources (National Guideline Clearinghouse, Center for Reviews and Dissemination Database) to identify pertinent clinical guidelines and previously performed assessments. We also hand searched the reference lists of relevant studies and the bibliographies of systematic reviews.

The clinical studies included in this report were identified using the algorithm shown in Appendix A. The search took place in four stages. The first stage of the study selection process consisted of the comprehensive electronic search and bibliography check. We then screened all possible relevant articles using titles and abstracts in stage two. This was done by two individuals independently. Those articles that met a set of *a priori* retrieval criteria were included. Articles were selected for full-text review if they included a comparison of an intervention and a control of interest for the treatment of chronic migraine, chronic tension-type headache, or chronic daily headache. We excluded conference abstracts, non-English-language articles, and studies of nonhuman subjects. Any disagreement between screeners that were unresolved resulted in the article being included for the next stage. Stage three involved retrieval of the full text articles remaining. The final stage of the study selection algorithm consisted of the selection of those studies using a set of *a priori* inclusion criteria, again, by two independent investigators. Discrepancies were resolved through discussion and if necessary adjudicated by a third investigator. A list of excluded articles along with the reason for exclusion is available in Appendix C. The remaining articles form the evidence base for this report, Figure 2.

Figure 2. Flow chart of literature search results



PET = positron emission tomography
*A publication may contribute data to both efficacy/effectiveness and safety. Further details provided in Section 4.2.

3.1.6. Data extraction

Reviewers extracted the following data from the clinical studies: study design, study period, setting, country, sample size, inclusion and exclusion criteria, study population characteristics, follow-up time, timing of PET assessment (e.g., initial staging, interim), criteria used to define PET positivity or negativity, study outcomes and adverse events. Information on how PET was used to make treatment modifications was also abstracted. For economic studies, data related to sources used, economic parameters and perspectives, results, and sensitivity analyses were abstracted. An attempt was made to reconcile conflicting information among multiple reports presenting the same data. Detailed study and patient characteristics and results are available in Appendices F–I.

3.1.7. Quality assessment: Overall Strength of evidence (SoE), Risk of Bias, and QHES evaluation

The method used by Aggregate Analytics, Inc. (AAI) for assessing the quality of evidence of individual studies as well as the overall strength of evidence (SoE) for each primary outcome from RCTs are based on criteria and methods established in the *Cochrane Handbook for Systematic Reviews of Interventions*,[89] precepts outlined by the Grades of Recommendation Assessment, Development and Evaluation (GRADE) Working Group,[28] and recommendations made by the Agency for Healthcare Research and Quality (AHRQ).[1,158] Economic studies were evaluated according to The Quality of Health Economic Studies (QHES) instrument developed by Ofman et al.[113] Based on these quality criteria, each study chosen for inclusion for a Key Question was given a RoB (or QHES) rating; details of each rating are available in Appendix E. Standardized, pre-defined abstraction guidelines were used to determine the RoB (or QHES) rating for each study included in this assessment.

The SoE for all primary health outcomes was assessed by two researchers following the principles for adapting GRADE (Grades of Recommendation Assessment, Development and Evaluation)[32,84,85] as outlined by the Agency for Healthcare Research and Quality (AHRQ).[1] The strength of evidence was based on the highest quality evidence available for a given outcome. In determining the strength of body of evidence regarding a given outcome, the following domains were considered:

- Risk of bias: the extent to which the included studies have protection against bias
- Consistency: the degree to which the included studies report results that are similar in terms of effect sizes, range and variability.
- Directness: describes whether the evidence is directly related to patient health outcomes or comparisons of interventions are direct (head to head).
- Precision: describes the level of certainty surrounding the effect estimates.
- Publication bias: is considered when there is concern of selective publishing or selective reporting.

When assessing the SoE for studies performing subgroup analysis, we also considered whether the subgroup analysis was preplanned *(a priori)* and whether a test for homogeneity or interaction was done.

Bodies of evidence consisting of RCTs were initially considered as High strength of evidence, while those comprised of nonrandomized studies began as Low strength of evidence. None of the included randomized controlled trials for this report randomized patients to receive PET/CT versus a strategy that did not include use of PET/CT, but rather, patients were randomized to various treatment options based

on results of their PET/CT. Consistent with a recent Cochrane Review, we assessed them as RCTs.[136] The strength of evidence could be downgraded based on the limitations described above. There are also situations where the observational studies could be upgraded if the study had large magnitude of effect (strength of association) or if a dose-response relationship is identified if there are no downgrades for the primary domains listed above and confounding is not a concern. Publication and reporting bias are difficult to assess, particularly with fewer than 10 RCTs.[1] Publication bias was unknown in all studies and thus this domain was eliminated from the strength of evidence tables. The final strength of evidence was assigned an overall grade of high, moderate, low, or insufficient, which are defined as follows:

- High - Very confident that effect size estimates lie close to the true effect for this outcome; there are few or no deficiencies in the body of evidence; we believe the findings are stable.
- Moderate – Moderately confident that effect size estimates lie close to the true effect for this outcome; some deficiencies in the body of evidence; we believe the findings are likely to be stable but some doubt remains.
- Low – Limited confidence that effect size estimates lie close to the true effect for this outcome; major or numerous deficiencies in the body of evidence; we believe that additional evidence is needed before concluding that findings are stable or that the estimate is close to the true effect.
- Insufficient – We have no evidence, are unable to estimate an effect or have no confidence in the effect estimate for this outcome; OR no available evidence or the body of evidence has unacceptable efficiencies precluding judgment.

Similar methods for determining the overall quality (strength) of evidence related to economic studies have not been reported, thus the overall strength of evidence for outcomes reported in Key Question 4 was not assessed.

Primary outcomes for this report were OS and PFS and treatment related toxicity; Strength of Evidence (SOE) was graded for these primary outcomes only. The results focus on the highest quality of evidence available. Where RCTs or higher quality evidence were available, these were used to assess the overall strength of evidence.

3.1.8. Analysis

We summarized evidence separately for initial staging, interim assessment, end-of-treatment assessment and surveillance, and within those categories, for Hodgkin lymphoma (HL) and non-Hodgkin lymphoma (NHL). HL was further divided into early-stage, advanced-stage, and relapsed/refractory disease, and NHL into aggressive (aNHL) and indolent (iNHL) subtypes. Some populations were comprised of mixed lymphoma types and results were not reported separately for the different subtypes; these are referred to in the report as “mixed” populations. The intervention of interest was positron emission tomography (PET) to measure glucose metabolism (^{18}FDG-PET) in addition to computed tomography (CT) (including combined PET/CT equipment). Except where noted the term PET will be construed to mean PET/CT. PET alone or using other tracers was excluded as an intervention for evaluation of primary outcomes. Combination of PET with MRI (including diffusion weighted MRI), though not the focus of this report, was included as appropriate with the contextual question.

For purposes of this report, the term "PET-adapted" refers to studies that explicitly describe what specific management changes or treatments were done or not done based on PET results, the number of patients impacted by the PET-based treatment decision(s) and describe the prognosis in terms of OS and PFS for those patients. Studies of PET-adapted treatment evaluated interim or end of treatment PET primarily. Studies that provided general information on PET's impact on staging (e.g. upstaging or down staging) or report non-specific treatment changes are briefly abstracted and summarized. Such outcomes were considered surrogates for treatment and therefore secondary and intermediate outcomes. Identified studies that were not PET-adapted were either prospective or retrospective cohort studies. Overall, these studies were considered at to be at moderately high or high risk of bias. Detailed risk of bias was not assessed for these studies; however selection bias and performance bias were generally noted across them during abstraction. They were not included in determination of overall strength of evidence.

The value of a medical test relates to its ability identify persons for whom there are appropriate and effective treatments. In the absence of studies randomizing patients to receive PET/CT or no PET/CT and following patients through treatments to final clinical outcomes, the impact of PET on clinical decision making leading to effective treatment and outcomes can only be assessed indirectly. The outcomes observed are a function of PET/CT results in combination with the treatments received on the basis of those results. Thus, our approach considers the clinical outcomes of OS and PFS in relation to the modifications of treatment received based on PET results, along with treatment related-toxicity. In other words, for this report the value of PET for assisting with decisions to escalate, de-escalate or continue therapy is indirectly assessed based upon the impact of the PET-adapted treatment.

Due to heterogeneity across studies with regard to designs, patient populations, treatments and clinical methods meta-analysis was not performed.

There is heterogeneity across included studies and guidelines with regard to how the timing of PET is defined and classified. For the purposes of this report the 2014 Consensus statement of the International Conference on Malignant Lymphomas Imaging Working Group[34] is used wherein 'interim PET' (iPET) is conceptualized as early (mid-treatment) assessment during the course of first-line treatment, and end of treatment PET (EoT-PET) is done for remission assessment (or response to treatment) after that initial induction therapy and prior to potential salvage therapy or consolidation therapy. For purposes of this report, PET/CT evaluation of patients in remission done in asymptomatic patients will be referred to as surveillance while PET/CT done if there is clinical suspicion of recurrence or relapse will be noted as such.

For the contextual question, data on diagnostic accuracy were summarized based on systematic reviews published subsequent to 2011 report to the extent possible. For completeness, data from primary studies were reported for lymphoma types included in the report if systematic review information was not available. These primary studies were identified by searching the EndNote library containing search results for our systematic review. For some combinations of lymphoma type and PET use, neither primary studies nor systematic reviews were available. To provide a general context for quality of included systematic reviews, a modified AMSTAR tool was used and for single studies of diagnostic accuracy and reliability, a modification of the QUADAS tool was used (Appendix D). Full formal evaluation of risk of bias was not done and overall strength of evidence was not assessed.

Outcomes are detailed in the evidence tables in the appendices and/or the body of the report.

4. Results

4.1. Contextual Questions

In the past decade ^{18}F-FDG PET/CT has become an integral part of clinical management of lymphoma. In evaluating evidence on the diagnostic accuracy (including estimation of prognosis) presented below and evidence presented in response to the key research questions, a number of factors should be kept in mind as they likely impact the values of these measures and interpretation of the studies. Such factors include heterogeneity in criteria for PET/CT interpretation across (and sometimes within) studies, the spatial resolution of PET and the frequency of false-positive results.

First, over the past decade, the technology and its application have evolved, including the methods for incorporating PET/CT information into initial staging, criteria for determining what constitutes a positive versus negative result and methods of evaluating tumor response to therapy.[49,99] For determination of PET positivity and evaluation of therapy response, there has been an evolution from the use of more visual criteria (e.g. IHP criteria) to more reproducible criteria such as the 5-Point Scale (formerly known as the Deauville Criteria) and more recently to more quantitative methods that evaluate standardized uptake value (SUV) measurements and changes in them from baseline.

While PET/CT has been considered a surrogate marker to assess therapy response and guide therapy, several limitations of PET need to be considered.[23,49,99] The spatial resolution of whole-body PET images may be 6-9 mm, thus small residual tumor deposits may be missed creating a false-negative result and not exclude residual disease.[23] False-positive results are not uncommonly reported during or after treatment and may vary with the type of lymphoma, the stage, type of therapy and timing of PET after the last cycle of therapy. Some meta-analyses[21] and narrative reviews[24,99] suggest that the false positive rate is high and due to inflammatory processes, immune response from treatment or tissue necrosis. As a consequence the predictive value for treatment failure may be over-estimated. Follow-up of false positives may include unnecessary procedures (e.g. biopsy), additional testing and/or additional therapies including those involving ionizing radiation and create anxiety in the patient.

With regard to study design many of the studies below, the reference standard for determining PET/CT accuracy was generally not a gold-standard such as histology.

Summary

Data are summarized below (Tables 9-19) and in the executive summary.

Terms and definitions

Diagnostic test performance measures were used in this report. Briefly for the contextual questions:

- **True positive (TP):** patients who test positive and who have the disease
- **False positive (FP):** patients who test positive but do not have the disease
- **False negative (FN):** patients who test negative but do have the disease

- **True negative (TN):** patients who test negative and do not have the disease
- **Sensitivity** (true positive rate) measures the ability of a diagnostic test to correctly identify patients with the disease, and can be calculated as follows: TP/(TP + FN)
- **Specificity** (true negative rate) measures the ability of a diagnostic test to minimize false positives, and can be calculated as follows: TN/ (FP + TN)
- Note that sensitivity and specificity are negatively correlated with one another: if the threshold of a positive test is set higher to maximize sensitivity, specificity will be lower, and vice versa.
- **Kappa statistic:** represents the extent to which the agreement observed exceeds that which would be expected by chance alone, and can be calculated as follows:

$$\text{Kappa} = \frac{(\%\ \text{agreement observed}) - (\%\ \text{agreement expected by chance alone})}{100\% - (\%\ \text{agreement expected by chance alone})}$$

- **Negative Predictive Value (NPV):** % of patients with negative test who do not have the disease; depends on prevalence can be calculated one of two ways: TN/(TN+FN) or

$$\text{NPV} = \frac{\text{specificity} \times (1\text{-prevalence})}{(1\text{-sensitivity}) \times \text{prevalence} + \text{sp1ecificity} \times (1\text{-prevalence})}$$

- **Positive Predictive Value (PPV):** % of patients with positive test who have the disease; depends on prevalence; can be calculated one of two ways: TP/(TP+FP) or

$$\text{PPV} = \frac{\text{sensitivity} \times \text{prevalence}}{\text{sensitivity} \times \text{prevalence} + (1\text{-specificity}) \times (1\text{-prevalence})}$$

- **Positive Likelihood Ratio (LR+):** how much odds of disease increase with a positive test; depends on sensitivity and specificity; can be calculated as follows: sensitivity/(1-specificity)
- **Negative Likelihood Ratio (LR-):** how much odds of disease decrease with a negative test; depends on sensitivity and specificity; can be calculated as follows: (1-sensitivity)/specificity)
- **Area Under the Curve (AUC):** area under the receiver operating characteristic (ROC) curve created by plotting sensitivity on the y axis and 1-specificity on the x axis; ranges from 0 to 1 where 0 represents a perfectly inaccurate test and 1 represents a perfectly accurate test
- **Diagnostic Odds Ratio (DOR):** ratio of odds of a positive test when disease is present relative to odds of a positive test if disease is not present; can be calculated as follows: (TP/FP)/(FN/TN)

Results for Diagnostic Accuracy

Overall, 22 SRs[13-15,17-20,22,52,54,70,103,117,125,141,153,156,161,164-166,170] and 8 primary studies[30,31,40,43,45,82,90,120] are summarized to provide context regarding the diagnostic accuracy and reliability of use PET/CT in patients. The focus is on the most common types of lymphoma and uses of PET represented by studies included to answer the full research Key Questions. Not all types of lymphoma or uses of PET/CT may be represented below. Across studies, there is substantial heterogeneity with regard study designs, PET/CT criteria and thresholds for positivity, patient populations and outcomes reported. Tables 9 through 19 provide additional detailed data. Data are organized by PET/CT use (timing) and by lymphoma type. Data from more recent studies as well as 8 SRs/HTAs from the prior review[53,68,97,100,146-148,167] are summarized

in the table. Risk of bias for primary studies ranged from moderately high to moderately low. Main concerns included lack of blinded comparisons, inadequate description of methods, and small sample size.

Tables 9-19 provides detailed results for diagnostic accuracy. Results summarized below focus on information available subsequent to the prior report unless otherwise noted. The following describes primary results relating to PET/CT:

Diagnosis

Two systematic reviews describing diagnostic accuracy were identified and included patients with primary CNS NHL[161,170] (most of which are diffuse large B-cell lymphomas (DLBCL). Sensitivity and specificity were high across studies of immunocompetent persons (pooled sensitivity 88% (95% CI 80% to 94% and specificity 86% (95% CI 73% to 94%, and AUC 0.9192) and in HIV patients (100% sensitivity, 75% to 100% specificity).

Initial staging

- In classic HL, two studies suggest that PET/CT is accurate for initial staging (sensitivity range 97.9% to 100%, specificity range 95.3% to 99%)[43,82]
- In aggressive NHL (extranodal natural killer/T-cell lymphoma) one SR including patient-level data from 6 studies found that PET/CT is accurate for initial staging with regard to sensitivity, but specificity was poor.(pooled sensitivity 95%, pooled specificity 40%)
- In indolent NHL, two systematic reviews provide limited information on the accuracy of PET for initial staging. Across 7 studies, one review evaluating the additional value of FDG-PET to CT, reported a pooled estimate of the proportion of patients restaged as a result of the addition of PET of 18.7% (95% CI 10.8% to 30.4%), but no data on diagnostic accuracy. A 2016 Australian MSAC review concluded that PET/CT is just as specific but more sensitive than CT for relapse and that PET/CT identified more lesions than bone marrow biopsy for new diagnosis (but true positive rates ranged from 15% to 100% for additional lesions) One study in patients with Marginal zone B cell lymphoma (MZL) suggests that PET/CT has high sensitivity in newly diagnosed patients.
- In mixed populations (NHL and HL), four SRs included in the previous HTA report conflicting results. Three SRs suggest high diagnostic accuracy for PET/CT (sensitivity range 71.4% to 100%, specificity range 86.2% to 100%) while one SR suggests a high rate of false negatives (ratio of true positive to true negative sites detected by PET was 3:1)

Interim PET

- In classic HL, one SR included in the previous HTA suggests that PET/CT outperforms both CT and FDG-PET alone (PET/CT sensitivity 100%, FDG-PET sensitivity range 85% to 100%, PET/CT specificity 90.7%, FDG-PET specificity range 57.1% to 100%)
- In aggressive NHL, one SR included in the previous HTA reported only on FDG-PET alone. Sensitivity ranged from 60% to 100% and specificity ranged from 80% to 100%

- In indolent NHL, one SR including one study reported that PET/CT has a higher specificity but lower sensitivity compared to CT (sensitivity 72%, specificity 100%)
- In mixed populations (NHL and HL), one SR including 11 studies reported that PET/CT scans result in a high number of false positive lesions (pooled FP rate for NHL 83%). Authors conclude that the role of interim PET/CT should be reconsidered

End of treatment PET

- In mixed populations (NHL and HL), one SR including 11 studies reported that PET/CT scans result in a high number of false positive lesions (pooled FP rate for NHL 31.5%, pooled FP rate for HL 23.1%). Authors conclude that the role of end-of-treatment PET/CT should be reconsidered

Surveillance with PET

- In classic HL, one study suggests that, PET/CT has lower specificity (86.3% vs. 96.3%) and PPV (72.7% vs. 90.7%) compared to US/chest radiography. Sensitivity and NPV were similar for both PET/CT and US/chest radiography (sensitivity 100% vs 97.5%, NPV 100% vs. 99.1%). Authors conclude that US/chest radiography should be used instead of PET/CT for routine surveillance imaging.
- In aggressive NHL, two studies reported discordant results. One study in patients with DLBCL suggests that PET/CT has a higher FP rate compared to CT (13.7% vs. 1.7%) and that PET/CT has limited value (sensitivity 100% vs 100%, specificity 86.1% vs. 98.3%, PPV: 11.5% vs. 30%, and NPV 100% vs. 100%). A second study among patients with transformed indolent NHL found that PET/CT had high sensitivity (83%) and specificity (94%), as well as a low FP rate (2% including indeterminate, 5% excluding indeterminate). Authors concluded that PET/CT surveillance for patients with transformed indolent NHL after achieving complete remission is not indicated.
- No studies reported on the clinical utility of PET/CT for post-treatment surveillance in indolent NHL
- In mixed populations (NHL and HL), two SRs reported a lack of evidence to support PET/CT in post-treatment surveillance (sensitivity range 75% to 100%, specificity range 43% to 92%, NPV range 98% to 100%, PPV range 20% to 100%)

Prognosis

- In classic HL, one SR on interim PET/CT suggests that diagnostic performance of interim PET/CT is limited (pooled sensitivity 67%, pooled specificity 89%, pooled NPV 93%), one SR on pre-SCT PET/CT suggests that patients with positive pre-transplant PET/CT scans have worse PFS and OS compared to patients with negative pre-transplant scans (PFS sensitivity 67.2%, OS sensitivity 74.4%, PFS specificity 70.7%, OS specificity 58.0%), and one SR on end-of-treatment PET/CT suggests that PET/CT-negative residual masses are not proven to be associated with worse outcomes compared to PET/CT-based complete remission without a residual mass (pooled disease relapse proportion with PET/CT-negative residual mass 6.8%)

- In aggressive NHL, four SRs on interim PET/CT reported conflicting results. In patients with DLBCL treated with R-CHOP, two SRs suggest suboptimal prognostic value of interim PET/CT (sensitivity range 21.2% to 89.7%, specificity range 27.4% to 91.5%). In contrast, two SRs including various types of aggressive NHL suggest that interim PET/CT can accurately assess disease prognosis (PFS HR range comparing PET/CT positive vs. PET/CT negative 2.93 to 4.4, PFS HR range comparing PET/CT positive vs. PET/CT negative 2.55 to 3.99). One SR on pre-SCT PET/CT reported that available evidence is of poor quality and diagnostic performance is low (PFS sensitivity 54.0%, PFS specificity 73.1%, OS sensitivity 54.5%, OS specificity 68.7%). For end-of-treatment PET/CT one SR suggests high prognostic value (PFS HR comparing PET/CT positive vs. PET/CT negative 4.05, OS HR comparing PET/CT positive vs. PET/CT negative 5.91) while a second SR reports poor prognostic value (5-year OS among patients with complete remission according to PET/CT 83%)
- In indolent NHL, one SR on interim PET/CT reported that evidence does not support the use of interim PET/CT in follicular lymphoma (no significant difference in PFS or OS between PET/CT positive and negative patients). Three SRs on end-of-treatment PET/CT suggest that PET/CT is predictive of both PFS and OS (PFS HR comparing PET/CT positive vs. PET/CT negative range: 3.9 to 5.1, OS HR comparing PET/CT positive vs. PET/CT negative 6.7%, PPV range 89% to 94%).
- In mixed populations (NHL and HL), one SR on pre- and post-SCT PET/CT suggests that PET/CT is a useful prognostic tool (Pre-SCT PFS HR comparing PET/CT positive vs. PET/CT negative 2.32, Post-SCT PFS HR comparing PET/CT positive vs. PET/CT negative 4.61, OS HR comparing PET/CT positive vs. PET/CT negative 2.64)

Pediatric lymphoma

- In pediatric classic HL, one study reported initial staging with PET/CT identifies more disease sites than CT alone (219 vs. 152), one study on the use of PET/CT for monitoring during treatment suggests that evidence for treatment escalation based on interim PET/CT is lacking (sensitivity range 0% to 25%, specificity range 61.5% to 91.4%, PPV range 0% to 4.7%, NPV 91.4%), and one study on end-of-treatment PET/CT suggests better some predictive value using the Deauville interpretation method (sensitivity 25%, specificity 95.7%), PPV 33.3%, NPV 94.1%)
- In pediatric aggressive HL, one study reported initial staging with PET/CT may be better than CT alone (number of sites detected 130 vs. 114), one study on the use of PET/CT for monitoring during treatment suggests that interim PET/CT findings cannot be used to prognosticate for PFS nor OS (sensitivity 44.9%, specificity 41.1%, PPV 41.2%, NPV 85.7%), and one study on end-of-treatment PET/CT suggests that post-treatment PET/CT can be used to prognosticate for PFS but not OS (sensitivity 75%, specificity 75.0%, PPV 33.3%, NPV 94.7%)
- In mixed populations (NHL and HL), one SR reported high sensitivity and specificity for initial staging with PET/CT (sensitivity range 96% to 99%, specificity range 95% to 100%) and suggests that more research is needed to determine optimal timing for interim PET/CT (sensitivity range 77.8% to 100%, specificity range 54.5% to 97.7%)

Three SRs evaluating MRI were identified. One SR (Fitzpatrick)[70] on interim PET included 8 studies comparing diffusion- weighted MRI (DW-MRI) to PET/CT as a reference standard. Patients (n = 221) in the included studies had either NHL or HL. Authors concluded that hat DW-MRI has good agreement with PET/CT for assessing treatment response in lymphoma (pooled Kappa based on 5 studies 0.88). Pooled sensitivity for DW-MRI was 94.7% and pooled specificity was 99.3%. Two SRs compared whole-body diffusion-weighted MRI (WB-MRI) to PET/CT for staging of both HL and NHL. Using histology as the reference standard, the first SR (Wang)[156] reported a pooled staging accuracy of 98% for both WB-MRI and PET/CT. For indolent NHL, the pooled staging accuracy was 96% for WB-MRI and 97% for PET/CT. Agreement between WB-MRI and PET/CT was excellent (range 0.82 to 0.95 for nodal involvement, range 0.83 to 1.00 for extra-nodal involvement). A second SR reported agreement between WB-MRI and PET/CT in 90.5% of cases (Kappa = 0.871). Sensitivity ranged from 59% to 100% for WB-MRI and from 63% to 100% for PET/CT. Authors reported inability to calculate specificity due to all patients having lymphoma.

Bone marrow involvement:

- Five systematic reviews (one included in the 2011 report) assessed the diagnostic value of FDG PET and FDG PET/CT for detecting bone marrow involvement in patients with lymphoma (Table 19). One review (Cheng 2013)[47] focused on patients with HL, two (Adams 2014; Chen 2011)[16,46] only reported results for patients with NHL, and two included a mix of patients with HL and NHL (Wu 2012; Pakos 2005).[114,160]
- Pooled sensitivities ranged from 46% (indolent NHL) to 94.5% (HL) and pooled specificities ranged from 87.3% (mixed HL and NHL) to 99.5% (HL). With the exception of one study, which reported low sensitivity for patients with indolent NHL, all authors concluded that FDG PET/CT and FDG-PET have high sensitivity for detecting bone marrow involvement in patients with lymphoma.

Results for Reliability

Summary

Eight studies evaluating the reliability of PET/CT for various types of lymphoma are summarized below. In general, four were considered good quality,[36,76,88,98] three were of moderate quality,[33,74,75] and one[152] was considered poor quality. Primary factors that led to determination of poor quality were lack of independent review, small sample size, and lack of adequate description of methods for replication.

Across studies and timing of PET/CT, interrater reliability in adult populations ranged from moderate to substantial based on the Landis and Koch criteria when the 5-Point Scale (5-PS) criteria were used. Interrater reliability may vary based on what threshold for positivity of the 5-PS is used and reader experience with oncology.[135] Results across three studies in pediatric patients suggest that interrater reliability ranged from fair to substantial, with one study reporting that concordance was lowest for middle values of the 5 PS (categories 2 and 3).

Detailed results of reliability studies

Hodgkin lymphoma (HL)

Three studies (Barrington 2016, Biggi 2013, Gallamini 2014)[33,36,76] assessed the reliability of PET/CT for patients with advanced-stage classic HL. One study (Barrington)[33] reported reliability of PET/CT for initial staging. All three studies described reliability of interim PET/CT and used a binary version of 5-PS with scores 1-3 regarded "negative" and scores 4-5 regarded "positive." Across the three reliability studies in persons with HL, interrater agreement was moderate to substantial for initial staging and for interim evaluation of treatment.

For initial staging, one moderate quality study (Barrington 2016),[33] to measure agreement between expert and local readers on 1,214 scans from the Response Adapted Therapy in Advance Hodgkin Lymphoma (RATHL) PET-adapted treatment trial. Overall agreement between the original staging from in the RATHL trial and staging based on PET/CT was reported as good to very good with RATHL and baseline PET/CT staging was concordant in 938 (80%) patients. Based on PET/CT upstaging would have occurred in 159 (14%) and downstaging in 74 (6%) patients.

For interim PET/CT, the most recent study (Barrington 2016)[33] was of moderate quality. PET/CT scans were done at baseline, after 2 treatment cycles (PET2), and after initiation of escalated therapy (PET3). Of the available scans, no baseline scans were dual reviewed, 300 PET2 scans were dual-reviewed by the study's core laboratory and the research site's local laboratory, 140 PET2 scans were dual-reviewed at the study's core laboratory, and 47 PET3 scans were dual-reviewed by the study's core laboratory and the research site's local laboratory. Authors evaluated interrater agreement with different thresholds. When the liver threshold was used, interrater agreement ranged from substantial (κ = 0.77, 95% CI 0.68 to 0.86) to almost perfect (κ = 0.91, 95% CI 0.78 to 1.00). When the mediastinal threshold was used, interrater agreement ranged from moderate (κ = 0.58, 95% CI 0.50 to 0.66) to substantial (κ = 0.64, 95% CI 0.55 to 0.73). The authors conclude that these results confirm PET/CT as the modern standard for staging HL and that response assessment using 5-PS is robust, enabling translation of RATHL results into clinical practice.

The second study of interim PET (N=260) (Gallamini 2014)[76] of good quality was conducted to assess the prognostic role of interim PET in patients with HL treated with doxorubicin, bleomycin, vinblastine, and dacarbazine as well as the reproducibility of 5-PS for the interpretation of interim PET. Scans were performed after 2 cycles of therapy. Six reviewers independently interpreted the scans and scored them using 5-PS. Agreement between pairs of reviewers was substantial (κ = 0.69 to 0.84). Authors conclude that 5-PS has good interrater reproducibility in HL and should be used in routine clinical practice.

The third study (Biggi 2013)[36] of interim PET was of good quality. The study aimed to validate the prognostic value of interim PET and to measure concordance among reviewers. Patients (N =260) underwent PET/CT scans after 2 cycles of therapy. Six nuclear medicine experts independently reviewed the scans. Interim scans were compared to baseline scans and analyzed using 5-PS. Overall agreement was excellent (Krippendorf α = 0.76) and agreement between pairs of reviewers ranged from substantial

(κ = 0.69) to almost perfect (κ = 0.84). Authors conclude that 5-PS, combined with detailed instructions, is robust enough to be used as a standard reporting tool for interpretation of interim PET scans.

Pediatric HL

Two studies (Kluge 2016, Furth 2011)[74,98] assessed interrater reliability of interim PET/CT for pediatric patients with HL. The first study (Kluge 2016)[98] was of good quality. Five expert readers independently scored 100 interim (after 2 cycles of therapy) PET/CT scans from pediatric HL patients with early- (n=40), intermediate- (n=21), or advanced- (n=39) stage. Using 5-PS (scores D1 through D5) criteria interrater agreement was only fair (κ = 0.24). Authors noted that probability of concordance was highest for extreme values (>50% for D1 and D5) and lowest for middle values (25% for D2, 36% for D3), concluding that the binary distinction for 5-PS is the most reliable criterion for clinical decision making.

The second study (N=39) (Furth 2011)[74] was of moderate quality. PET/CT scans were performed after 2 cycles of therapy. Using 5-PS for interpretation, interrater agreement was substantial (κ = 0.748), however the review was not entirely independent (2 reviewers from each center came to a consensus).

Aggressive non-Hodgkin lymphoma (aNHL)

One good quality study (Han 2016)[88] evaluated interrater variability of interim and post-treatment PET/CT for patients with DLBCL. Two nuclear medicine physicians independently interim (after 3 cycles of therapy) (n=55) and post-treatment (n=57) scans for 59 adult patients with DLBCL. Authors defined the IHP threshold for negative PET as moderately sized or large residual lesions with FDG uptake as intense as or less intense than mediastinal blood pool structures or lesions smaller than 2 centimeters in diameter with FDG uptake that is not differentiable from surrounding background activity. Using IHP criteria, interrater agreement was substantial (κ = 0.76) with 89% absolute agreement. Authors defined 5-PS thresholds as follows: 1, no FDG uptake above background; 2, FDG uptake ≥ mediastinum; 3, FDG uptake > mediastinum but ≤ liver; 4, FDG uptake moderately > liver; and 5, FDG uptake markedly higher than liver and/or new lesions. Using 5-PS, interrater agreement was moderate (κ = 0.54) with 62% absolute agreement. Authors conclude that interrater agreement for visual interpretation of PET/CT scans for response assessment was unsatisfactory.

Pediatric NHL (all types)

One moderate quality study (Furth 2013)[75] evaluated reliability of interim PET/CT (after 2 cycles of therapy) for 18 pediatric patients with NHL. PET/CT scans were scored using 5-PS with the following thresholds: 1, sites with no uptake; 2, lesions with slightly increased uptake but ≤ uptake of MBPS; 3, lesions with uptake > uptake of MBPS but < uptake of liver; 4, lesions with uptake moderately > uptake of liver; and 5, lesions with uptake markedly > uptake of liver or new lesions. Interrater agreement was substantial (κ=0.618). Authors conclude that 5-PS criteria for response evaluation at interim should thus be used for multicenter trials.

Mixed Population (HL and NHL)

One poor quality study (Toriihara 2017)[152] assessed whether nuclear medicine physicians' interpretations of PET/CT scans reliably suggest lymphoma at diagnosis. Two nuclear physicians read the reports and reviewed the PET/CT scans for 70 patients with suspected lymphoma and came to a consensus on the degree of suspicion (Grade 3: definitely suspicious, Grade 2: probably suspicious, and Grade 1: possibly suspicious). Criteria for grading was: Grade 3, no other diseases referred to in the report; Grade 2, one or two diseases, except for lymphoma, suggested as differential diagnoses and evaluators suspected lymphoma in a degree equal to or greater than other diseases; and Grade 1, more than three differential diagnoses, except for lymphoma, proposed or evaluators clearly suspected other diseases compared to lymphoma. When the degree of suspicion was low, the likelihood of lymphoma was also low (none of the patients assessed as Grade 1 received a confirmed lymphoma diagnosis). However, the false positive rate for cases rated Grade 3 (definitely suspected) was 21.4% (compared to confirmed diagnosis). Methodological limitations of this study include: subjective criteria for image evaluation, readings done based on knowledge of the previous report, lack of independent review, and unclear timing of PET/CT scans.

Diagnostic Accuracy

Table 9. Summary of diagnostic accuracy information: PET for diagnosis of Non-Hodgkin lymphoma

Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
NHL								
Zou (2017)[170] AMSTAR score: 7/11	8 studies (8 retrospective); N=129; QUADAS-2 tool used for appraisal of included studies; overall risk of bias of studies considered to be low 4 Studies used FDG-PET, 3 used FDG-PET/CT and 1 used both. Heterogeneity checked with Chi-Square tests	MRI & histopathology (3 studies); histopathology & clinical-radiological f/u (1 study); histopathology (4 studies)	NHL (primary CNS lymphoma, immunocompetent patients)	84/129 = 65% (calculated)	Range 87.5% to 100%; Pooled 88% (95% CI 80% to 94%)	Range NR-96.3%; Pooled 86% (95% CI 73% to 94%)	Pooled LR+ 3.99 (95% CI 0.73 to 0.94) Pooled LR- 0.11 (95% CI 10.40 to 107.3) Pooled DOR 33.40 (95% CI 10.40 to 107.3) AUC 0.9192	F-FDG PET and PET/CT are valuable diagnostic tools in immunocompetent PCNSL patients; F-FDG-PET and PET/CT are recommended for routine diagnostic imaging in PCNSL
Yang (2017)[161] AMSTAR score: 9/11	6 studies; N=108; QUADAS-2 tool used for appraisal of included studies; overall quality of studies considered acceptable	Pathology, serology & clinical f/u	NHL (primary CNS lymphoma, HIV patients)	NR	100%	Range 75% to 100%	NR	PET may be superior to SPECT in terms of sensitivity and specificity, but it is more expensive and has less supporting data

PET: positron emission tomography; HIV: human immunodeficiency virus; HL: Hodgkin's lymphoma; CNS: central nervous system; CT: computed tomography; FDG: fluorodeoxyglucose; AMSTAR: Assessing the Methodological Quality of Systematic Reviews; NHL: non-Hodgkin's lymphoma; CNS: central nervous system; QUADAS: Quality Assessment of Diagnostic Accuracy Studies; MRI: magnetic resonance imaging; DOR: diagnostic odds ratio; LR+: positive likelihood ratio; LR-: negative likelihood ratio; AUC: area under the curve; SPECT: single-photon emission computed tomography

Initial Staging

Table 10. Summary of diagnostic accuracy information: PET for initial staging of HL and aNHL

Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
Classic HL								
Data from new individual studies								
Cerci (2011)[43] ROB: moderately high	Prospective cohort; N=210	Clinical follow-up, response to treatment	Classic HL	Correctly staged lymphoma: 90% (189/210)	PET vs. CT: 97.9% (95% CI, 95% to 98%) vs. 87.3% (95% CI 84% to 89%)	PET vs. CT: 95.3% (95% CI 91% to 97%) vs. 96.8% (95% CI 93% to 98%)	PPV PET vs. CT: 97.9% (95% CI 96% to 98%) vs.98.4% (95% CI 96% to 99%) NPV PET vs. CT: 93.8% (95% CI 90% to 96%) vs. 77.0% (95% CI 72% to 81%) False Negatives PET vs. CT: 14/210 (6.7%) vs. 55/210 (26.2%) False Positives PET vs. CT: 7/210 (3.3%) vs 6/210 (2.9%)	FDG PET is highly accurate for initial staging in patients with HL
Grellier (2014)[82] ROB: moderately low	Retrospective case series; N=35	Histological confirmation or clinical and biological follow-up when histology could not be obtained	Nodular lymphocyte-predominant Hodgkin lymphoma (NLPHL)	Positive lesions: 103/455 (22.6%)	100% (per-patient analysis) 100% (per-site analysis)	99% (per-site analysis)	PPV (per-site analysis) 97% NPV (per-site analysis) 100%	FDG PET/CT is a useful clinical tool for initial staging of NLPHL
SRs from 2011 report								
Kirby (2007)[97] (included in MSAC 2010)	20 studies; N=854 patients (from 13 patient-based studies) and N=3,658 lesions (based on 6 lesion-based studies);	CT, Gallium, clinical follow-up, bone marrow biopsy, or conventional staging	HL	NR	Median (from 9 patient-based studies): 93.2%	Median (from 9 patient-based studies): 87.7%	NR	PET has a high diagnostic accuracy in the staging and restaging of

Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
Reported in previous HTA	no appraisal of individual studies							lymphoma patients
Aggressive NHL								
Data from new SRs								
Zhou (2014)[164] AMSTAR score: 7/11	8 studies (7 retrospective, 1 not documented); N=135; QUADAS tool for appraisal of included studies; overall risk of bias of studies not reported Heterogeneity assessed via chi-square test	Pathology, or clinical and imaging follow-up	aNHL (extranodal natural killer/T-cell lymphoma)	Positive diagnosis Patient-based data (calculated): 116/118 = 98%	6 "patient based" studies only: Range 81% to 97% Pooled 95% (95%CI 89% to 98%)	6 "patient based" studies only Range 17% to 100% Pooled 40% (95%CI 9% to 78%)	DOR Patient-based data: 13.28 (95% CI 2.60 to 67.82) Pooled LR+ Patient-based data: 1.63 (95% CI 0.91 to 2.91) Pooled LR+ Patient-based data: 0.13 (95% CI 0.05 to 0.38) AUC Patient-based data: 0.8537	PET/CT has a high diagnostic accuracy for staging patients with NK/T-cell lymphoma

PET: positron emission tomography; HL: Hodgkin's lymphoma; CT: computed tomography; FDG: fluorodeoxyglucose; AMSTAR: Assessing the Methodological Quality of Systematic Reviews; NHL: non-Hodgkin's lymphoma; QUADAS: Quality Assessment of Diagnostic Accuracy Studies; DOR: diagnostic odds ratio; LR+: positive likelihood ratio; LR-: negative likelihood ratio; AUC: area under the curve; ROB: risk of bias; HTA: health technology assessment; MSAC: Medical Service Advisory Committee; NK: natural killer

Table 11. Summary of diagnostic accuracy information: PET for initial staging of iNHL and mixed HL and NHL

Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
Indolent NHL								
Data from new SRs								
Adams (2017a)[20] AMSTAR score: 7/11	7 studies (1 prospective, 6 retrospective); N=349; QUADAS-2 tool for appraisal of included studies; overall risk of bias of studies considered to be high	BMB, clinical examination, CT, clinical-radiological confirmation	iNHL (Follicular lymphoma); additional value of FDG-PET to CT	See other outcomes	NR	NR	Proportion of patients upstaged based on FDG-PET results Range: 0.0% to 45.2% Pooled summary: 18.7% (95% CI 10.8% to 30.4%)	Available studies have many methodological errors; future research is needed before FDG-PET can be recommended for routine staging of follicular lymphoma
MSAC (2016)[54]	Relapse and new diagnosis 1 study; N=45; no appraisal of individual studies	CT	Newly diagnosed and relapsed iNHL	Positive nodal lesions	77% (95% CI 69% to 84%)	98% (95% CI 97% to 99%)	NR	PET/CT has the same specificity as CT and higher sensitivity than CT for detecting nodal regions or a per-lesion basis
	New diagnosis 3 studies; N=130; no appraisal of individual studies	Bone marrow biopsy	iNHL	Positive nodal lesions	NR	NR	PET/CT identified additional lesions in 16-36% of patients who were negative on bone marrow biopsy 15% to 100% of additional lesions were true positives	None

Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
Data from new individual studies								
Carrillo-Cruz (2015)[40] ROB: moderately high	Retrospective case series; N=25	Histology	Marginal zone B cell lymphoma (MZL)	Positive diagnosis	Overall: 96% (24/25) Extranodal sites: 95.5%	NR	NR	PET/CT has high sensitivity in newly diagnosed patients with MZL, especially for staging patients with nodal and extranodal subtypes
Mixed HL & NHL								
SRs from 2011 Report								
Kwee (2008)[100] (included in MSAC 2010) Reported in previous HTA	FDG-PET 17 studies; N=860; adapted checklist from prior literature for appraisal of included studies; overall quality of studies considered moderate	CT, histology, or clinical follow-up	HL & aggressive NHL	Correct restaging (criteria unclear)	HL 87.5% (based on one lesion-based study) NHL 83.3% (based on one lesion-based study) Mixed HL and NHL Range: 71.4% (patient-based) and 100%	HL 100% (based on one lesion-based study) NHL 100% (based on one lesion-based study) Mixed HL and NHL Range: 86.2% (patient-based) and 100% (patient-based)	NR	FDG-PET/CT fusion outperforms both CT alone and FDG-PET alone

Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
					(patient-based)			
	FDG-PET/CT 4 studies; N=234; adapted checklist from prior literature for appraisal of included studies; overall quality of studies considered moderate	CT, histology, or clinical follow-up	HL & aggressive NHL	Correct restaging (criteria unclear)	Mixed HL and NHL Range: 97.9% (region-based) to 100% (region-based)	Mixed HL and NHL 100% (based on one, region-based study)	NR	
Facey (2007)[68] (included in MSAC 2010) Reported in previous HTA	1 SR and 7 studies;	CT or Gallium; concordance with conventional work-up	HL and NHL	Lymph-node involvement, correctly staged lymphoma, positive nodal lesions	Range 79% to 100%	≥90%	NR	PET has greater sensitivity that Gallium-67 citrate; PET was better than or equal to CT for pooled HL and NHL patients
Kirby (2007)[97] (included in MSAC 2010) Reported in previous HTA	20 studies; N=854 patients (from 13 patient-based studies) and N=3,658 lesions (based on 6 lesion-based studies); no appraisal of individual studies	CT, Gallium, clinical follow-up, bone marrow biopsy, or conventional staging	HL and NHL (all types)	Correct staging	HL and NHL Median (from 14 patient-based studies): 90.3% NHL Median (from 9 patient-based	HL and NHL Median (from 14 patient-based studies): 91.1% NHL Median (from 9 patient-based studies): 93.8%	True positives: 91% False positives: 10%	PET has a high diagnostic accuracy in the staging and restaging of lymphoma patients; diagnostic accuracy may be higher in patients with HL than in patients with NHL

Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
					studies): 87.5.2%			
MSAC (2010)[53] Reported in previous HTA	2 studies; N=79; QUADAS tool for appraisal of included studies; overall quality of studies considered fair to poor	CT	HL and aggressive NHL	Positive disease sites	NR	NR	PET detected additional sites in 18-60% of patients; ratio of true positive to true negative sites detected by PET was 3:1	A large proportion of additional negative PET results were false negatives; these studies were conducted with older stand-alone PET scanners and results may not be applicable to combined PET/CT

PET: positron emission tomography; HL: Hodgkin's lymphoma; CT: computed tomography; FDG: fluorodeoxyglucose; NHL: non-Hodgkin's lymphoma; DLBCL: diffuse large B-cell lymphoma; QUADAS: Quality Assessment of Diagnostic Accuracy Studies; ROB: risk of bias; HTA: health technology assessment; MSAC: Medical Service Advisory Committee AMSTAR: Assessing the Methodological Quality of Systematic Reviews

Interim PET

Table 12. Summary of diagnostic accuracy information: interim PET for HL and aNHL

Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
Classic HL								
SRs from 2011 report								
Kwee (2008)[100] (included in MSAC 2010) Reported in previous HTA	FDG-PET 17 studies; N=860; adapted checklist from prior literature for appraisal of included studies; overall quality of studies considered moderate	CT, histology, or clinical follow-up	HL	Correct restaging (criteria unclear)	Range: 85% (lesion-based) to 100% (patient-based)	Range: 57.1% (patient-based) to 100% (patient-based)	NR	FDG-PET/CT fusion outperforms both CT alone and FDG-PET alone
	FDG-PET/CT 4 studies; N=234; adapted checklist from prior literature for appraisal of included studies; overall quality of studies considered moderate	CT, histology, or clinical follow-up	HL	Correct restaging (criteria unclear)	100% (based on one patient-based study)	90.7% (based on one patient-based study)	NR	
Aggressive NHL								
SRs from 2011 report								
Kwee (2008)[100] (included in MSAC 2010) Reported in previous HTA	FDG-PET 17 studies; N=860; adapted checklist from prior literature for appraisal of included studies; overall quality of studies considered moderate	CT, histology, or clinical follow-up	aNHL	Correct staging (criteria unclear)	Range: 60% (patient-based) to 100% (lesion-based)	Range: 80% (patient-based) to 100% (patient-based and lesion-based	NR	FDG-PET/CT fusion outperforms both CT alone and FDG-PET alone

PET: positron emission tomography; HL: Hodgkin's lymphoma; CT: computed tomography; FDG: fluorodeoxyglucose; aNHL: aggressive non-Hodgkin's lymphoma; HTA: health technology assessment; MSAC: Medical Service Advisory Committee AMSTAR: Assessing the Methodological Quality of Systematic Reviews

Table 13. Summary of diagnostic accuracy information: interim PET for iNHL and mixed HL and NHL

Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
Indolent NHL								
Data from new SRs								
MSAC (2016)[54]	1 study; N=45; no appraisal of individual studies	CT±BMAT (not well defined)	iNHL	Positive nodal and extranodal lesions	72% (95% CI 47% to 90%)	100% (95% CI 87% to 100%)	NR	PET/CT has higher specificity but lower sensitivity than CT in patients with indolent NHL; evidence considered insufficient for PET/CT as a replacement for assessment of treatment response
SRs from 2011 report								
MSAC (2010)[53] Reported in previous HTA	2 studies; N=167; tool adapted from literature for appraisal of independent studies; overall quality of studies considered fair	CT	iNHL	Positive disease sites	NR	NR	PET positive for disease not evident on CT in 3% to 28% of patients	PET is more accurate than CT for initial staging and assessment of suspected relapse in patients with indolent NHL
Mixed HL & NHL								
Data from new SRs								
Fitzpatrick (2018)[70] AMSTAR score: 8/11	8 studies (8 prospective); N=221; QADAS-2 tool for appraisal of included studies; overall risk of bias of studies considered to be unclear	Diffusion weighted MRI was compared with PET/CT as reference standard	5 studies with NHL (3 DLBCL); 3 studies with HL & NHL	Positive lesions (based on 5 studies, calculated):71% (114/161)	DW-MRI Pooled (from 5 studies with lesion based data) 94.7% (95% CI 88.9 to 98)	DW-MRI Pooled (from 5 studies with lesion based data) 99.3% (95% CI 98.9 to 99.6)	NR	DWI-MRI has good agreement with PET/CT for assessing treatment response in lymphoma

Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
Adams (2016a)[22] AMSTAR score: 7/11	11 studies (5 prospective, 6 retrospective); N=139; QUADAS-2 tool for appraisal of included studies; overall risk of bias of studies considered to be high	Biopsy	HL & NHL (various types)	Positive lesions	NR	NR	False positives FDG-PET for NHL during or after treatment: Range 7.7% to 90.5% Pooled 55.7% (95% CI 32.6% to 76.6%) Interim FDG-PET for HL: No studies available Interim FDG-PET for NHL: Pooled 83% (95%CI 72% to 90.2%) End of treatment FDG-Pet for HL: Pooled 23.1% (95%CI 4.7% to 64.5%) End of treatment FDG-PET for NHL:	FDG-PET scans result in a high number of false-positive FDG-avid lesions; the role of interim and end-of-treatment FDG-PET should be reconsidered

Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
							Pooled 31.5% (95%CI 3.9 to 83.9%)	
SRs from 2011 Report								
MSAC (2010)[53] Reported in previous HTA	2 studies (1 prospective, 1 retrospective); N=71; tool adapted from literature for appraisal of independent studies; overall quality of studies considered fair	Clinical follow-up	HL and aNHL	Residual masses, relapse	Range: 96% to 100%	Range: 70% to 99%	TP (based on one study): 9% (3/33) TN (based on one study): 36% (12/33)	PET is more accurate than CT for assessment of treatment response
Kwee (2008)[100] (included in MSAC 2010) Reported in previous HTA	FDG-PET/CT 4 studies; N=234; adapted checklist from prior literature for appraisal of included studies; overall quality of studies considered moderate	CT, histology, or clinical follow-up	HL and aNHL	Correct restaging (criteria unclear)	Range: 91.3% (patient-based) to 100% (patient-based)	Range: 90.6% (patient-based) to 100% (patient-based and region-based)	NR	FDG-PET/CT fusion outperforms both CT alone and FDG-PET alone

PET: positron emission tomography; HL: Hodgkin's lymphoma; CT: computed tomography; FDG: fluorodeoxyglucose; NHL: non-Hodgkin's lymphoma; aNHL: aggressive non-Hodgkin's lymphoma; iNHL: indolent non-Hodgkin's lymphoma; DLBCL: diffuse large B-cell lymphoma; QUADAS: Quality Assessment of Diagnostic Accuracy Studies; DW-MRI: diffusion-weighted magnetic resonance imaging; TP: true positive; TN: true negative; HTA: health technology assessment; MSAC: Medical Service Advisory Committee AMSTAR: Assessing the Methodological Quality of Systematic Reviews

Post-treatment surveillance

Table 14. Summary of diagnostic accuracy information: post-treatment surveillance for HL and aNHL

Author (year)	Evidence Base Available	Comparator	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
Classic HL								
Data from new individual studies								
Picardi (2014)[120] ROB: moderately low	RCT; N=300	Histology	Advanced-stage HL	Relapse: 40/40 (100%)	PET/CT vs. US/Chest Radiography 100% (95% CI 91.2% to 100%) vs. 97.5% (95% CI 87.1% to 99.5%)	PET/CT vs. US/Chest Radiography 86.3% (95% CI 78.7% to 91.5%) vs. 96.3% (95% CI 91.0% to 98.5%)	PET/CT vs. US/Chest Radiography PPV: 72.7% (95% CI 59.0% to 83.8%) vs. 90.7% (95% CI 77.8% to 97.4%) NPV: 100% (95% CI 96.2% to 100%) vs. 99.1% (95% CI 94.9% to 99.9%) FN: 0/40 (0%) vs. 1/40 (2.5%) FP: 15/40 (13.7%) vs. 4/40 (3.7%) LR+: 7.3 (95% CI 4.5 to 11.7) vs. 26.8 (95% CI 10.2 to 70.2) LR-: 0.00 vs. 0.02 (95% CI 0.04 to 0.18)	Compared with US/chest radiography, PET/CT had lower specificity and PPV; US/chest radiography should be used in routine surveillance imaging instead of PET/CT

Author (year)	Evidence Base Available	Comparator	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
Aggressive NHL								
Data from new individual studies								
Hong (2014)[90] ROB: moderately high	Retrospective cohort; N=106 patients (856 visits)	Histological or cytological confirmation	DLBCL	Relapse: 14.2% (15/106)	PET/CT vs. CT (per-visit) 100% (95% CI 30.5% to 100.0%) vs. 100% (95% CI 30.5% to 100.0%)	PET/CT vs. CT (per-visit) 86.1% (95% CI 79.8 to 91.0) vs. 98.3% (95% CI 96.5 to 99.3)	PET/CT vs. CT (per-visit) PPV: 11.5% (95% CI 2.6% to 30.2%) vs. 30.0% (95% CI 7.0% to 65.2%) NPV: 100% (95% CI 97.4% to 100.0%) vs. 100% (95% CI 99.1% to 100.0%) FP (based on 23 scans in 21 patients): 23/165 (13.7%) vs. 7/407 (1.7%)	Compared to CT, FDG PET/CT had a higher FP rate, resulting in unneeded procedures; routine imaging has limited value in patients with DLBCL
Cheah (2014)[45] ROB: moderately low	Retrospective case series; N=55	Histology	Transformed iNHL	Relapse: 10% (18/180)	83%	94%	TP: 15/180 (9%) FP excluding indeterminate: 4/180 (2%) scans FP including indeterminate: 9/180 (5%) of scans Indeterminate: 7/180 (4%) TN: 153/180 (85%)	In patients with transformed iNHL, PET/CT has high sensitivity but a low FP rate for detecting subclinical relapses; surveillance of patients with transformed iNHL after achieving complete remission is not indicated

Author (year)	Evidence Base Available	Comparator	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
							FN including indeterminate: 3/180 (1.7%) FN excluding indeterminate: 1/180 (0.5%) PPV: 63% NPV: 98%	

PET: positron emission tomography; HL: Hodgkin's lymphoma; CT: computed tomography; US: ultrasound; FDG: fluorodeoxyglucose; NHL: non-Hodgkin's lymphoma; aNHL: aggressive non-Hodgkin's lymphoma; iNHL: indolent non-Hodgkin's lymphoma; DLBCL: diffuse large B-cell lymphoma; NPV: negative predictive value; PPV: positive predictive value; TP: true positive; TN: true negative; FP: false positive; FN: false negative; LR+: positive likelihood ratio; LR-: negative likelihood ratio; ROB: risk of bias

Table 15. Summary of diagnostic accuracy information: post-treatment surveillance for mixed HL and NHL

Author (year)	Evidence Base Available	Comparator	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
Mixed HL & NHL								
Data from new SRs								
Cohen (2017)[52] AMSTAR score: 6/11	6 studies; N=621; no appraisal of individual studies	Observation without imaging	HL and DLBCL	Relapse (calculated based on 5 studies reporting data): 21% (129/621)	NR	NR	NR	Patients with HL and DLBCL who achieve complete remission should not receive routing surveillance imaging; further research is needed to determine subpopulations which might benefit from surveillance imaging
Patel (2013)[117] AMSTAR score: 8/11	4 studies (4 retrospective); N=312; QUADAS tool for appraisal of included studies; three grade B	Some combination of: other imaging modalities, laboratory tests, clinical examination, and biopsy	HL and NHL	Relapse	Range: 75% to 100%	Range: 43% to 92%	NPV Range: 98% to 100% PPV Range: 20% to 100%	There is a lack of evidence supporting PET/CT in post-treatment surveillance; risks of PET/CT include false-positives and radiation exposure; more research is needed to assess the clinical impact of PET/CT in surveillance

Author (year)	Evidence Base Available	Comparator	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
	studies and 1 grade C study							
SRs from 2011 report								
Facey (2007)[68] (included in MSAC 2010) Reported in previous HTA	1 HTA and 3 studies;	Clinical follow-up	HL and aNHL	Relapse	75% to 80%	90%	NR	Post-therapy PET has similar sensitivity and better specificity than CT to evaluate residual masses

PET: positron emission tomography; HL: Hodgkin's lymphoma; CT: computed tomography; FDG: fluorodeoxyglucose; NHL: non-Hodgkin's lymphoma; aNHL: aggressive non-Hodgkin's lymphoma; DLBCL: diffuse large B-cell lymphoma; QUADAS: Quality Assessment of Diagnostic Accuracy Studies; NPV: negative predictive value; PPV: positive predictive value; HTA: health technology assessment; MSAC: Medical Service Advisory Committee; AMSTAR: Assessing the Methodological Quality of Systematic Reviews

Prognosis

Table 16. Summary of diagnostic accuracy information: prognosis for HL and aNHL

Timing of PET	Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
Classic HL									
Data from new SRs									
Interim (after 1 to 4 cycles of therapy)	Ziakas (2012)[166] AMSTAR score: 9/11	14 studies; N=1,328; QUADAS-2 tool for appraisal of included studies; overall quality of studies	Clinical follow-up	HL	Treatment failure: 212/1,328 (16%)	Pooled: 67% (95% CI 57% to 76%)	Pooled: 89% (95% CI 84% to 93%)	TP: 144/1,328 (11%) TN: 982/1,328 (74%) FP: 134/1,328 (10%) FN: 68/1,328 (5%)	Diagnostic performance of interim PET is limited and should not be used as a substitute for conventional

Timing of PET	Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
		considered moderate						Pooled LR+: 6.2 (95% CI 3.9 to 10.0) Pooled LR-: 0.37 (95% CI 0.27 to 0.50) NPV: 93% (95% CI 85% to 100%) AUC: 0.85 (95% CI 0.82 to 0.88)	staging and follow-up
Pre-SCT	Adams (2016b)[13] AMSTAR score: 7/11	PFS: 5 studies (2 prospective, 3 retrospective); N=626; QUIPS tool for appraisal of included studies; overall quality of included studies considered moderate	Response/non-response before SCT confirmed by histological biopsy	Refractory/ relapsed HL treated with autologous SCT	Range of PFS for FDG-PET positive patients: 0% to 52% Range of PFS for FDG-PET negative patients: 55% to 85%	Pooled: 67.2% (95 % CI 58.2% to 75.3%)	Pooled: 70.7% (95 % CI 64.2% to 76.5%)	NR	Patients with positive pre-transplant FDG-PET scans tend to have worse outcomes (both PFS and OS) than patients with negative pre-transplant FDG-PET scans; It is unclear whether pre-transplant FDG-PET should be incorporated into routine clinical practice
		OS: 2 studies (1 prospective, 1 retrospective); N= 238; QUIPS tool for appraisal of included studies; overall quality of included studies considered moderate	Response/non-response before SCT confirmed by histological biopsy	Refractory/ relapsed HL treated with autologous SCT	Range of OS for FDG-PET positive patients: 17% to 77% Range of OS for FDG-PET negative patients: 78% to 100%	Pooled: 74.4% (95% CI 58.8% to 86.5%)	Pooled: 58.0% (95 % CI 49.3% to 66.3 %)	NR	

Timing of PET	Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
End of treatment (after range of 2 to 19 cycles of therapy)	Adams (2015a)[17] AMSTAR score: 7/11	5 studies (4 prospective, 1 retrospective); N=727; QUIPS tool for appraisal of included studies; overall quality of included studies considered moderate	Disease relapse confirmed by histological biopsy, CT, or induction of second line therapy	HL	See outcomes	NR	NR	Pooled disease relapse proportion with FDG-PET-negative residual mass: 6.8% (95% CI 2.6% to 12.5%)	FDG-PET-negative residual masses are not proven to be associated with worse outcomes compared to post-treatment FDG-PET–based complete remission without a residual mass
SRs from 2011 report									
Interim (after 1 to 4 cycles of therapy)	MSAC (2010)[53] Reported in previous HTA	1 study (1 prospective); N=99; tool adapted from literature for appraisal of independent studies	Clinical follow-up	Advanced-stage HL	PFS	NR	NR	2 year PFS in PET non-responders: 0% 2 year PFS in PET responders: 96%	Interim PET is predictive of 2 year PFS for patients with advanced-stage HL
	Terasawa (2009)[147] Reported in previous HTA	13 studies (8 prospective, 5 retrospective); N=360 HL; QUADAS tool for appraisal of included studies; overall quality of	Biopsy or clinical follow-up	HL	Relapse or progression (range): 20% to 50%	Range: 67% to 100% Pooled: 81% (95% CI 72% to 89%)	Range: 94% to 100% Pooled: 97% (95% CI 94% to 99%)	LR+: 28.4 (95% CI 14.2 to 56.7) LR-: 0.19 (95% CI 0.12 to 0.30)	Interim PET should not yet be employed in a routine setting; future prospective trials should evaluate PET

Timing of PET	Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
		studies considered poor							for patients with advanced-stage HL;
End of treatment (after completion of therapy)	MSAC (2010)[53] Reported in previous HTA	1 study (1 NR); N=40; tool adapted from literature for appraisal of independent studies	Clinical follow-up	HL	PFS	NR	NR	3 year PFS in PET non-responders: 33% 3 year PFS in PET responders: 94%	End of treatment PET is predictive of 3 year PFS for patients with HL
	Terasawa (2008)[148] (included in MSAC 2010) Reported in previous HTA	15 studies (3 prospective, 12 retrospective; N=474; QUADAS tool for appraisal of included studies; overall quality of studies considered poor	Biopsy or clinical follow-up	HL	Residual mass: 35% to 72% Relapse: 0% to 55%	Post-therapy evaluations (range): 50% to 100% Residual mass evaluations (range): 43% to 100%	Post-therapy evaluations (range): 67% to 100% Residual mass evaluations (range): 67% to 100%	AUC Post-therapy evaluations: 0.94 Residual mass evaluations: 0.93	FDG-PET has good accuracy for detecting residual disease in patients with HL after first-line therapy
	Zijlstra (2006) Reported in previous HTA	15 studies (5 prospective, 9 retrospective, 1 both prospective and retrospective); N=705; criteria list recommended by the Cochrane Methods Working Group on Systematic Review of Screening and Diagnostic test	Clinical follow-up	HL	Relapse: 14% to 46%	Pooled: 84% (95% CI 71% to 91%)	Pooled: 90% (95% CI 84% to 94%)	NR	FDG-PET is useful in evaluating response to treatment

Timing of PET	Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
		for appraisal of included studies; overall quality of studies considered moderate							
Aggressive NHL									
Data from new SRs									
Interim (after 2 to 4 cycles of therapy)	Liao (2017)[103] AMSTAR score: 9/11	PFS: 8 studies (4 prospective, 4 retrospective); N=832; Newcastle-Ottawa tool for appraisal of included studies; overall quality of included studies considered to be high	NR	aNHL	PFS	NR	NR	PET/CT positive vs. PET/CT negative Pooled HR (Deauville method, based on 2 studies): 2.91 (95% CI 1.68 to 5.04) Pooled HR (IHP method, based on 3 studies): 2.92 (95% CI 1.81 to 4.70) Pooled HR (no standard method, based on 3 studies): 3.47 (95% CI 1.53 to 7.87) Pooled HR (all methods): 2.93 (95% CI 2.93 to 3.90)	PET/CT with visual interpretation can accurately assess disease prognosis after 3-4 cycles of chemotherapy; more research is needed to determine if the outcome of PET/CT can be used to change treatment.

Timing of PET	Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
		OS: 9 studies (6 prospective, 3 retrospective); N=925; Newcastle-Ottawa tool for appraisal of included studies; overall quality of included studies considered to be high	NR	aNHL	OS	NR	NR	PET/CT positive vs. PET/CT negative Pooled HR (Deauville method, based on 2 studies):1.99 (95% CI 0.67 to 5.93) Pooled HR (IHP method, based on 3 studies): 2.29 (95% CI 1.38 to 3.790) Pooled HR (no standard method, based on 4 studies): 3.63 (95% CI 1.66 to 7.90) Pooled HR (all methods): 2.55 (95% CI 1.76 to 3.68)	
	Adams (2016c)[14] AMSTAR score: 8/11	PFS: 6 studies (3 prospective, 3 retrospective); N=519; QUIPS tool for appraisal of included studies; overall quality of studies considered to be moderate	None	DLBCL treated with R-CHOP	Relapse (calculated): 24.5% (127/519)	Range: 27.3% to 62.5%	Range: 55.8% to 90.7%	AUC: 0.651	Interim FDG-PET has some correlation with outcomes in R-CHOP treated DLBCL but prognostic value is suboptimal; there is no scientific base to support the use of interim FDG-PET in R-
		OS: 3 studies (1 prospective, 2 retrospective; N=323; QUIPS tool for appraisal	NR	DLBCL treated with R-CHOP	Death (calculated): 21.7% (70/323)	Range: 35.3% to 82.2%	Range: 67.1% to 91.5%	AUC: 0.817	

Timing of PET	Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
		of included studies; overall quality of studies considered to be moderate							CHOP treated DLBCL
	Sun (2015)[141] AMSTAR score: 8/11	6 studies; N=605; QUADAS tool for appraisal of included studies; overall quality of studies considered to be high	NR	DLBCL treated with R-CHOP	PFS	Range: 21.2% to 89.7% Pooled: 52.4%	Range: 37.4% to 90.7% Pooled: 67.8%	TP 0.6978	Pooled sensitivity and specificity of interim PET/CT for predicting outcome of DLBCL patients treated with R-CHOP is not satisfactory
	Zhu (2013)[165] AMSTAR score: 8/11	13 studies (5 prospective, 8 retrospective); N=1,160; modified scale based on European lung cancer working party quality scale for biological prognostic factors for lung cancer for appraisal of included studies	NR	aNHL	PFS and OS	NR	NR	PET/CT positive vs. PET/CT negative PFS DLBCL (9 studies): Pooled HR 4.4 (95% CI 3.34 to 5.81) OS DLBCL (8 studies): Pooled HR 3.99 (95% CI 2.63 to 6.06)	Interim PET is a prognostic factor for survival and recurrence in non-DLBCL NHL; future cost-effectiveness analyses are needed
Pre-SCT	Adams (2017b)[15]	PFS: 7 studies (4 prospective, 3	NR	aNHL	Treatment failure in FDG-PET-	Pooled: 54.0% (95% CI	Pooled: 73.1%	NR	Pre-SCT FDG-PET cannot be recommende

Timing of PET	Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
	AMSTAR score: 7/11	retrospective); N=374; QUIPS tool for appraisal of included studies; overall quality of studies considered low			positive group (calculated): 61/125=49% Treatment failure in FDG-PET-negative group (calculated): 52/226=23%	44.4% to 63.4%) Newly diagnosed (based on 2 studies): 20.0% Refractory/relapsed (based on 3 studies): 68.1%	(95% CI: 67.0% to 78.6%) Newly diagnosed (based on 2 studies): 70.0% Refractory/relapsed (based on 3 studies): 72.1%		d for predicting outcomes in aggressive NHL after autologous SCT because available studies are poor quality and reported diagnostic performance is low
		OS: 4 studies (2 prospective, 2 retrospective); N=196; QUIPS tool for appraisal of included studies; overall quality of studies considered low	NR	aNHL	Death in FDG-PET-positive group (calculated): 24/70=34% Death in FDG-PET-negative group (calculated): 20/121=17%	Pooled: 54.5% (95% CI: 38.8% to 69.6%) Newly diagnosed (based on 1 study): 8.3% Refractory/relapsed (based on 2 studies): 77.31%	Pooled: 68.7% (95% CI: 60.5% to 76.1%). Newly diagnosed (based on 1 study): 30.5% Refractory/relapsed (based on 2 studies): 69.6%	NR	
End of treatment (after completion of therapy)	Adams (2015b)[18] AMSTAR score: 7/11	PFS: 5 studies (1 prospective, 4 retrospective); N=887; QUIPS tool for appraisal of included	NR	DLBCL treated with R-CHOP	Relapse rate among all patients with CR according to end-of-treatment FDG-PET	NR	NR	NR	Negative end-of-treatment FDG-PET does not accurately predict a favorable outcome in a

Timing of PET	Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
		studies; overall quality of studies not reported			range 7.0% to 20.0%				non-negligible proportion of patients
		OS: 3 studies (3 retrospective); N=730; QUIPS tool for appraisal of included studies; overall quality of studies not reported	NR	DLBCL treated with R-CHOP	See outcomes	NR	NR	5-year OS among all patients with CR according to end-of-treatment FDG-PET (based on 1 study): 83% Range of estimated 3-year OS among all patients with CR according to end-of-treatment FDG-PET (based on 2 studies): 90% to 93.6%	
	Zhu (2013)[165] AMSTAR score: 8/11	13 studies (5 prospective, 8 retrospective); N=1,160; modified scale based on European lung cancer working party quality scale for biological prognostic factors for lung cancer for appraisal of included studies	NR	NHL	PFS and OS	NR	NR	PET/CT positive vs. PET/CT negative PFS DLBCL (3 studies): Pooled HR 6.75 (1.72 to 26.5) Non-DLBCL (4 studies): Pooled HR 4.05 (2.68 to 6.1) OS DLBCL (4 studies): Pooled HR 5.91 (95% CI 3.15 to 11.09)	End of treatment PET in DLBCL and end of treatment PET in non-DLBCL are independent prognostic factors for survival and recurrence; future cost-effectiveness analyses are needed
SRs from 2011 report									
Interim (after 1 to 4 cycles	Terasawa (2009)[147]	13 studies (8 prospective, 5 retrospective);	Biopsy or clinical follow-up	DLBCL	Relapse or progression	Range: 50% to 100%	Range: 73% to 100%	LR+: 5.9 (95% CI 2.8 to 12.3)	Interim PET should not yet be employed

Timing of PET	Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
of therapy)	Reported in previous HTA	N=311; QUADAS tool for appraisal of included studies; overall quality of studies considered poor			(range): 27% to 47%	Pooled: 78% (95% CI 64% to 87%)	Pooled: 87% (95% CI 75% to 93%)	LR-: 0.26 (95% CI 0.15 to 0.46)	in a routine setting; evidence of clinical utility for interim PET in patients with DLBCL is lacking
End of treatment (after completion of therapy)	MSAC (2010)[53] Reported in previous HTA	1 study (1 NR); N=61; tool adapted from literature for appraisal of independent studies	Clinical follow-up	aNHL	PFS	NR	NR	3 year PFS in PET non-responders: 7% 3 year PFS in PET responders: 87%	End of treatment PET is predictive of 3 year PFS for patients with aNHL
	Terasawa (2008)[148] (included in MSAC 2010) Reported in previous HTA	8 studies (1 prospective; 7 retrospective); N=254; QUADAS tool for appraisal of included studies; overall quality of studies considered poor	Biopsy or clinical follow-up	aNHL	Residual mass: 26% to 56% Relapse: 33% to 67%	Post-therapy evaluations (range): 33% to 77% Residual mass evaluations (range):33% to 87%	Post-therapy evaluations (range): 82% to 100% Residual mass evaluations (range): 75% to 100%	NR	Current literature is of poor quality; clinicians should be cautious about making decisions based solely on PET results
	Zijlstra (2006) Reported in previous HTA	15 studies (5 prospective, 9 retrospective, 1 both prospective and retrospective); N=705; criteria list recommended by the Cochrane	Clinical follow-up	aNHL	Relapse: 14% to 46%	Pooled: 72% (95% CI 61% to 82%)	Pooled: 100% (95% CI 97% to 100%)	NR	FDG-PET is useful in evaluating response to treatment

Timing of PET	Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
		Methods Working Group on Systematic Review of Screening and Diagnostic test for appraisal of included studies; overall quality of studies considered moderate							

PET: positron emission tomography; HL: Hodgkin's lymphoma; CT: computed tomography; FDG: fluorodeoxyglucose; AMSTAR: Assessing the Methodological Quality of Systematic Reviews; PFS: progression free survival; OS: overall survival; QUIPS: Quality in Prognosis Studies; SCT: stem cell transplant; HR: hazard ratio; NHL: non-Hodgkin's lymphoma; aNHL: aggressive non-Hodgkin's lymphoma; QUADAS: Quality Assessment of Diagnostic Accuracy Studies; LR+: positive likelihood ratio; LR-: negative likelihood ratio; HTA: health technology assessment; MSAC: Medical Service Advisory Committee; R-CHOP: rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone; TP: true positive; TN: true negative; FP: false positive; FN: false negative; IHP: International Harmonization Project

Table 17. Summary of diagnostic accuracy information: prognosis for iNHL and mixed HL and NHL

Timing of PET	Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
Indolent NHL									
Data from new SRs									
Interim (not defined)	Adams (2016d)[19] AMSTAR score: 7/11	8 studies (2 prospective, 6 retrospective); N=748; QUIPS tool for appraisal of included studies	NR	Indolent follicular lymphoma	Death (based on 2 studies): Range for FDG-PET positive 6.1% to 21.8% Range for FDG-PET negative 3.3% to 20% Relapse (based on 3 studies): Range for FDG-PET positive 84.6% to 100% Range for FDG-PET negative 5.9% to 15.6%	NR	NR	Interim FDG-PET: 2/3 studies reported no significant difference in PFS between FDG-PET positive and negative patients; 2/2 studies reported no significant difference in OS between FDG-positive and negative patients	Evidence does not support the use of interim FDG-PET in follicular lymphoma
End of treatment (after completion of therapy)	Adams (2016d)[19] AMSTAR score: 7/11	8 studies (2 prospective, 6 retrospective); N=748; QUIPS tool for appraisal of included studies	NR	Indolent follicular lymphoma	Death (based on 2 studies): Range for FDG-PET positive 6.1% to 21.8% Range for FDG-PET negative 3.3% to 20% Relapse (based on 3 studies): Range for FDG-PET positive 84.6% to 100%	NR	NR	End-of-treatment FDG-PET: 5/6 studies reported FDG-PET positive patients to have significantly worse PFS; 3/8 studies reported worse OS in FDG-PET positive patients	Some studies suggest end-of-treatment FDG-PET is predictive of PFS and OS but the quality of the evidence is poor

Timing of PET	Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
					Range for FDG-PET negative 5.9% to 15.6%				
	MSAC (2016)[54]	1 study; N=246 scans; no appraisal of individual studies	CT	Indolent follicular lymphoma	PFS and OS	NR	NR	PET/CT positive vs. PET/CT negative Pooled PFS HR: 3.9 (95% CI 2.5 to 5.9) Pooled OS HR: 6.7 (95% CI 2.4 to 18.5)	PET/CT is predictive of both PFS and OS while CT was weakly predictive of PFS but not OS
	Pyo (2013)[125] AMSTAR score: 9/11	8 studies (1 prospective; 7 retrospective); N= 577; REMARK tool for appraisal of included studies; overall quality of included studies considered moderate	Clinical follow-up, with or without pathologic confirmation of the suspected residual/relapsed/ progressed lesion	Follicular lymphoma	PFS	NR	NR	Pooled HR for PFS with positive PET 5.1 (95% CI 3.7 to 7.2) PPV range: 89% to 94%	PET-based response criteria are more predictive of progression free survival than CT-based response criteria
Mixed HL & NHL									
Data from new SRs									
Pre- and post-SCT	Wang (2016)[156] AMSTAR score: 9/11	PFS: 16 studies (3 prospective, 13 retrospective); N=1,132; Newcastle-Ottawa tool for appraisal of included studies; overall	NR	Pre-SCT HL and NHL (various types)	PFS	NR	NR	PET/CT positive vs. PET/CT negative Pre-SCT Pooled HR: 2.32 (95% CI 1.52 to 3.12)	PET scan is a useful prognostic tool before SCT in HL but not NHL; patients who had a negative

Timing of PET	Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
		quality of included studies considered to be high						HL pooled HR (based on 4 studies): 3.28 (95% CI 1.61 to 4.95) NHL pooled HR (based on 7 studies): 2.00 (95% CI 0.71 to 3.29) Post-SCT Pooled HR (based on 5 studies): 4.61 (95% CI 1.59 to 7.64)	PET scan had a better prognosis than patients who had a positive PET scan (both PFS and OS)
		OS: 12 studies (2 prospective, 10 retrospective); N=932; Newcastle-Ottawa tool for appraisal of included studies; overall quality of included studies considered to be high	NR	Pre-SCT HL and NHL (various types)	PFS	NR	NR	PET/CT positive vs. PET/CT negative Pooled HR: 2.64 (95% CI 1.55 to 3.72)	
SRs from 2011 report									
Pre-HDCT with SCT	Terasawa (2010)[146] Reported in previous HTA	12 studies (3 prospective, 9 retrospective); N=630; QUADAS tool for appraisal of included studies	Clinical follow-up	Relapsed/refractory HL and aNHL	Treatment failure	Range: 32% to 100% Pooled: 69% (95% CI 56% to 81%)	Range: 48% to 100% Pooled: 81% (95%	LR+: 3.6 LR-: 0.38 PET/CT positive vs. PET/CT negative	FDG-PET before high dose chemotherapy with SCT has good accuracy

Timing of PET	Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
							CI 73% to 87%)	PFS (based on 11 studies): 4.3 (95% CI 3.1 to 6.0)	for predicting progression and relapse during the first 2 years following the completion of therapy for patients with lymphoma

PET: positron emission tomography; HL: Hodgkin's lymphoma; CT: computed tomography; FDG: fluorodeoxyglucose; AMSTAR: Assessing the Methodological Quality of Systematic Reviews; PFS: progression free survival; OS: overall survival; QUIPS: Quality in Prognosis Studies; SCT: stem cell transplant; HR: hazard ratio; NHL: non-Hodgkin's lymphoma; iNHL: indolent non-Hodgkin's lymphoma; aNHL: aggressive non-Hodgkin's lymphoma; HDCT: high dose chemotherapy; QUADAS: Quality Assessment of Diagnostic Accuracy Studies; LR+: positive likelihood ratio; LR-: negative likelihood ratio; HTA: health technology assessment; MSAC: Medical Service Advisory Committee; REMARK: reporting recommendations for tumor marker prognostic studies

Table 18. Summary of diagnostic accuracy information: pediatric lymphoma

Timing of PET	Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Primary Conclusions
Classic HL									
Data from new individual studies									
Initial staging	Bakhshi (2017)[30] ROB: moderately low	Prospective cohort; N=57	Clinical follow-up	Pediatric HL	Positive sites	NR	NR	PET/CT vs. CT Number of detected sites: 219 vs. 152	Initial staging with PET/CT may affect treatment decision; further studies are needed
Monitoring during treatment	Bakhshi (2017)[30] ROB: moderately low	Prospective cohort; N=57	Clinical follow-up	Pediatric HL	Relapse: 4/57 (57%)	Revised International Working Group (based on 56 patients) 25% (95% CI 13.1% to 78%) Deauville Criteria (based on 52 patients) 0% (95% CI 0% to 60.4%)	Revised International Working Group (based on 56 patients) 61.5% (95% CI 47% to 74.3%) Deauville Criteria (based on 52 patients) 91.4% (95% CI 78.7% to 97.2%)	Revised International Working Group PPV: 4.7% (95% CI 0.2% to 25.8%) NPV: 91.4% (95% CI 0.2% to 25.8%) Deauville Criteria: PPV: 0% (95% CI 0% to 60.4%) NPV: 91.4% (95% CI 78.7% to 97.2%) PFS Positive scan: 93.3±4.1% Negative scan: 89.6±3.8%	There is not enough evidence to support escalation therapy based on interim PET in pediatric HL
End of treatment	Bakhshi (2017)[30] ROB: moderately low	Prospective cohort; N=55	Clinical follow-up	Pediatric HL	Relapse: 4/57 (57%)	Revised International Working Group (based on 54 patients) 25% (95% CI 13.1% to 78%)	Revised International Working Group (based on 56 patients) 88% (95% CI 74.9% to 95%)	Revised International Working Group PPV: 14.2% (95% CI 0.7% to 57.9.8%) NPV: 93.6% (95% CI 81.4% to 98.3%) Deauville Criteria:	Post-treatment PET/CT using the Deauville interpretation method predicts overall survival and has better

Timing of PET	Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Primary Conclusions
						Deauville Criteria (based on 51 patients) 25% (95% CI 13.1% to 78%)	Deauville Criteria (based on 52 patients) 95.7% (95% CI 84.2% to 99.2%)	PPV: 33.3% (95% CI 1.7% to 87.4%) NPV: 94.1% (95% CI 82.7% to 98.4%)	specificity than conventional imaging
Aggressive NHL									
Data from new individual studies									
Initial staging	Bakhshi (2012)[31] ROB: moderately low	Prospective cohort; N=34	Clinical follow-up	Pediatric nonlymphoblastic NHL	Positive sites	NR	NR	Number of detected sites: 130 vs. 114	PET/CT may be better than CT for initial staging
Monitoring during treatment	Bakhshi (2012)[31] ROB: moderately low	Prospective cohort; N=31	Clinical follow-up	Pediatric nonlymphoblastic NHL	PFS and OS	PET/CT vs. CT 7/9 (77.8%) (95% CI 44.9% to 95.9%) vs. 7/9 (77.8%) (95% CI 47.8% to 94.2%)	PET/CT vs. CT 12/22 (54.5%) (95% CI 41.1% to 62.0%) vs. 10/22 (45.5%) (95% CI 32.2% to 52.9%	PET/CT vs. CT PPV: 7/17 (41.2%) [95% CI 23.8% to 50.8%) vs. 7/19 (36.8%) (95% CI 21.5% to 45.4%) NPV: 12/14 (85.7%) (95% CI 64.6,% to 97.4%) vs. 10/12 (83.3%) (95% CI 59.0% to 96.9%) PET/CT complete metabolic remission vs. no complete metabolic remission PFS: 85.7% (95% CI 53.9% to 96.2%) vs. 57.0% (95% CI 30.1% to 76.9%)	Interim PET/CT findings cannot be used to prognosticate for PFS nor OS

Timing of PET	Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Primary Conclusions
								OS: 78.6% (95% CI 47.2% to 92.5%) vs. 50.4% (95% CI 24.9% to 71.4%)	
End of treatment	Bakhshi (2012)[31] ROB: moderately low	Prospective cohort; N=28	Clinical follow-up	Pediatric nonlymphoblastic NHL	PFS and OS	PET/CT vs. CT 3/4 (75.0%) (95% CI 23.5% to 98.7%) vs. 6/7 (85.7%) (95% CI 46.3% to 99.2%)	PET/CT vs. CT 18/24 (75.0%) (95% CI 66.4% to 78.9%) vs. 18/24 (75.0%) (95% CI 63.5% to 78.9%)	PET/CT vs. CT PPV: 3/9 (33.3%) %95% CI 10.4% to 43.9%) vs. 6/12 (50.0%) (95% CI 27.0% to 57.9%) NPV: 18/19 (94.7%) (95% CI 83.9% to 99.9% vs. 18/19 (94.7%) (95% CI 80.2% to 99.7%) PET/CT complete metabolic remission vs. no complete metabolic remission PFS: 94.7% (95% CI 68.1% to 99.2%) vs. 66.6% (95% CI 28.2% to 87.8%) OS: 84.2% (95% CI 58.6% to 94.6%) vs. 55.6% (95% CI 20.4% to 80.4%)	Post-treatment PET/CT findings can be used to prognosticate for PFS but not for OS
Mixed HL & NHL									
Data from SRs									
Initial staging	Uslu (2015)[153]	5 studies (3 prospective, 2	Conventional imaging modalities	Pediatric HL & NHL	NR	Range: 96% to 99%	Range: 95% to 100%	NR	NR

Timing of PET	Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Primary Conclusions
	AMSTAR score: 4/11	retrospective); N= 260; no appraisal of individual studies							
Monitoring during treatment	Uslu (2015)[153] AMSTAR score: 4/11	5 studies (3 prospective, 2 retrospective); N = 288; no appraisal of individual studies	CT scans, therapeutic response	Pediatric HL & NHL	Relapse, PFS	Range: 77.8% to 100%	Range: 54.5% to 97.7%	NR	More research is needed to determine optimal timing for interim PET scans to assess treatment response for pediatric lymphomas

PET: positron emission tomography; HL: Hodgkin's lymphoma; CT: computed tomography; FDG: fluorodeoxyglucose; AMSTAR: Assessing the Methodological Quality of Systematic Reviews; ROB: risk of bias; PFS: progression free survival; OS: overall survival; NHL: non-Hodgkin's lymphoma

Table 19. Summary of diagnostic accuracy information: lymphoma with bone marrow involvement

Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
HL								
Cheng (2013)[47]	6 studies (1 prospective, 5 retrospective); N=687; no appraisal of individual studies	Bone marrow biopsy	HL	18.5% (127/687)	Pooled 94.5% (95% CI 89% to 97.8%)	Pooled 99.5% (95% CI 98.4% to 99.0%)	DOR 1,591 (95% CI 421 to 6,010) AUC 0.996 Pooled LR_ 79.65 (95% CI 35.88 to 176.83) Pooled LR- 0.06 (95% CI 0.02 to 0.21)	FDG PET is better than iliac bone marrow biopsy at detecting bone marrow involvement during initial staging of patients with HL
Aggressive NHL								
Adams (2014)[16]	7 studies (2 prospective, 3 retrospective, 2 NR); N=654; QUADAS-2 tool for appraisal of included studies; overall quality of studies considered moderate	Bone marrow biopsy	DLBCL	Weighted summary proportion among BMB positive patients: 3.1% (95% CI 1.8% to 5.0%) Weighted summary proportion among BMB negative patients: 3.1% (95% CI 1.8% to 5.0%)	Pooled 88.7% (95% CI 82.5% to 93.3%) Range 70.8% to 95.8% Focal uptake only (based on 4 studies) Pooled 78.4% (95% CI 69.9% to 85.5%) Both focal and diffuse uptake (based on 4 studies) Pooled 87.0% (95% CI 79.2% to 92.7%)	Pooled 98.8 (95% CI 82.5 to 93.3) Range 98.8% to 100% Focal uptake only (based on 4 studies) Pooled 99.7% (95% CI 98.3% to 100%) Both focal and diffuse uptake (based on 4 studies) Pooled 100% (95% CI 98.82% to 100%)	AUC 0.9983	FDG PET/CT is accurate for detecting bone marrow involvement in patients with newly diagnosed DLBCL; negative FDG PET/CT findings with positive bone marrow biopsy findings may not have treatment consequences

Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
Chen (2011)[46]	7 studies; N=134 PET; N=237 PET/CT; N=321 PET or PET/CT; criteria list recommended by the Cochrane Methods Working Group on Systematic Review of Screening and Diagnostic test for appraisal of included studies	Bone marrow biopsy	Aggressive NHL	PET 37% (50/134) PET/CT 38% (91/237) PET or PET/CT 28% (91/321)	PET pooled 0.74 (95% CI 0.62 to 0.86) PET/CT pooled 0.74 (95% CI 0.65 to 0.83) PET or PET/CT pooled 0.74 (95% CI 0.65 to 0.83)	PET pooled 0.92 (95% CI 0.86 to 0.98) PET/CT pooled 0.80 (95% CI 0.74 to 0.87) PET or PET/CT pooled 0.84 (95% CI 0.80 to 0.89)	Accuracy PET 0.85 (95% CI 0.79 to 0.91) PET/CT 0.78 (95% CI 0.72 to 0.83) PET or PET/CT 0.81 (95% CI 0.77 to 0.86)	FDG PET or PET/CT have high diagnostic accuracy in patients with aggressive NHL
Indolent NHL								
Chen (2011)[46]	4 studies; N=156; criteria list recommended by the Cochrane Methods Working Group on Systematic Review of Screening and Diagnostic test for appraisal of included studies	Bone marrow biopsy	Indolent NHL	37% (57/156)	Pooled 0.46 (95% CI 0.33 to 0.59)	Pooled 0.93 (95% CI 0.88 to 0.98)	Accuracy 0.76 (95% CI 0.69 to 0.82)	The sensitivity of FDG PET or PET/CT in patients with indolent NHL is low; in cases of negative FDG findings, a bone marrow biopsy should still be performed
Mixed HL and NHL								
Wu (2012)[160]	PET/CT 5 studies (1 prospective, 4 retrospective); N=383; QUADAS tool for appraisal of included studies	NR	HL and NHL	NR	Pooled overall: 0.916 (95% CI 0.851 to 0.959) Pooled prospective: 0.583 (95% CI 0.366 to 0.779) Pooled retrospective:	Pooled overall: 0.903 (95% CI 0.859 to 0.937) Pooled prospective: 0.952 (95% CI 0.903 to 0.980) Pooled retrospective:	**DOR** Pooled overall: 68.891 (95% CI 15.877 to 298.92) Pooled prospective: 39.767 (95% CI 2.614 to 605.05)	PET/CT is highly sensitive and specific for detecting bone marrow involvement in lymphoma

Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
					1.000 (95% CI 0.962 to 1.000)	0.835 (95% CI 0.749 to 0.901)	Pooled retrospective: 157.42 (95% CI 26.776 to 925.44) ***Q index** Pooled overall: 0.8912 Pooled prospective: 0.5 Pooled retrospective: 0.8840	
	PET 20 studies (9 prospective, 11 retrospective); N=1,171; QUADAS tool for appraisal of included studies	NR	HL and NHL	NR	Pooled overall: 0.815 (95% CI 0.773 to 0.853) Pooled prospective: 0.791 (95% CI 0.690 to 0.871) Pooled retrospective: 0.905 (95% CI 0.857 to 0.941)	Pooled overall: 0.873 (95% CI 0.849 to 0.895) Pooled prospective: 0.961 (95% CI 0.933 to 0.980) Pooled retrospective: 0.710 (95% CI 0.654 to 0.762)	**DOR** Pooled overall: 49.656 (95% CI 18.667 to 132.09) Pooled prospective: 89.213 (95% CI 11.629 to 684.41) Pooled retrospective: 22.545 (95% CI 6.680 to 76.088) ***Q index** Pooled overall: 0.8813 Pooled prospective: 0.9041 Pooled retrospective: 0.8175	

Author (year)	Evidence Base Available	Reference Standard	Lymphoma Type	Prevalence (outcome)	Sensitivity	Specificity	Other Outcomes	Authors' Primary Conclusions
SRs from 2011 report								
Pakos (2005)[114] (included in MSAC 2010) Reported in previous HTA	13 studies (7 prospective, 4 retrospective, 2 NR); N=587; no appraisal of individual studies	Bone marrow biopsy	HL and NHL	53/587=18%	51%	91%	TP: 53 FN: 50 TN: 449 FP: 35	FDG-PET has moderately good concordance with bone marrow biopsy for detection of bone marrow infiltration

PET: positron emission tomography; HL: Hodgkin's lymphoma; CT: computed tomography; FDG: fluorodeoxyglucose; SCT: stem cell transplant; HR: hazard ratio; NHL: non-Hodgkin's lymphoma; DLBCL; diffuse large B-cell lymphoma; QUADAS: Quality Assessment of Diagnostic Accuracy Studies; LR+: positive likelihood ratio; LR-: negative likelihood ratio; TP: true positive; TN: true negative: FN: false negative; FP: false positive; AUC: area under the curve; DOR: diagnostic odds ratio

4.2. Key Questions: Number of Studies Retained and Overall Quality of Studies

Overall 58 studies (across 62 publications) met our inclusion criteria. For efficacy and effectiveness, 18 nonrandomized observational studies provided data for initial staging with PET/CT[12,27,29,31,35,40,42,43,48,69,71,73,82,107,115,119,132,154]; only four of these studies were in pediatric populations[31,48,119,132] one of which provided data separately for children with HL and NHL.[48] For interim PET evaluation, a total of 26 studies (across 28 publications) were identified and included eight RCTs (10 publications) of PET-adapted therapy[26,37,38,41,63,77,94,101,126,128]; 15 observational studies of PET-adapted therapy[39,58,60,78,95,109-111,116,124,133,139,140,143,169]; and three clinical nonrandomized observational studies.[31,123,150] With the exception of one clinical observational studies in children,[31] all studies evaluating interim PET were in adult populations. For end-of-treament PET evaluation, a total of six studies (across 7 publications) were identified and included two RCTs of PET-adapted therapy[63,77]; two nonrandomized observational studies (3 publications) of PET-adapted therapy[65,66,168]; and two clinical observational studies.[31,138] With the exception of one clinical observational studies in children,[31] all studies evaluating end-of treatment PET were in adult populations. For surveillance imaging with PET, a total of eight studies, one RCT[120] and seven nonrandomized observational studies (8 publications) provided data.[45,59,67,90,118,121,144,145]

For safety, a total of nine studies met inclusion criteria: one RCT,[120] seven observational, [59,83,90,116,118,121,130] and one case series.[56] Of the studies that provided data for efficacy or effectiveness, only three comparative observational studies reported safety outcomes.[59,116,121]

Two full economic evaluations were identified and included one cost-effectiveness study[43] and one cost utility study.[93]

Table 20 below provides an overview of included studies by timing of PET and type of lymphoma.

Table 20. Overview of included studies

Timing of PET and Type of Lymphoma		Studies (adults unless otherwise indicated)*
EFFICACY/EFFECTIVENESS		
INITIAL STAGING (18 studies)		**All Clinical Observational:**
Early-Stage HL		1 study[69]
Mixed-stage HL (Early and Advanced)		7 studies (4 adults, 3 pediatric)[27,35,43,48,82,119,132]
aNHL		2 studies[42,73]
iNHL		5 studies (4 adults, 1 pediatric)[12,31,40,71,107]
HL and NHL (Mixed population)		2 studies[29,154]
aNHL and iNHL (Mixed population)		2 studies[48,115]
INTERIM (26 studies, 28 publications)		
Early-Stage HL	PET-adapted trials - RCTs	2 RCTs (3 publications)[26,126,128]
	PET-adapted trials- Observational	2 studies[58,140]

Timing of PET and Type of Lymphoma		Studies (adults unless otherwise indicated)*
	Clinical Observational	1 study[123]
Advanced-Stage HL	PET-adapted trials - RCTs	3 RCTs (4 publications)[37,38,77,94]
	PET-adapted trials- Observational	8 studies[39,58,60,78,95,124,133,169]
Relapsed/Refractory HL	PET-adapted trials- Observational	3 studies[109-111]
aNHL	PET-adapted trials – RCTs	3 RCTs[41,63,101]
	PET-adapted trials – Observational	3 studies[116,139,143]
HL and NHL (Mixed population)	Clinical Observational	1 study[150]
iNHL	Clinical Observational	1 study (pediatric)[31]
END-OF-TREATMENT (6 studies, 7 publications)		
Advanced-Stage HL	PET-adapted trials – RCTs	1 RCT[77]
	PET-adapted trials – Observational	1 study (2 publications)[65,66]
aNHL	PET-adapted trials – RCT	1 RCT[63]
	PET-adapted trials – Observational	1 study[168]
iNHL	Clinical Observational	1 study (pediatric)[31]
aNHL and iNHL (Mixed population)	Clinical Observational	1 study[138]
SURVEILLANCE (8 studies, 9 publications)		**(All Clinical Observational)**
Early-Stage HL		1 study[118]
Advanced-Stage HL		1 RCT[120]
Mixed-Stage HL (Early and Advanced)		2 studies[59,121]
aNHL		2 studies[67,90]
Transformed NHL		1 study[45]
aNHL and iNHL (Mixed population)		1 study (2 publications)[144,145]
SAFETY		
		9 studies (1 RCT, 7 observational, 1 case series) [56,59,83,90,116,118,120,121,130]
COST-EFFECTIVENESS		
		2 studies (1 cost-effectiveness, 1 cost utility)[43,93]

aNHL = aggressive non-Hodgkin lymphoma; HL = Hodgkin lymphoma; iNHL = indolent non-Hodgkin lymphoma; NHL = non-Hodgkin lymphoma; RCT = randomized controlled trial.
*Some studies provided data for more than one type of lymphoma and/or timing of PET.

With regard to the overall quality of retained studies, the PET-adapted RCTs were primarily low to moderately low risk of bias and the PET-adapted observational studies were considered primarily moderately high risk of bias. The clinical observational studies were all moderately high to high risk of

bias. For the two economic studies, one was considered to be poor quality and the other moderate quality. Detailed descriptions of study quality are provided in Appendix E.

4.3. Key Question 1: Efficacy and Effectiveness

The number of studies retained and results regarding efficacy and effectiveness are provided below.

Primary outcomes for this report were OS and PFS and treatment related toxicity. Not all studies reported on all outcomes. No studies reported on health-related quality of life.

In the past decade ^{18}F-FDG PET/CT has become an integral part of clinical management of lymphoma. In evaluating evidence included in this report, a number of factors should be kept in mind. Such factors include heterogeneity in criteria for PET/CT interpretation across (and sometimes within) studies, the spatial resolution of PET and the frequency of false-positive results.

For initial staging, end of treatment and surveillance, data are limited to observational studies. For therapeutic decision making, studies of PET-adapted treatment were available. No trials compared treatment strategies based on use of PET with strategies that did not use PET.

There was substantial heterogeneity of patients within and across studies and the impact of false-positives, false-negatives and other limitations of PET are unclear. The consistency and generalizability of these findings for other treatment regimens, different clinical settings and patient populations is unknown.

4.3.1. Initial staging

Summary of results

- No studies of PET/CT-guided initial treatment based on incorporation of PET/CT results in initial staging were identified. No study provided detail on how specific changes in clinical management affected patient outcomes.
- A total of 18 observational studies at moderately high to high risk of bias compared initial staging with PET/CT or PET versus CT alone or with conventional staging in patients with lymphoma. Outcomes reported included the proportions of patients upstaged or downstaged and/or proportions of patients for whom the change in stage would impact treatment. These are considered surrogate outcomes for treatment modification and therefore indirect.
- In general across lymphoma types, more patients were upstaged versus downstaged when PET/CT was incorporated into the initial staging process. Changes in stage varied widely (8% to 50%) as did actual or theoretical change in management (6% to 70%). Studies that include mixed populations (e.g. mixed HL and NLH) are described in the full report.
- Survival outcomes based initial staging with PET/CT were reported in two studies of adults.
    - One study in patients with early-stage HL found that pretreatment PET imaging was associated with a greater probability of relapse-free survival (RFS) compared to no pretreatment PET imaging: 4-year RFS, 97% versus 67%, p=0.001.

- One study in patients with early-stage follicular lymphoma found no significant difference in overall survival (data not provided) or the probability of progression-free survival (PFS) over a median follow-up of 4.8 years when initial staging was done with PET/CT compared with conventional methods (hazard ratio [HR] for PFS for conventional vs. PET/CT: 0.87, 95% CI 0.47 to 1.62.

Studies included

A total of 18 observational studies[12,27,29,31,35,40,42,43,48,69,71,73,82,107,115,119,132,154] that met our inclusion criteria were identified that compared initial staging with PET/CT or PET versus with CT alone or with conventional staging in patients with lymphoma. Conventional staging usually included clinical examination and laboratory tests, contrast-enhanced whole-body CTs and bone marrow biopsy. Specifics of treatment changes were not provided in most studies; available details are found in Table 21. No study provided information how the change in clinical stage or management effected outcomes in those patients. These studies were all moderately high to high risk of bias, primarily due to a lack of or an inability to blind assessors to clinical data and failure to control for confounding factors (Appendix E).

Detailed results

4.3.1.1. Change in clinical stage and/or management

4.3.1.1.1. Hodgkin Lymphoma (HL)

Seven comparative observational studies[27,35,43,48,82,119,132] were identified that evaluated initial staging with PET/CT in patients with mixed early- and advanced-stage HL (Table 21). Four studies included adults (median age range 29 to 36 years)[27,35,43,82] and three studies included children (median age range 14 to 15 years).[48,119,132]

For adults, three studies compared PET/CT with CT alone and found that PET/CT resulted in a change in clinical stage in almost a third of the populations (range 32% to 34% across studies) with more patients being upstaged (range, 24% to 31%) than being downstaged (range, 3% to 8%).[35,43,82] Additionally, a change in management as a result of PET/CT staging was reported in 11% to 21% of patients across the studies. The one study comparing PET/CT with conventional staging reported a change in clinical stage in 22% patients (16% upstaged; 6% downstaged) and a subsequent modification of treatment in 6%.[27]

For children, the proportion of patients who had a change in clinical stage on the basis on PET/CT vs. CT alone in two studies was 23% and 50% with upstaging in 23% and 27% and downstaging in 0% and 23%, respectively.[48,132] One study reported that staging with PET/CT resulted in radiotherapy field volume adjustments in 70% of patients, primarily to include more sites.[132] Compared with conventional staging in one study, the addition of PET/CT only changed the clinical stage in 9% of patients (5% upstaged vs. 4% downstaged); radiation fields were changed in 21% based on PET/CT information, again, primarily to make the fields more extensive.

4.3.1.1.2. Aggressive Non-Hodgkin Lymphoma (aNHL)

Two comparative observational studies were identified that evaluated initial stating with PET/CT in adults (median age range 54 to 61 years) with aNHL (Table 21). One study compared PET/CT with CT alone in patients with peripheral T-cell lymphoma.[42] PET/CT changed the clinical stage in 5% of patients (upstaged 2%; downstaged 3%); no patient had a their treatment modified as a result, however, largely due to the use of combination chemotherapy regardless of stage in this study. In the second small study which included patients with mixed early- and advanced-stage extranodal natural killer/T-cell lymphoma, compared with conventional staging, PET/CT upstaged two patients (11%; from stage I-II to III-IV). Patients with stage III-IV disease in this study were treated with a more intensive chemotherapy regimen (compared with stage I-II) with the possibility of stem cell hematopoietic stem cell transplantation.[73]

4.3.1.1.3. Indolent Non-Hodgkin Lymphoma (iNHL)

Five comparative observational studies[12,31,40,71,107] were identified that evaluated initial staging with PET/CT in patients with mixed early- and advanced-stage iNHL (Table 21). Three studies included adults (median age range 56 to 63 years) [12,40,71,107] and included children with a median age of 10.5 years.[31]

For adults, three studies compared PET/CT with CT alone; subtypes of iNHL included follicular, marginal zone B-cell, and other lymphoma.[12,40,107] Two of the studies reported that the clinical stage changed in 29% and 42% of patients, with PET/CT upstaging the vast majority (29% and 37% compared with 0% and 5% that were downstaged).[40,107] Only one of these reported whether a subsequent change in management occurred: 39% of patients had their treatment switched to observation (as opposed to the planned radiotherapy).[107] The third study reported that staging with PET/CT (as compared with CT alone) resulted in an increased likelihood of receiving early therapy (OR 1.9; 95% CI 1.3 to 2.7) and that PET/CT was associated with a significant difference in choice of initial therapy (p=0.02). [12] The fourth study compared PET/CT with conventional staging and found no significant difference in initial treatments received following the two staging methods.[71] No information on upstaging or downstaging was provided by the latter two studies.

The one study that included children with primarily advanced-stage B-cell non-lymphoblastic lymphoma reported that 15% of children were upstaged (none were downstaged) by PET/CT compared to CT alone; however, there were no treatment modifications because no patient was upstaged from early- to advanced-stage disease (likely due to the large proportion of patients who had advanced-stage disease, 91%).[31]

4.3.1.1.4. Populations with mixed lymphoma types

Four studies included populations comprised of different types of lymphoma: mixed Hodgkin (HL) and non-Hodgkin lymphoma (NHL) in two studies [29,154] and mixed aggressive (aNHL) and indolent non-Hodgkin lymphoma (iNHL) in two studies,[48,115] (Table 21).

Mixed Hodgkin HL and NHL

The two studies that included mixed populations of both HL and NHL differed in that one included mostly younger (median age 26 years) patients with HL (66%)[29] while the other included mostly older (median age 55 years) patients with NHL (79%).[154] Both compared initial staging with PET/CT to staging with CT alone. One study (of mostly advanced-stage disease, primarily HL)[29] reported that 8% of

patients were upstaged (none were downstaged) using PET/CT; however, there were no treatment modifications because no patient was upstaged from early- to advanced-stage disease and the treatment recommendations for stage III and IV are similar. The other study in primarily NHL patients reported that PET/CT correctly staged 93% (95% CI 78%–98%) of patients[154]; in 9 patients for whom the PET/CT and CT alone disagreed, PET correctly staged 78% and CT 22% (differences in staging were resolved using bone marrow biopsy and post-treatment PET/CT studies as the standard of reference). For the two cases for which CT correctly provided another stage than PET/CT, therapy would not have been changed.

Mixed aNHL and iNHL

For the two studies that included mixed populations of both aNHL and iNHL, one enrolled adults (median age 59 years) with primarily advanced-stage disease (to include DLBCL, peripheral T-cell, follicular, marginal zone, mycosis fungoides)[115] while the other enrolled children (median age 15 years) with equal proportions of early- and advanced-stage disease (to include DLBCL and Burkitt).[48] Both compared PET/CT with CT alone for initial staging. In the study evaluating adults, 9% of patients had a change in clinical stage as a result of PET/CT with 4% upstaged and 5% downstaged; a modification of treatment occurred in 3% (no details provided). In the study evaluating children, one-third (33%) were upstaged due to PET/CT results; there was no information on treatment modification.

4.3.1.2. Survival outcomes

Two comparative observational studies that met our inclusion criteria were identified that reported survival outcomes following initial staging with PET/CT.[69,71] One study included adult patients with early-stage HL and found that patients with pretreatment PET imaging had a greater probability of relapse-free survival (RFS) compared to those without pretreatment PET imaging: 4-year RFS, 97% versus 67%, p=0.001 (i.e., they are less likely to experience disease relapse).[69] The second study in adults with early-stage follicular lymphoma (a subtype of iNHL) reported that there was no significant difference in overall survival (data not provided) or the probability of progression-free survival (PFS) over a median follow-up of 4.8 years when initial staging was done with PET/CT compared with conventional methods (hazard ratio [HR] for PFS for conventional vs. PET/CT: 0.87, 95% CI 0.47 to 1.62).[71] Overall survival was not reported by the other study.

Table 21: Initial staging: Proportion of Patients where PET result resulted in a Change in Clinical Stage and/or Change in Management

Change in clinical stage due to PET/CT						
Study, year Study design N	**pre-PET/CT lymphoma stage**	**PET/CT or PET staging compared with:**	**Total**	**Upstage**	**Downstage**	**Change in management due to PET/CT***
HL (mixed early- and late-stage)						
Adults *(median age range 29–36 years)*						
Bednaruk-Mlynski 2015 Retrospective cohort N=96	NR	CT alone	34% (33/96)	28% (27/96)	6% (6/96)	21% (20/96); treatment* extended (n=16) or shortened (n=4)
Cerci 2011 Prospective cohort N=210	early 48%; advanced 52%	CT alone	32% (67/210)	23.8% (50/210)	8.1% (17/210)	15% (32/210)
Grellier 2014† Retrospective database N=35	early 77%; advanced 23%	CT alone	34.3% (12/35)	31.4% (11/35)	2.9% (1/35)	11% (4/35); all changed from local (RT) to systemic (chemotherapy) treatment
			Range, 32%–34%	*Range, 24%–31%*	*Range, 3%–8%*	*Range, 11%–21%*
Angelopoulou 2017 Retrospective cohort N=162	early 52%; advanced 48%	Conventional‡	21.6% (35/162)	16% (26/162)	6% (9/162)	14% (23/162) (theoretical) 6% (10/162) (actually modified)
Pediatric *(mean age range 14–15 years)*						
Cheng 2013§ Retrospective cohort N=30	early 63%; advanced 37%	CT alone	23.3% (7/30)	23.3% (7/30)	0% (0/30)	NR
Robertson 2011 Prospective cohort N=30	early 47%; advanced 53%	CT alone	50% (15/30)	27% (8/30)	23% (7/30)	70% (21/30); RT field volume adjustments (32 sites added, 15 excluded)
			Range, 23%–50%	*Range, 23%–27%*	*Range, 0%–23%*	*NA*
Paulino 2012 Retrospective cohort N=53	NR	Conventional**	9.4% (5/53)	5.7% (3/53)	3.8% (2/53)	21% (11/53); RT fields more (n=9) and less (n=2) extensive
aNHL (adults, median age range 54–61 years)						
Casulo 2013 Retrospective cohort	NR	CT alone	5.2% (5/95)	2.1% (2/95)	3.2% (3/95)	0% (0/95)

Change in clinical stage due to PET/CT						
Study, year Study design N	**pre-PET/CT lymphoma stage**	**PET/CT or PET staging compared with:**	**Total**	**Upstage**	**Downstage**	**Change in management due to PET/CT***
N=95 aNHL subtype: peripheral T cell						(largely due to the use of combination chemotherapy regardless of stage, and short course chemotherapy was not performed)
Fujiwara 2011†† Retrospective cohort N=19 aNHL subtype: extranodal natural killer/ T-cell	early 42%; advanced 58%	Conventional‡	10.5% (2/19)	10.5% (2/19)	0% (0/19)	10.5% (2/19); both upstaged from I-II to III-IV; treatment for stage I-II was chemotherapy plus local irradiation and for stage III-IV was intensive chemotherapy with or without hematopoietic stem cell transplantation
iNHL						
Adults *(median age range 56–63 years)*						
Abou-Nassar 2013 Retrospective database (NCCN) N=935 (PET/CT, n=532; CT, n=421) iNHL subtype: grade 1-2 follicular	early 25%; advanced 75%	CT alone	NR	NR	NR	(PET/CT, n=532; CT only, n=421) Increased likelihood of receiving early therapy if staged with PET/CT: adjusted OR 1.87 (95% CI 1.31-2.66), p=0.0006; PET/CT associated with significant differences in choice of initial therapy; p=0.02 (more anthracycline based chemotherapy ± immunochemotherapy and

Change in clinical stage due to PET/CT						
Study, year Study design N	**pre-PET/CT lymphoma stage**	**PET/CT or PET staging compared with:**	**Total**	**Upstage**	**Downstage**	**Change in management due to PET/CT***
						less treatment with radiation only)
Carrillo-Cruz 2015 Retrospective cohort N=25 (17 had both imaging results) iNHL subtype: marginal zone B-cell	early 32%; advanced 68%	CT alone	29.4% (5/17)	29.4% (5/17)	0% (0/17)	NR
Metser 2017 Prospective registry N=197 iNHL subtypes: follicular (76%); marginal zone (11%) and other (13%)	early 88%; advanced 12%	CT alone	41.6% (82/197)	36.5% (72/197)	5.1% (10/197)	38.7% (60/155); plan changed to no treatment (from RT)
			Range, 29%–42%	*Range, 29%–37%*	*Range, 0%–5%*	*Range, NA*
Friedberg 2012 Retrospective cohort N=206 (PET/CT, n=218; conventional, n=78) iNHL subtype: marginal zone B-cell	early 100%	Conventional‡‡	NR	NR	NR	No significant difference (p=0.47) in treatments received following staging with PET/CT vs. conventional: • Watchful waiting: 15% vs. 21% • Rituximab monotherapy: 12% vs. 13% • Rituximab/chemotherapy: 25% vs. 32% • Radiation therapy: 30% vs. 23%

Change in clinical stage due to PET/CT						
Study, year Study design N	**pre-PET/CT lymphoma stage**	**PET/CT or PET staging compared with:**	**Total**	**Upstage**	**Downstage**	**Change in management due to PET/CT***
						• Combined modality with radiation therapy: 14% vs. 10% • Other: 5% vs. 1%
Pediatric *(median age 10.5 years)*						
Bakhshi 2012 Retrospective cohort N=34 iNHL subtype: non-lymphoblastic (B-cell 82%; T-cell 18%)	early 9%; advanced 91%	CT alone	14.7% (5/34)	14.7% (5/34)	0% (0/34)	0% (0/34); no changes in treatment since no patient was upstaged from early- to advanced-stage (1 patient went from stage I to II and 4 from III to IV).
HL and NHL (mixed population; adults)						
Awan 2013 Prospective cohort N=53 HL (66%) and NHL (34%), median age 26 years	NR (mostly advanced according to authors)	CT alone	7.5% (4/53)	7.5% (4/53)	0% (0/53)	0% (0/53); change in number of chemotherapy cycles
van Hamersvelt 2014 Prospective cohort N=29 HL (21%) and NHL (79%), median age 55 years	NR	CT alone	PET/CT correctly staged 27/29 patients (93%; 95% CI 78%–98%); disagreement on staging in 9 of 29 patients; of these, PET correctly staged 78% (n=7) and CT 22% (n=2)§§	NR	NR	NR; authors report that in the two cases (7%, 95% CI 2%–22%) for with CT correctly provided another stage than PET/CT, therapy would not have been changed.
aNHL and iNHL (mixed population)						
Adults *(median age 59 years)*						
Papajik 2011b	early 22%; advanced 78%	CT alone	9.4% (11/117)	4.3% (5/117)	5.1% (6/117)	2.6% (3/117); details NR

Change in clinical stage due to PET/CT						
Study, year Study design N	**pre-PET/CT lymphoma stage**	**PET/CT or PET staging compared with:**	**Total**	**Upstage**	**Downstage**	**Change in management due to PET/CT***
Prospective cohort N=117 Various subtypes to include DLBCL, peripheral T-cell, follicular, marginal zone, mycosis fungoides						
Pediatric *(median age 15 years)*						
Cheng 2013§ Retrospective cohort N=117 Various subtypes to include DLBCL, Burkitt	early 48%; advanced 52%	CT alone	33.3% (7/21)	33.3% (7/21)	0% (0/23)	NR

aNHL = aggressive non-Hodgkin lymphoma; CT = computed tomography; DLBCL = diffuse large B-cell lymphoma; HL = Hodgkin Lymphoma; iNHL = indolent non-Hodgkin lymphoma; NHL = non-Hodgkin lymphoma; NR = not report; PET = positron emission tomography.

*Treatment details are given if they were provided by the authors.

†This study included patients undergoing initial staging (n=27) and relapse staging (n=8); of those undergoing initial staging a total of 8 had a change in stage due to PET [upstaged: 25.9% (7/27); downstaged: 3.7% (1/27)] and of those undergoing relapse staging a total of 4 had a change in stage due to PET [upstaged: 50% (4/8); downstaged: 0% (0/8)].

‡Conventional staging including clinical examination, contrast-enhanced whole-body CTs and bone marrow biopsy.

§Cheng 2013 reported clinical outcomes for patients with Hodgkin and non-Hodgkin lymphoma separately.

**Conventional staging included CT and bone marrow biopsy; some had bone scan, MRI, gallium.

††Median age of the population was 61 years with a range of 13-90 years; since the median age was 61 we considered this to be an adult population (even though range went down to 13 years).

‡‡ Conventional staging included both bone marrow aspirate and biopsy and a CT scan; authors refer to this as “rigorous” staging.

§§ Differences in staging between unenhanced low-dose FDG-PET/ CT and contrast enhanced CT (CECT) were resolved using bone marrow biopsy and posttreatment FDG-PET/CT (including CECT) studies as the standard of reference

4.3.2. Interim

Summary of results

Early-Stage Hodgkin Lymphoma (HL)

PET-Adapted treatment for early-stage HL: RCTs

- No trials compared treatment strategies based on use of PET with strategies that did not use PET. No studies in children met the inclusion criteria.
- Two randomized trials of PET-adapted therapy in persons with early-stage HL (stage I, IA, II, IIA) were identified. One trial randomized patients with positive interim PET after two cycles of initial therapy (ABVD) to escalated (BEACOPP) versus standard treatment[26,128] (1-2 additional ABVD cycles plus involved node radiotherapy (INRT). The same trial randomized persons with negative interim PET to different de-escalation regimens based on favorable or unfavorable prognostic status at initial staging. The de-escalation component (omission of INRT) was stopped early and a safety amendment issued based on futility analyses for this non-inferiority trial. Better PFS in those in the standard treatment group (those with ABVD + INRT) vs. those receiving the ABVD only led to closing the arms which omitted INRT. The second trial randomized persons with a negative interim PET to no further treatment versus one additional cycle of ABVD and INRT.[126]
- Overall, in persons with early-stage Hodgkin Lymphoma, interim PET may be of most value for identifying those who may benefit from treatment escalation based on positive interim PET findings in one trial. For results across the two trials evaluating treatment de-escalation (elimination of consolidative RT) in persons with negative interim PET the impact of PET-directed treatment is less clear.
    - **Escalated therapy in PET positive groups**: In the one trial[26,128] evaluating escalation of treatment versus standard treatment in persons with a positive interim PET, the probability of both OS and PFS was greater in those who had the escalated therapy. There was, however, substantially greater treatment-related toxicity reported in those with escalated therapy. (SOE MODERATE for all outcomes)
    - **De-escalation of therapy in PET negative groups:**
        - Results across the two trials that randomized persons with a negative interim[26,126,128] PET were generally consistent for OS and PFS; the overall benefit of PET-adapted treatment is less clear compared with escalation strategies described above.
        - At 3 years the difference in OS was small between those who received no further treatment vs. those who had INRT in the RAPID trial. In the H10 trial at 5 years, the differences in OS between groups randomized to de-escalated therapy omitting INRT and standard therapy with it were small regardless of initial staging as favorable (0.4/100) or unfavorable (1.6/100) and were not statistically different. Across the favorable and unfavorable groups the difference between de-escalated (99%) and standard treatment with INRT (98%) was estimated to be small, 1/100 and of unknown clinical significance (SOE MODERATE)

- PFS in both trials was less common in groups receiving de-escalated therapy where INRT was omitted. In the RAPID trial an absolute difference of -3.8% at 3 years was seen but was not statistically significant. Similarly in the H10 trial PFS was lower in those with an initially favorable prognosis (difference -11.9/100) and for those with an initially favorable prognosis (difference - 2.5/100). Across prognostic groups the overall difference was -7/100. Based on PFS, results suggest that exclusion of INRT may have a clinically important impact, particularly in those with an initially favorable prognosis and should be retained in the treatment regimen. (SOE MODERATE)
- Evidence on treatment-related toxicity was sparse and no clear conclusions regarding the impact of interim PET-adapted treatment to reduce such toxicities are possible. Follow-up time may have been insufficient to capture some RT-related toxicities. Power to detect rare events and differences between strategies was limited. (SOE LOW)

▪ Difference in therapeutic regimens, study design, patient populations and length of follow-up may contribute to differences across trials.

- Heterogeneity of patients within and across studies is likely and the impact of false-positives, false-negatives and other limitations of PET are unclear. The consistency and generalizability of these findings for other treatment regimens, different clinical settings and patient populations is unknown.
- Data through 2014 from the National Cancer Institute's SEER database indicate that 5-year relative survival for stages I and II HL are 92.3% and 93.4% respectively.

PET-Adapted Treatment for early-stage HL: Observational studies (SOE not assessed)

- Across two observational studies[58,140] in which patients with HL stages I or II (without bulky tumor) with positive interim PET received escalated treatment and those with negative interim PET received de-escalated treatment after two ABVD cycles, PFS was substantially worse in PET-positive individuals. Prognosis in terms of OS, was also slightly worse. It is unclear how much improvement in survival might be expected from the escalated therapies and the ability of PET to guide effective treatment are unclear from the data provided.

Other observational studies evaluating interim PET for early-stage HL (SOE not assessed)

- In one observational study of primarily adults with early-stage HL, no significant difference was seen in the proportion of treatment plan modifications for patients who had a pre-treatment PET/CT compared with those who did not.

Advanced-Stage Hodgkin Lymphoma (HL)

PET-Adapted treatment for advanced-stage HL: RCTs (SOE MODERATE for all outcomes)

- No trials compared treatment strategies based on use of PET with strategies that did not use PET. All trials were in adult populations
- A total of three randomized PET-adapted treatment trials in persons with advanced-stage HL were identified. Randomization in the trials was to treatment regimens based on interim PET results. Two trials compared randomized escalated treatment regimens in persons with positive interim PET results.[37,38,77] One of these also compared persons with negative interim PET results randomized to de-escalation of treatment versus continuation of an escalated treatment.[37,38] The third trial randomized those with PET negative results to de-escalated therapy[94] versus continuation of standard treatment. There is substantial heterogeneity across trials with regard to initial treatments, PET-adapted treatment regimens, patient populations and differences in criteria for determining positivity of PET results.
- Overall, in persons with advanced-stage HL, interim PET appears to be of most value for the identification of persons who may be candidates for de-escalation of therapy based on negative PET scan results. Limited data across two trials suggest that the probability of OS and PFS in those receiving scaled back treatment similar to those continuing more intensive treatment while many toxicities related to treatment were substantially reduced in those whose treatment was de-escalated.
- On the other hand, the value of PET to direct to escalation treatments when interim PET results are positive is less clear in the identified trials. Across trials of two different types of escalated therapy estimated differences per 100 patients between treatments for OS or PFS were small at 3 years and in one trial at 5 years and not beyond what may be expected by chance. Toxicities based on treatments received were also generally similar between treatment groups. The role of false positives and other limitations of PET are unclear.
- With regard to prognosis based on interim PET results alone, in one trial no difference in survival for PET positive and negative groups was observed. This may be interpreted in several ways. One interpretation is that PET may not adequately discriminate between individuals who would benefit most from escalated therapy perhaps due to population heterogeneity and/or limitations of PET. Another explanation may be that the PET + groups that receive escalated therapy in fact had better survival following escalated therapy versus survival had they not received it. There are differences in patient characteristics and disease severity between PET-positive and PET-negative groups.
- The consistency and generalizability of these findings to other treatment regimens, clinical settings and patient population is unknown. Each trial used different induction therapies. PET positivity did not appear to identify high risk patients who may benefit most from escalate treatment in the HD18 trial.
- Data from 2008 through 2014 from the National Cancer Institute's SEER database indicate that 5-year relative survival for stage II is 93%, stage III is 83% and stage IV HL is 73%.[2] Across stages, 5-year relative survival is about 87% and 10 year survival 78%. Death rates have been falling on average 2.8% each year over 2006-2015. Authors of included trials indicate that standard ABVD

treatment for advanced-stage classic HL results in a 3-5 year PFS that ranges from 61% to 76% with cure rates of 70% to -80%, and that 5-year freedom from treatment failure may be achieved in up to 90% of patients receiving more intensive regimens (i.e., escalated BEACOPP).

PET-Adapted treatment for advanced-stage HL: Observational studies (SOE not assessed)

- The benefit of interim PET across eight observational studies reporting PET-adapted treatment is unclear. Patient characteristics, treatments and methods across studies varied substantially making conclusions across studies difficult. Detailed data on the impact of specific PET-adapted treatments were limited, including information on impact of escalated or de-escalated treatment on survival or treatment-related toxicity. Studies were at moderately high or high risk of bias.
- In all but two studies,[39,60] prognosis in terms of OS or PFS at various time frames was worse in those with a positive interim PET. Only a limited number of patients with positive interim PET were available in the two studies in which prognosis was similar for those with a positive versus a negative interim PET.

Relapsed or Refractory Hodgkin Lymphoma (HL)

Pre-Autologous Stem Cell Transplantation (ASCT) PET (SOE INSUFFICIENT)

- No RCTs comparing treatment strategies based on use of PET with strategies that did not use PET were identified and no randomized PET-adapted studies were identified.
- Three observational studies by the same author,[109-111] at moderately high risk of bias, used PET results to allocate PET negative patients directly to high dose radiochemotherapy (HDT) and ASCT versus using extended salvage chemotherapy in those with positive PET were identified. Limited information from these studies suggests that pre-ASCT PET may facilitate identification of those who may be able to progress to HDT/ASCT if results are negative from those who may require extended savage therapy prior to HDT/ASCT; pre-transplant PET results may provide valuable prognostic information as achievement of PET negativity prior to HDT/ASCT have better probability of survival.
    - The 2012 study reports that EFS for persons with pre-HDT/ASCT negative FDG-PET, after 1 or 2 cycles of salvage therapy was > 80%, versus 28.6% for patient who had a positive scan ($P < .001$); no data comparing the impact of PET-adapted strategies on toxicity were provided.
    - EFS at 3 years for those with pre-ASCT PET positive results was 60%, significantly less compared with those having a with pre-ASCT PET negative result ($p = 0.05$) in persons who had ASCT based on combined data from the 2015 and 2017 publications.[109,110] Across patients, 76 % (95% CI 62–89) achieved PET-negative status at the end of all treatments.
- The reports do not provide adequate data to evaluate the impact of PET-adapted therapy in terms of OS, PFS and toxicities relative to specific PET-adapted therapies.

- A recent meta-analysis of ASCT in patients with HL reported pooled estimates at 2 years for OS (58% and 50%) and for relapse free survival (37% and 31% respectively).[129]

Aggressive Non-Hodgkin Lymphoma (aNHL): Diffuse Large B-cell Lymphoma (DLBCL)

PET-Adapted treatment for aNHL (DLBCL): RCTs

- No trials randomizing patients to PET/CT versus non-PET/CT treatment strategies were identified in persons with aggressive NHL. One RCT of PET-adapted escalation of treatment for those with positive and negative interim PET and one RCT of induction treatments in patients with stage III or IV NLH were identified. A third trial in early stage (I or II with no bulky mass) low risk patients was also identified. Conclusions across these trials regarding the value of interim PET differ, due to differences in patient populations, study designs and effectiveness of treatment escalation strategies.
- The PETAL trial[63] randomized patients with aNHL to treatments based on interim PET/CT findings. PET/CT was done after two cycles of CHOP or R-CHOP (rituximab use was restricted to patients with CD20 positive lymphomas), persons with positive interim PET were randomly assigned to receive six blocks of an intensive Burkitt's lymphoma treatment protocol or six additional cycles of R-CHOP. Individuals with a negative interim PET with CD20 positive lymphomas were randomly assigned to receive four additional cycles of R-CHOP or the same treatment with two additional doses rituximab. Recruitment was stopped early because PET-positivity was rare and the Burkitt's protocol was more toxic and possibly less effective than R-CHOP. (SOE Moderate for all outcomes)
    - In persons with a positive interim PET, OS, PFS and EFS were lower in patients who received the Burkitt's protocol versus those receiving additional RCHOP; the differences were within what might be expected by chance. Higher risk of treatment-related toxicities was seen in those receiving the Burkitt's protocol.
    - In persons with a negative interim PET, additional doses of rituximab did not appear to substantially improve OS, PFS or ES. Risks for treatment-related adverse events did not consistently favor one or the other treatment and differences may not have been beyond what might be expected by chance.
    - The prognosis for those with a positive interim PET was substantially worse than for those with a negative interim PET.
    - Based on this trial, the value of interim PET to identify those who may benefit most from escalated therapy from this trial is unclear.
- The LNH-3B trial randomized high-risk patients with CD20 positive DLBCL to different induction treatments and used results of interim PET after 2 and 4 cycles of therapy to adapt additional treatment[41]. Individuals with negative PET results at both time frames, received standard immunochemotherapy (SIC) consolidation, where as those with positive PET after cycle 2 received escalated therapy including autologous stem cell transplantation. Treatment of participants with positive PET results after 4 cycles of therapy was at the discretion of the local investigator. Based on this trial, interim PET positive results may facilitate identification of the subgroup of patients for whom escalation of treatment may improve PFS. This should be

interpreted with caution as it assumes that the prognosis of PET-positive patients would be worse in the absence of escalated treatment. No data on the impact of these strategies on PET-adapted treatment-related toxicities were reported, making it difficult to draw firm conclusion regarding the overall value of interim PET in aNHL. (SOE low for survival outcomes, insuffiicent for toxicity)

  - At 4 years, the probability of OS was similar between those who received escalated therapy and those who had standard immunochemotherapy (SIC) consolidation (difference 0.8/100), however PFS was better in those receiving escalated therapy (difference 10/100).
  - Information on treatment-related toxicities by PET-adapted treatments was not provided.

- The LYSA/GOELAMS trial randomized patients at initial staging to receive RT or not after 4-6 cycles of R-CHOP.[101] Patients with interim PET negative results (those in complete remission) patients continued with the treatments as randomized. OS at 5 years was similar between those who did and those who did not receive RT (90% vs. 92%) as was the frequency of relapse (RD 1%) suggesting that RT might be omitted in these patients. (SOE Moderate).
- Specific to DLBCL, clinical guidelines suggest that 3-year overall survival may range from 59% for high risk patients to 91% for low risk patients[151] and has improved in the last decade due to the addition of rituximab.[134] Untreated, median survival is estimated at less than 1 year.[134] Data through 2014 from the National Cancer Institute's SEER database indicate that 5-year relative survival across types of NHL is around 71%[112] and may be as high as 82% in those with stage I or as low as 62% in those with stage IV%.
- The consistency across and generalizability of the findings from the trials to other treatment regimens and patient population is unknown.

PET-Adapted treatment for aNHL (DLBCL): observational studies

- Across three observational studies[116,139,143] evaluating three different treatment escalation strategies in those with positive interim PET, probabilities of OS, PFS and EFS, were lower than for those with negative interim PET who continued with standard therapy. Probabilities for survival outcomes were generally in the same range as those reported in the RCTS. The impact of early escalated therapy in these patient populations is unclear. No study compared escalated therapy to standard therapy for interim PET+ patients and therapy regimens differed substantially across studies as did patients characteristics.
- Two studies[139,143] provided comparative data related to toxicity. The incidence of toxicity-related adverse events was greater in patients whose treatment was presumably escalated (PET+) versus not-escalated (PET–), as were treatment-related deaths.

Detailed results

The following results sections contains abbreviations for various chemotherapy regimens or other treatments for the management of lymphomas; please see Appendix M for descriptions.

4.3.2.1. Early-Stage HL

Studies included

Two trials of PET-adapted therapy in persons with early-stage HL (stage I-II) were identified (3 publications)[26,126,128]. The H10 trial (reported across two publications)[26,128] randomized patients after two cycles of ABVD therapy to interventional or standard treatment strategies based on determination of unfavorable (presence of ≥1 of the following risk factors: age ≥ 50 years, > four involved nodal areas, presence of mediastinal bulk, or erythrocyte sedimentation rate ≥ 50 mm without B symptoms or ≥ 30 mm with B symptoms) or favorable (all others) prognostic status at initial staging in addition to interim PET findings. The primary objective was to evaluate whether involved-node radiotherapy (INRT) could be omitted without loss of progression-free survival (PFS) in persons with negative interim PET results. The secondary objective was to evaluate whether therapy escalation to BEACOPP would improve PFS. Overall, the population (n=1950) was 52% female with a median age of 30.5 (range 15-70) years; 61% had unfavorable prognostic status. Those whose interim PET findings were positive received escalation treatment of BEACOPP versus standard ABVD with INRT. Patients with negative interim PET results were randomized to different de-escalation strategies. Additionally, a safety protocol amending treatment was issued towards the end of study recruitment that reduced the number of chemotherapy cycles for patients recruited afterwards. A total of 505 PET negative patients were then treated according to this safety amendment, receiving either one or two additional ABVD cycles based on favorable or unfavorable status, respectively. The second trial, the RAPID trial, is ongoing and to date has only reported results for those with negative PET findings.[126]. The negative PET group (n=420) had equal numbers of male and female patients (50%), with ages ranging between 16 and 75 (median age 34), and similarly began treatment with two cycles of ABVD. Negative interim PET patients were randomized to receive no further treatment or INRT, while the positive interim PET patients (an additional n=145) were nonrandomly allocated to receive one additional cycles of ABVD before continuing to INRT. Only results from the negative patients are available for analysis. Both trials were considered at low risk of bias. In the H10 trial, independent or blinded assessment of PET/CT was not described.

Additionally, two prospective observational studies[58,140] evaluating PET-adapted treatment in patients with early-stage HL (stage I-II without bulky tumor) were identified. Both studies included adult patients (age 18-59 years, median 30 years) with females comprising 62% and 46% of the populations. In both studies, the induction treatment was ABVD and interim PET was performed after two cycles. One study (N=170)[58] categorized early-stage HL patients into favorable (14%) and unfavorable (86%) prognosis groups, based on the presence of 1 or more risk factors (extra-nodal disease, bulky mediastinal mass >10 cm, erythrocyte sedimentation rate >50, ≥3 disease sites, and age ≥50 years). Patients with a positive interim PET result continued first-line therapy for an additional 2 cycles (favorable group) or 4 cycles (unfavorable group) followed by involved site radiotherapy (ISRT); those with a negative interim PET result were allocated based on physician discretion to ISRT or to ABVD for two additional cycles without ISRT (favorable group) or to ABVD for two additional cycles followed by ISRT or ABVD for four additional cycles without ISRT (unfavorable group). The number alterations were based on the presence of multiple or extensive disease sites or fields overlapping breast tissue (women). The second observational study (n=149)[140] allocated patients with positive interim PET results to receive two cycles of escalated BEACOPP followed by involved field (IFRT); those with negative interim PET results continued first-line therapy for an additional 2 cycles . Both studies were rated as moderately high risk

of bias; both lacked independent or blinded assessment of PET/CT and one study failed to control for confounding while the other did not appear to account for time at risk.

4.3.2.1.1. PET-adapted treatment: results from RCTs

Escalated therapy

Overall, escalated therapy in patients with positive interim PET may lead to improved OS and PFS. The H10 trial in persons with early-stage HL randomized persons with a positive interim PET (N=331) to either escalated therapy of BEACOPP and INRT or to continuation of standard therapy of ABVD with INRT.[26,128] At five years, OS was more common in those randomized to receive escalated treatment (96.0%, 95% CI 91.1%–98.2%) versus those continuing with standard treatment (89.3%, 95% CI 83.4%–93.2%) with an estimated difference of 6.7% and a corresponding reduction in overall mortality (RD -5%, 95% CI -10% to -0.1%). PFS at 5 years was substantially more common (estimated difference 12.9%) in those who had the escalated therapy (90.6%, 95% CI 84.7%–94.3%) versus standard therapy (77.4%, 95% CI 70.4%–82.9%), Table 22

Patients receiving escalated treatment had significantly higher risk of grade 3 or 4 hematologic toxicities and infections compared with those receiving standard treatment (Table 23). In particular, patients receiving escalate treatment had substantially higher risks for neutropenia (RD 23%, 95% CI 13% to 33%), febrile neutropenia (RD 23%, 95% CI 16% to 29%) and thrombocytopenia 19.7% vs. 0%; RD 20%. Grade 3 or 4 infection (without neutropenia) was also more substantially more common RD 4% (95% CI 1% to 8%) but risk of secondary malignancy was similar (2.4% vs. 2.1%). There may have not been sufficient statistical power to compare toxicity related deaths or death due to secondary malignancy in the two groups. (Tables 22, 23)

De-escalated therapy

Two trials at low risk of bias in persons with early-stage HL randomized those with a negative interim PET to de-escalation treatment strategies versus standard treatment as describe above. In the H10 trial[26,128] patients with negative interim PET (N=381) were randomized to different regimens based on favorable or unfavorable prognostic status at initial staging, but for both prognostic groups, de-escalated treatment involved omission of INRT. The RAPID trial[126] randomized persons with negative interim PET findings (N=420) to no further treatment (omission of INRT) or to INRT. A 5-PS of 1 or 2 was used to classify scans as negative in both trials (Tables 24, 25).

The probability of OS was reported in both trials with somewhat differing results. At 3 years in the RAPID trial, OS tended to be slightly greater in those who received no further treatment vs. those who had INRT (99% vs. 97.1 %) and the difference may not be beyond what might be expected by chance. At 5 years in the H10 trial, in patients with favorable prognostic status at initial staging a very small difference decrease in OS (-0.4%) between groups randomized to de-escalated therapy omitting INRT (99.6%) and standard therapy with it (100%) was seen. In persons with unfavorable prognostic status at initial staging OS was slightly more common among those receiving de-escalated treatment (98.3% vs. 96.7%, estimated difference 1.6%). Across the favorable and unfavorable groups the difference between de-escalated (99%) and standard treatment with INRT (98%) was small, 1%.

Results for PFS were somewhat more consistent across the trials and suggest that PFS may be diminished when INRT is omitted in those with a negative interim PET. In the RAPID trial[126] at 3 years base on intention-to-treat analysis, PFS was less common in the group that received no further treatment (90.8%) versus those that had INRT (94.6%, estimated difference -3.8%). Findings from the per-protocol analysis also suggest lower PFS when no further treatment is given and INRT is omitted 90.8% (95% CI, 86.8 to 94.7) versus having INRT 97.1% (95% CI, 94.7 to 99.6), with a RR comparing INRT with no INRT of 2.36 (95% CI, 1.13 to 4.95) in favor of doing INRT (P = 0.02).

In the H10 trial preliminary results at one year suggested that PFS may be compromised when INRT was omitted compared with those who continued with therapies that included INRT. In those with initially favorable prognosis, PFS was less common when INRT was omitted was 94.9% versus 100% when INRT was retained. Similarly, in those with initially unfavorable prognosis, PFS was again less common with the omission of INRT (94.7% vs. 97.3 %)[128]. As noted above, the study arm omitting INRT was closed early and a safety amendment issued to require that INRT be done. The 5 year follow-up for the H10 trial[26] across participants enrolled prior to the amendment is reported here and reflected in Table 24. At 5 years, PFS was lower in those with an initially favorable prognosis (estimated difference -11.9%) and for those with an initially favorable prognosis (difference - 2.5%). Across prognostic groups the estimated difference was -7% which may be clinically important. Authors report that preplanned sensitivity analyses including a per-protocol analysis, were performed and led to similar conclusions. Authors of both studies indicate that noninferiority of the treatment strategies could not be demonstrated.

In the observational continuation of the H10 trial in 505 patients enrolled following the safety, all patients with a negative interim PET received INRT, OS and PFS improved and did not differ from those of the standard treatment which included INRT in the randomized portion of the study. (Table 26)

A 2015 Cochrane Review evaluated data for individuals with negative interim PET results from three PET-adapted treatment trials[136]. They concluded that PFS was shorter in persons with early-stage HL and a negative interim PET scan who received the de-escalated treatment (chemotherapy) versus standard treatment (inclusion of radiotherapy). This review is summarized in the background section.

Treatment-related toxicity

Treatment-related toxicity was poorly reported in both trials of PET-adapted de-escalation of treatment (Tables 24, 25). In the RAPID trial,[126] deaths related to cardiovascular events, secondary malignancy, pneumonia, pneumonitis, cerebral hemorrhage and mycosis fundoides were reported, but it is not clear to what extent they may or may not be related to PET-adapted treatment. With the exception of death due to secondary malignancy (1.4%), all others occurred in ≤1% of the study population. The H10 trial reported that there were no unexpected toxicities were observed in those with negative interim PET, but provide no detail beyond death due to various causes (Table 25).[26,128] Overall, there were no statistical differences between de-escalated strategies and standard treatments, but studies may have lacked statistical power to detect rare events.

Table 22. Early Hodgkin (stage I-II): RCTs of PET-adapted therapy evaluating *escalation* of therapy: overall and progression-free survival and mortality

Outcome	Andre 2017/Raemaekers 2014[26,128] *EORTC/LYSA/FIL H10 Trial* Initial therapy: ABVD, 2 cycles		
	PET + (randomized) (N=361) **(IHP used, corresponding 5-PS ≥3)** Combined Favorable (n=97) and Unfavorable (n=264)*		
	Intervention: eBEACOPP 2 cycles + INRT (n=169)	*Comparator*: ABVD 1-2 additional cycles + INRT (n=192)	Effect size Estimate (95% CI)
5-year OS (95% CI)	96.0% (91.1%–98.2%)	89.3% (83.4%–93.2%)	HR 0.45 (0.19–1.07), p=0.062 (improved OS with eBEACOPP) Estimated difference 6.7%
5-year PFS (95% CI)	90.6% (84.7%–94.3%)	77.4% (70.4%–82.9%)	HR 0.42 (0.23–0.74), p=0.002 (improved PFS with eBEACOPP) Estimated difference 12.9%
			Intervention vs. comparator
Deaths†	4.1% (7/169)	9.4% (18/192)	RR 0.4 (0.2–1.0); RD -5% (-10% to -0.1%); p=0.05
Progression/relapse	1.8% (3/169)	5.7% (11/192)	RR 0.3 (0.1–1.1); RD -4% (-8% to -0.1%); p=0.05
Toxicity of protocol treatment	0.6% (1/169)	0.5% (1/192)	RR 1.1 (0.1–18.0); RD 0.1% (-1% to 2%); p=0.93
Toxicity of second-line treatment	0%	1.0% (2/192)	RR NC; RD -1%; p=0.18
Cardiovascular event	0%	0%	---
Second malignancy	0.6% (1/169)	1.0% (2/192)	RR 0.6 (0.1–6.2); RD -0.4% (-2% to 1%); p=0.64
Other/unknown	1.2% (2/169)	1.0% (2/192)	RR 1.1 (0.2–8.0); RD 0.1% (-2% to 2%); p=0.90

ABVD =doxorubicin, bleomycin, vinblastine, and dacarbazine; CI = confidence interval; eBEACOPP = escalated bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine and prednisone; HR = hazard ratio; IFRT = involved-field radiotherapy; INRT = involved-node radiotherapy; NC = not calculable; NR = not report; OS = overall survival; PET = positron emission tomography; PFS = progression-free survival; RD = risk difference; RR = relative risk.

*For evaluation of treatment escalation, interim PET(+) favorable and unfavorable patients were pooled because of their presumed common poor prognosis.

- *Favorable status* indicates age <50 years with ≤3 involved nodal areas, absence of mediastinal bulk (mediastinum-to-thorax ratio <0.35), and erythrocyte sedimentation rate (ESR) <50 mm without B symptoms or ESR <30 mm with B symptoms.
- *Unfavorable status* indicates age ≥50 years, >4 involved nodal areas, presence of mediastinal bulk (mediastinumto- thorax ratio ≥0.35), or ESR ≥50 mm without B symptoms or ESR ≥30 mm with B symptoms.

†RR/RD (95% CI) and p-value calculated by AAI.

Table 23. Early Hodgkin (stage I-II): RCTs of PET-adapted therapy evaluating *escalation* of therapy: Toxicity and secondary malignancy

Outcome	Andre 2017/Raemaekers 2014[26,128] *EORTC/LYSA/FIL H10 Trial* Initial therapy: ABVD, 2 cycles		
	PET + (randomized) (N=361)		
	Intervention: eBEACOPP 2 cycles + INRT	*Comparator*: ABVD 1-2 additional cycles + INRT	Intervention vs. Comparator RR/RD (95% CI)*
Toxicity, hematologic grade 3/4†‡			
Neutropenia	53.3% (90/169)	30.2% (58/192)	RR 1.8 (1.4–2.3); RD 23% (13% to 33%); p<0.0001
Febrile neutropenia	23.7% (40/169)	1.0% (2/192)	RR 22.7 (5.6–92.6); RD 23% (16% to 29%); p<0.0001
Thrombocytopenia	19.7% (33/169)	0%	RR NC; RD 20%; p<0.0001
Anemia	4.9% (8/169)	0%	RR NC; RD 5%; p=0.002
Death from toxicity	0.6% (1/169)	1.6% (3/192)	RR 0.4 (0.04–3.6); RD -1% (-3% to 1%); p=0.38
of protocol treatment	0.6% (1/169)	0.5% (1/192)	RR 1.1 (0.1–18.0); RD 0.1% (-1% to 2%); p=0.93
of second-line treatment	0%	1.1% (2/192)	RR NC; RD -1%; p=0.18
Infections, grade 3/4‡ (without neutropenia)	5.6% (9/169)	1.1% (2/192)	RR 5.1 (1.1–23.3); RD 4% (1% to 8%); p=0.02
Second malignancy	2.4% (4/169)	2.1% (4/192)	RR 1.1 (0.3–4.5); RD 0.3% (-3% to 3%); p=0.86
Death from secondary malignancy	0.6% (1/169)	1.0% (2/192)	RR 0.6 (0.1–6.2); RD -0.4% (-2% to 1%) p=0.64

ABVD =doxorubicin, bleomycin, vinblastine, and dacarbazine; CI = confidence interval; eBEACOPP = escalated bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine and prednisone; HR = hazard ratio; IFRT = involved-field radiotherapy; INRT = involved-node radiotherapy; NC = not calculable; NR = not report; OS = overall survival; PET = positron emission tomography; PFS = progression-free survival; RD = risk difference; RR = relative risk.

*Calculated by AAI.

†Numerators were back-calculated using percentages and denominators given.

‡Grading of Adverse Events (AE) is according to the Common Terminology Criteria for Adverse Events (CTCAE) v3.0. Grades of 1-5 refer to the severity of the AE. Grade 1 Mild AE, Grade 2 Moderate AE, Grade 3 Severe AE, Grade 4 Life-threatening or disabling AE, Grade 5 Death related to AE.

Table 24. Early Hodgkin (stage I-II): RCTs of PET-adapted therapy evaluating *de-escalation* of therapy

Outcome	Andre 2017/Raemakers 2014[26,128] *EORTC/LYSA/FIL H10 Trial* Initial therapy: ABVD, 2 cycles								
	Favorable PET – Group* (randomized) (5-PS score of 1 or 2)			PET +	PET + vs. PET –	Unfavorable PET – Group† (randomized) (5-PS score of 1 or 2)			PET +	PET + vs. PET –
	Intervention: 2 additional ABVD cycles (n=193)	Comparator: 1 ABVD cycle + INRT (n=188)	Comparator vs. Intervention	NR		Intervention: 4 additional ABVD cycles (n=268)	Comparator: 2 ABVD cycle + INRT (n=251)	Comparator vs. Intervention	NR	
1-year PFS‡	94.9% (79.6% CI, 91.9%–96.9%)§	100.0%	HR 9.36 (79.6% CI, 2.5%–35.7%), p=0.017§	NR	NR	94.70% (80.4% CI, 92.1%–96.5%)§	97.28% (80.4% CI, 95.2%–98.5%)§	HR 2.42 (80.4% CI, 1.35–4.36), p=0.026§	NR	NR
	Intervention: (n=238)	***Comparator: (n=227)***				***Intervention: (n=302)***	***Comparator: (n=292)***			
5-year OS (95% CI) (prior to protocol amendmen)	99.6% (97.0%–99.9%)	100%	NR	NR	NR	98.3% (96.0% –99.3%)	96.7% (93.7% – 98.3%)	NR	NR	NR
5-year PFS (95% CI)	87.1% (82.1%–90.8%)	99.0% (95.9% –99.7%)	HR 15.8 (3.8–66.1), p=NR**	NR	NR	89.6% (85.5% –92.6%)	92.1% (88.0% –94.8%)	HR 1.45 (95% CI, 0.8 to 2.5), p=NR††	NR	NR

Outcome	Radford 2015 *RAPID trial* Initial therapy: ABVD, 3 cycles				
	PET – (randomized) (5-PS score of 1 or 2)			PET +	PET + vs. PET –
	Intervention: No further treatment (n=211)	Comparator: IFRT (n=209)	Comparator vs. Intervention	ABVD (1 cycle) + IFRT (n=145)	
3-year OS (95% CI)	99.0% (95% CI, 97.6%–100%)	97.1% (95% CI, 94.8%–99.4%)	HR 0.51 (95% CI, 0.15%–1.68%), p=0.27	NR	NR

Outcome	Radford 2015

Outcome	Andre 2017/Raemakers 2014[26,128] *EORTC/LYSA/FIL H10 Trial* Initial therapy: ABVD, 2 cycles							
	Favorable PET – Group* (randomized) (5-PS score of 1 or 2)		PET +	PET + vs. PET –	**Unfavorable PET – Group† (randomized) (5-PS score of 1 or 2)**		PET +	PET + vs. PET –
	RAPID trial Initial therapy: ABVD, 3 cycles							
3-year PFS (95% CI)	90.8% (86.9%–94.8%)	94.6% (91.5%–97.7%)	HR 1.57 (95% CI, 0.84%–2.97%) p=0.16; Absolute RD 3.8% (95% CI -1.3 to 8.8)	NR	NR			

ABVD =doxorubicin, bleomycin, vinblastine, and dacarbazine; CI = confidence interval; HR = hazard ratio; IFRT = involved-field radiotherapy; INRT = involved-node radiotherapy; NR = not report; OS = overall survival; PET = positron emission tomography; PFS = progression-free survival;

*Favorable status indicates age <50 years with ≤3 involved nodal areas, absence of mediastinal bulk (mediastinum-to-thorax ratio <0.35), and erythrocyte sedimentation rate (ESR) <50 mm without B symptoms or ESR <30 mm with B symptoms.

†Unfavorable status indicates age ≥50 years, >4 involved nodal areas, presence of mediastinal bulk (mediastinumto- thorax ratio ≥0.35), or ESR ≥50 mm without B symptoms or ESR ≥30 mm with B symptoms.

‡Confidence level adjusted to significance level used in interim test: 79.6% CI for favorable group and 80.4% CI for unfavorable group.

§Results represent the data up to the protocol amendment.

**Noninferiority could not be demonstrated as the upper bound of the 95% CI for the estimated HR (66.07) exceeded the prespecified noninferiority margin (3.2).

††Noninferiority could not be demonstrated as the upper bound of the 95% CI for the estimated HR (2.50) exceeded the prespecified noninferiority margin (2.10).

Table 25. Early Hodgkin (stage I-II): RCTs of PET-adapted therapy evaluating *de-escalation* of therapy: Mortality and secondary malignancy

Outcome	Andre 2017/Raemaekers 2014[26,128] *EORTC/LYSA/FIL H10 Trial* Initial therapy: ABVD, 2 cycles **PET – (randomized) (N=1059)**						Radford 2015[126] *RAPID trial* Initial therapy: ABVD, 3 cycles **PET – (randomized) (N=420)**		
	Favorable (n=465)			Unfavorable (n=594)					
	Intervention: 2 additional ABVD cycles (n=238)	*Comparator*: 1 ABVD cycle + INRT (n=227)	Intervention vs. Comparator RR/RD (95% CI)*	*Intervention*: 4 additional ABVD cycles (n=302)	*Comparator*: 2 ABVD cycle + INRT (n=292)	Intervention vs. Comparator RR/RD (95% CI)*	*Intervention*: No further treatment (n=211)	*Comparator*: IFRT (n=209)	Intervention vs. Comparator RR/RD (95% CI)*
Deaths, due to:	1.3% (3/238)	0%	RR NC; RD 1%, p=0.09	2.0% (6/302)	3.4% (10/292)	RR 0.6 (0.2–1.6); RD -1% (-4% to 1%); p=0.28	1.9% (4/211)	3.8% (8/209)	RR 0.5 (0.2–1.6); RD -2% (-5% to 1%); p=0.24
Progression/ relapse	0%	0%	-----	1.0% (3/302)	1.0% (3/292)	RR 1.0 (0.2–4.8); RD 0% (-2% to 2%); p=0.98	0%	0.5% (1/209)	RR NC; RD -0.5%, p=0.32
Toxicity of protocol treatment	0%	0%	-----	0.3% (1/302)	0%	RR NC; RD 0.3%, p=0.33	-----	-----	-----
Toxicity of second-line treatment	0.4% (1/238)	0%	RR NC; RD 0.4%, p=0.33	0%	0.3% (1/292)	RR NC; RD -0.3%, p=0.31	-----	-----	-----
Cardio-vascular event	0%	0%	-----	0%	0.7% (2/292)	RR NC; RD -0.7%, p=0.15	0%	0.5% (1/209)	RR NC; RD -0.5%, p=0.32
Second malignancy	0.8% (2/238)	0%	RR NC; RD 1%, p=0.17	0.3% (1/302)	0.7% (2/292)	RR 0.5 (0–5.3); RD -0.4% (-2% to 1%); p=0.54	1.4% (3/211)	0.5% (1/209)	RR 3.0 (0.3–28.3); RD 1% (-1% to 3%), p=0.32
Other/ unknown	0%	0%	-----	0.3% (1/302)	0.7% (2/292)	RR 0.5 (0–5.3); RD -0.4% (-2% to 1%); p=0.54	-----	-----	-----
Pneumonia	-----	-----	-----	-----	-----	-----	0.5% (1/211)	0.5% (1/209)	RR 1.0 (0.1–15.7); RD 0%

Outcome	**Andre 2017/Raemaekers 2014[26,128]** *EORTC/LYSA/FIL H10 Trial* Initial therapy: ABVD, 2 cycles **PET – (randomized) (N=1059)**						**Radford 2015[126]** *RAPID trial* Initial therapy: ABVD, 3 cycles **PET – (randomized) (N=420)**		
	Favorable (n=465)			Unfavorable (n=594)					
									(-1% to 1%), p=0.99
Pneumonitis	-----	-----	-----	-----	-----	-----	0%	1.0% (2/209)	RR NC; RD -1.0%, p=0.15
Cerebral hemorrhage	-----	-----	-----	-----	-----	-----	0%	0.5% (1/209)	RR NC; RD -0.5%, p=0.32
Mycosis fundoides	-----	-----	-----	-----	-----	-----	0%	0.5% (1/209)	RR NC; RD -0.5%, p=0.32
Second malignancy	2.9% (7/238)	1.3% (3/227)	RR 2.2 (0.6–8.5); RD 2% (-1% to 4%), p=0.23	3.0% (9/302)	3.4% (10/292)	RR 0.9 (0.4–2.1); RD -0.4% (-3% to 2%); p=0.76	-----	-----	-----

ABVD =doxorubicin, bleomycin, vinblastine, and dacarbazine; CI = confidence interval; eBEACOPP = escalated bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine and prednisone; HR = hazard ratio; IFRT = involved-field radiotherapy; INRT = involved-node radiotherapy; NC = not calculable; NR = not report; OS = overall survival; PET = positron emission tomography; PFS = progression-free survival; RD = risk difference; RR = relative risk.
*Calculated by AAI.

Table 26. Early Hodgkin (stage I-II): Nonrandomized observational subanalysis (case series)* from Andre 2017/Raemaekers 2014 RCT of PET-adapted therapy following interim PET

Author (year) Study Design	N Initial treatment/ cycles prior to interim PET F/U PET criteria	Treatment received in interim PET (–) patients following *safety amendment*	Primary Outcomes		Authors' conclusion
			OS (95% CI)	PFS (95% CI)	
Andre 2017/Raemaekers 2014[26,128] Nonrandomized observational (observational continuation of study after protocol - all patients got radiation) *Subanalysis of PET(-) patients treated following safety amendment protocol**	N=505 PET - ABVD/2 cycles	**PET (–) Favorable , n=185:** 1 additional cycle of ABVD + INRT	100%	98.9% (95.6%–99.7%)	The outcomes of the 505 ePET-negative patients who were included after the safety amendment and all treated with CMT did not differ from those of the standard CMT in the randomized setting (Table 2).
		PET (–) Unfavorable , n=320: 2 additional cycles of ABVD + INRT	99.7% (97.7%–100.0%)	95.5% (92.5%–97.3%)	
			p=NR	p=NR	

ABVD =doxorubicin, bleomycin, vinblastine, and dacarbazine; CI = confidence interval; F/U = follow-up; HR = hazard ratio; NR = not reported; NS = not statistically significant; OS = overall survival; PET = positron emission tomography; PFS = progression-free survival; RT = radiotherapy

* All early PET-negative patients who were included after the safety amendment received treatment with ABVD + INRT, as the independent data monitoring committee recommended to close the ABVD only arm

4.3.2.1.2. PET-adapted treatment: results from observational studies

Two observational studies[58,140] (N=170 and 149) reported the use of interim PET to guide both escalation and de-escalation of therapy in adult patients (age 18-59 years, median 30 years) with early-stage HL (stage I-II without bulky tumor)Tables 27 and 28. In both studies, the induction treatment was ABVD and interim PET was performed after two cycles. Patients with positive interim PET received escalated treatment; those with negative interim PET received de-escalated treatment. Without comparisons to groups receiving standard treatment to understand the impact of escalated or de-escalated treatment, conclusions regarding the impact of these treatments based on interim PET findings are difficult.

In the one study reporting OS, 5 year prognosis for those with a positive interim PET was somewhat lower (5/100) compared with those having a negative interim PET. The small number of PET-positive patients limits the ability to draw firm conclusions. Across both studies, the probability of progression-free survival (PFS) was significantly greater for patients with a negative versus a positive interim PET: 3-year PFS[140] (HR 3.8, 95% CI 1.5 to 9.8) in one study and 5-year PFS[58] (HR 3.7, 95% CI 1.4 to 9.7; p=0.008) in the other (Table 27). Grade 3 or 4 toxicities were reported according to interim PET status (PET+ and PET-) in one study. In general, toxicity, particularly hematological toxicity, was greater following the escalated treatment given to those with a positive interim PET compared with the deescalated therapy for those with a negative interim PET (Table 28).

Table 27. Early HL (stage I-II): Survival outcomes from observational studies of PET-adapted therapy following interim PET

		Author	N PET +	PET + Estimate (95%CI)	N PET –	PET – Estimate (95%CI)	Overall or comparative if given (95%CI); notes
Escalated and De-escalated therapy							
Adults							
OS	3-year	Straus 2018[140]	14	NR	135	NR	1 death (due to suicide in the PET+ group)
	5-year	Dann 2017[58]	20	95% (85.4%–104.6%)	150	100%	NS difference PET + vs PET -
PFS	1-year	Straus 2018[140]	14	75% (41%–91%)	135	96% (91%–98%)	NR
	2-year			67% (34%–86%)		93% (87%–96%)	NR
	3-year			67% (34%–86%)		91% (84%–95%)	PET – vs. PET+, HR 3.8 (1.5–9.8); p=0.01
	5-year	Dann 2017[58]	20	69.2% (48.6%–89.8%)	150	91.4% (86.7%–96.1%)	PET – vs. PET+, HR 3.7 (1.4–9.7); p=0.008

aHR = adjusted hazard ratio; CI = confidence interval; EFS = event-free survival; NR = not report; OS = overall survival; PET = positron emission tomography; PFS = progression-free survival.

Table 28. Early HL (stage I-II): Comparative data on treatment toxicities from one observational studies of PET-adapted therapy following interim PET in adults

	Straus 2018[140]		
Grade 3-4 Toxicities	**PET+ eBEACOPP + RT (escalated)**	**PET – ABVD only (de-escalated)**	RR (95% CI)*
Hematological			
Anemia	15% (2/13) (*both grade 3*)	0% (0/131)	RR NC, p<0.0001
Thrombocytopenia	8% (1/13) (*grade 3*)	1% (1/131) (*grade 3*)	RR 10.0 (0.7–151.8), p=0.04
Neutropenia	69% (9/13) *grade 3*: 23% (3/13) *grade 4*: 46% (6/13)	70% (92/131) *grade 3*: 28% (37/131) *grade 4*: 42% (55/131)	RR 1.0 (0.7–1.4), p=0.94 RR 0.8 (0.3–2.3), p=0.69 RR 1.1 (0.6–2.0), p=0.77
Febrile neutropenia	15% (2/13) *grade 3*: 8% (1/13) *grade 4*: 8% (1/13)	5% (7/131) *grade 3*: 5% (7/131) *grade 4*: 0% (0/131)	RR 2.9 (0.7–12.5), p=0.16 RR 1.4 (0.2–10.8), p=0.73 RR NC, p=0.002
Lymphopenia	23% (3/13) *grade 3*: 15% (2/13) *grade 4*: 8% (1/13)	5% (6/131) *grade 3*: 3% (4/131) *grade 4*: 2% (2/131)	RR 5.0 (1.4–17.8), p=0.01 RR 5.0 (1.0–24.9), p=0.03 RR 5.0 (0.5–51.9), p=0.14
Leukopenia	54% (7/13) *grade 3*: 8% (1/13) *grade 4*: 45% (6/13)	35% (46/131) *grade 3*: 32% (42/131) *grade 4:* 3% (4/131)	RR 1.5 (0.9–2.7), p=0.18 RR 0.2 (0.04–1.6), p=0.07 RR 15.1 (4.9–46.8), p<0.0001
Cardiovascular			
Left ventricular systolic dysfunction	0% (0/13)	1% (1/131)	RR NC, p=0.75
Decreased ejection fraction	0% (0/13)	1% (1/131)	RR NC, p=0.75
Gastrointestinal			
Constipation	0% (0/13)	1% (1/131)	RR NC, p=0.75
Mucositis oral	8% (1/13)	0% (0/131)	RR NC, p=0.002
Nausea	0% (0/13)	2% (3/131)	RR NC, p=0.58
Vomiting	0% (0/13)	4% (5/131)	RR NC, p=0.47
General			
Fatigue	8% (1/13)	2% (3/131)	RR 3.4 (0.4–30.0), p=0.26
Infections			
Kidney infection	0% (0/13)	1% (1/131)	RR NC, p=0.75
Sepsis	0% (0/13)	1% (1/131)	RR NC, p=0.75
Skin infection	0% (0/13)	2% (2/131)	RR NC, p=0.58

	Straus 2018[140]		
Grade 3-4 Toxicities	**PET+ eBEACOPP + RT (escalated)**	**PET – ABVD only (de-escalated)**	RR (95% CI)*
Metabolic and nutrition disorders			
Glucose intolerance	0% (0/13)	1% (1/131)	RR NC, p=0.75
Hyperglycemia	0% (0/13)	2% (2/131)	RR NC, p=0.58
Hypoglycemia	0% (0/13)	1% (1/131)	RR NC, p=0.75
Musculoskeletal and connective tissue disorders			
Bone pain	8% (1/13)	0% (0/131)	RR NC, p=0.002
Myalgia	0% (0/13)	1% (1/131)	RR NC, p=0.75
Nervous system disorders			
Peripheral motor neuropathy	0% (0/13)	1% (1/131)	RR NC, p=0.75
Peripheral sensory neuropathy	0% (0/13)	3% (4/131)	RR NC, p=0.52
Vasovagal reaction	0% (0/13)	1% (1/131)	RR NC, p=0.75
Respiratory disorders			
Dyspnea	8% (1/13)	1% (1/131)	RR 10.0 (0.7–151.8), p=0.04
Pneumonitis	0% (0/13)	1% (1/131)	RR NC, p=0.75
Vascular disorders			
Thromboembolic event	0% (0/13)	2% (2/131)	RR NC, p=0.58

eBEACOPP = escalated bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine and prednisone; NC = not calculable; PET = Positron Emission Tomography; RR = Risk Ratio.

4.3.2.1.3. Other observational studies

One comparative observational study (N=123) was identified that reported the impact of PET/CT on radiation treatment (RT) planning for early-stage HL.[123] The study included primarily adult patients (87%, median age 33 years; 13% children, median age 16 years) receiving RT following completion of chemotherapy.[123] No significant difference was seen in the proportion of treatment plan modifications for patients who had a pre-treatment PET/CT compared with those who did not: 78.7% (95% CI, 66.3% to 88.1%) to include three cancellations and 10 modifications, versus 85.5% (95% CI, 74.2% to 93.1%) to include 3 cancellation and 6 modifications, respectively. Modifications to treatment included changes to the planned dose, main clinical target volume (CTV), number of beam incidences, and number of CTVs.

4.3.2.2. Advanced-Stage HL

Studies included

A total of three randomized PET-adapted treatment trials in persons with advanced-stage HL were identified.[37,38,94] Two trials compared randomized escalated treatment regimens in persons with positive

interim PET results.[37,38,77] One of these also compared persons with negative interim PET results randomized to de-escalation of treatment[37,38] versus continuation of an escalated treatment. The third trial randomized those with PET negative results to de-escalated therapy versus continuation of standard treatment.[94] Randomization in the trials was to treatment regiments based on interim PET results. Trials did not compare escalated treatments with standard treatments. There is substantial heterogeneity across trials with regard to initial treatments, PET-adapted treatment regimens, patient populations and criteria for determining positivity of PET results.

The GHSG HD18 trial (n=1964, with data across two publications)[37,38] randomized patients with interim PET positive and negative results. First-line regimen for all 1964 patients in the trial consisted of two cycles of eBEACOPP. The median age in the trial was 32, and the majority of patients were male (62%). After randomization based on interim PET results, positive patients either continued standard eBEACOPP (8 cycles) or were escalated to eBEACOPP with rituximab (8 cycles), while negative patients continued standard eBEACOPP (8 cycles) or were de-escalated to a reduced number of cycles (eBEACOPP x4). Additionally, a protocol change was introduced during the study, which reduced the number of cycles (from 8 cycles to 6 cycles) for those allocated to standard continuation therapy from then onward. After initial therapy was completed no further treatment was offered during the course of the study.

The GITIL/FIL HD 0607 trial (n=782)[77] randomized solely those with positive interim PET results (19%), who were randomly escalated to eBEACOPP with (n=72) or without rituximab (n=76). Negative interim PET patients continued first-line treatment (ABVD) for an additional four cycles. Additional PET scans were conducted at the end of first-line therapy for both positive and negative interim groups. In the positive arm, final PET results further allocated interim positive patients to no further treatment (n=108 final PET negative patients) or RT, stem cell transplant or no further treatment (n=28 final PET positive patients). In the interim PET negative arm, those who were PET positive at the time of final PET received either no further treatment, RT or stem cell transplant, while those who had continued to have a negative PET scan either had no further treatment or were randomized between RT and no further treatment based on the presence of residual large nodal mass. Overall, the study had slightly more females (51%), and the median age of patients was 31.

Another trial[94] (n=1203, median age 33, 55% male) began treatment with two cycles ABVD before randomizing interim PET negative patients to either continuation of first-line treatment (four additional cycles) or reduction by removal of bleomycin (AVD for four cycles). Interim PET positive patients received BEACOPP (six cycles) or eBEACOPP (four cycles). All iPET negative patients received no further treatment after completion of adapted therapy. Positive interim PET patients went on to receive an additional follow-up PET scan which allocated patients between no further treatment and escalation to salvage therapy.

Additionally, eight nonrandomized observational studies (six prospective, two retrospective) of interim PET for advanced-stage HL were identified.[39,58,60,78,95,124,133,169] Across all eight studies, first-line treatment regimens varied, with four using ABVD,[78,124,133,169] three using eBEACOPP,[58,60,95] and one using VABEM,[39] all for two initial cycles. The number of patients in the studies ranged from 50 to 336, the median ages across studies ranged from 25 to 41 years and all but one study[78] had between 40% and 60% males.

Six studies escalated therapy on interim PET positive results.[39,58,78,124,133,169] Across these studies, the proportion of PET positive patients at interim PET ranged from 14.5% to 23.5% of patients. Therapy

adaptations for PET positive patients included BEACOPP or eBEACOPP (between four and eight cycles), IGEV (four cycles), and PDG as escalation options, while interim PET negative patients continued respective first-line treatments for four cycles in most cases (one cycle only in Carras). Additional PET scans were conducted in three of the six studies.[39,133,169] One of these studies conducted further interim scans at after two cycles of PDG for PET positive patients, and further escalated therapy for one continued positive patient with MINE-R (3 cycles), and continued PDG for one additional cycle followed by BEAM for second interim PET negative patients. All patients continued afterwards onto HDT/ASCT and some even further to RT. Among the patients who were PET negative at the first interim scan, most received no further treatment while a subset received RT consolidation. In another study,[169] additional PET scans were conducted for both positive and negative groups; interim PET positive patients received an end of treatment scan that allocated EoT-PET positive patients to HDT and BEAM with autologous bone marrow transplant (BMT) or HDT and reduced intensity allogenic BMT, while EoT-PET negative patients were allocated to BEAM with autologous BMT. For interim PET negative patients, an additional PET scan removed continued positive patients from the study, and randomized negative patients between RT and no further treatment. The final study conducted a final PET scan at the end of adapted therapy to assess treatment response but did not allocate patients to any further treatment afterwards. Three other studies did not feature additional PET scans and all but one of these had no further treatment after interim PET. In the study that did feature additional treatment,[58] a portion (n=15) of interim PET positive patients were allocated to RT (this study also featured a cohort of early-stage HL patients, whose data is reported in the results for early-stage HL).

Of the two retrospective observational studies that de-escalated treatment based on interim PET results, one study (n=64)[60] allocated interim PET positive patients (n=9, 14% of total) to additional eBEACOPP cycles while interim PET negative patients (n=55) were de-escalated to ABVD for 4 cycles.[60] The other study followed a similar de-escalation protocol, except de-escalation to four cycles of ABVD was applied to both interim PET positive (n=17) and interim PET negative patients (n=52).[95] The median ages in these studies were 25 and 30 years, and males comprised between 42% and 65% of the populations. In the first study further PET scans were conducted for PET positive patients after the additional eBEACOPP cycles, after which patients who remained PET positive at the second scan were allocated to salvage therapy (n=5, off-study), while negative patients finished with either two additional cycles of eBEACOPP or de-escalation to ABVD for two final cycles.[60] In the second study, no further scans were conducted and only a subset (n=4) received any further treatment (RT).[95]

4.3.2.2.1. PET-adapted treatment: results from RCTs

Escalated therapy

Across two trials in which patients with advanced-stage HL and positive interim PET were randomized to different escalated therapies, there were no statistical differences in estimated probabilities of OS or PFS between escalated treatments through the time points reported. Each trial used different induction treatments and the number of escalation cycles differed. The HD 0607 trial randomized persons with a positive interim PET (N= 148) after two ABVD cycles to either escalated therapy with R-BEACOPP or BEACOPP for 8 cycles.[77] The HD18 trial randomized persons with a positive interim PET (N= 434) after two BEACOPP cycles to an additional 6 cycles of R-BEACOPP or BEACOPP.[37,38] At 3 years, a small

estimated difference in OS between treatments was seen in each trial (-1/100 and -1.4/100), but these differences were not beyond what might have occurred by chance. Somewhat higher OS was seen in the HD18 trial than the HD0607 and OS was similar for interim PET + and PET - groups in this trial while there was an estimated 10/100 difference between PET + and PET – groups in the HD 0607 trial. At 5 years in the HD18 trial, OS was slightly lower in the R-BEACOPP group (93.9%) versus the BEACOPP group (96.4%, reported absolute RD -2.5%) (Table 29).

Estimated probability of PFS at 3 years was similar across treatment groups for PET + group in the HD 18 trial (estimated difference 0.5/100) but was somewhat higher for escalated treatment with R-BEACOPP in the HD0607 trial vs. BEACOPP (Table 29); however, the difference (estimate, 6/100) was not be beyond what would be expected by chance. PFS overall was substantially higher in the HD18 trial and as was seen for OS, there was no statistical difference in PFS between PET + and PET – groups in HD 18; however, a significant difference between PET + and PET – was reported for the HD 0607 trial. At 5 years, PFS was slightly lower than what was seen at 3 years, but was similar across the escalated treatment groups.

Authors of both trials report that the addition of rituximab to BEACOPP did not improve patient outcomes in persons with a positive interim PET compared with BEACOPP alone; The HD18 trial,[38] in which initial treatment was BEACOPP, reported that positivity of interim PET did not identify higher risk persons when the German Hodgkin Study Group (GHSG) standard protocol is used and that the positive PET findings did not necessitate treatment modification. In the HD0607 trial,[77] patients received ABVD as the initial treatment authors felt that interim PET positivity facilitated escalation to BEACOPP in higher risk individuals.

The risk of grade 3 or 4 treatment-related toxicity across the escalated treatment groups in the HD18 trial was statistically similar, but limited detail was provided (Table 30). No toxicity data specific to escalated PET-adapted therapy in persons with a positive interim PET were reported in the HD 0607 trial fewer toxic events occurred in persons with interim PET negative versus interim PET positive therapies.

The RATHL trial[94] described included a non-randomized arm involving escalation of treatment among persons with positive interim PET, comparing escalated BEACOPP an accelerated version (BEACOPP-14) involving growth-factor support and had additional PET/CT imaging after four and three cycles of these treatments respectively. Authors report that there were no substantial differences between the two treatment that 3-year progression-free survival rate for the group as a whole was 67.5% (95% CI, 59.7 to 74.2), and the overall survival rate was 87.8% (95% CI, 81.5 to 92.1).

De-escalated therapy

De-escalation of therapy for persons with a negative interim PET was evaluated in two trials. The HD18 randomized those with negative interim PET (N =1005) to a reduced number of BEACOPP cycles (4 cycles) or continuation of 6 to 8 BEACOPP cycles.[37,38] In the RATHL trial (N = 935) persons were randomized to de-escalated treatment where bleomycin was omitted or to continuation of ABVD.[94] Across both trials, OS and PFS were either somewhat better for the group receiving deescalated treatment (HD 18 Trial) or similar (RATHL Trial) and overall, treatment-related toxicity was substantially reduced for those receiving the deescalated treatment versus those continuing with more intensive treatment.

Across both trials through 3 years, differences between de-escalated therapies and continuation of more intensive treatment were not beyond what might be expected by chance. In the RAHTL trial,[94] difference in OS between the group for which bleomycin was omitted and the group for which was retained was small (estimated difference 0.4/100). OS for those receiving a reduced number of BEACOPP cycles in the HD18 trial was slightly greater versus those who continued with 6-8 cycles (absolute RD 3.1%).[37,38] Through 5 years in the HD18 trial, OS in those receiving fewer cycles was slightly greater (97.7%) versus those who received 6-8 cycles (95.4%, absolute difference, 2.3%) (Table 31).

Across both trials, through 3 years, PFS was statistically similar in groups receiving de-escalated treatment versus those with more intensive therapy in both trials. In the HD18 trial,[37,38] PFS was somewhat greater in those with de-escalated treatment (95.3% vs. 91.7%; absolute RD: 3.6%) and in the RATHL trial[94] the difference between groups was smaller (84.4% vs. 85.7%, absolute RD -1.6%). Through 5 years in the HD 18 trial, the slightly better PFS in the group receiving the de-escalated treatment (92.2% vs. 90.8%, absolute RD 1.4%) was again seen, however the difference was not statistically significant. (Table 31).

Both trials set out to test the non-inferiority of de-escalated treatments using PFS as primary outcome. Although results from both trials suggest that OS and PFS in those receiving scaled back treatment are better than or similar to those continuing standard treatment, the RATHL trial[94] concluded that although the results fall just short of the specified noninferiority margin, the omission of bleomycin from the ABVD regimen after negative findings on interim PET resulted in a lower incidence of pulmonary toxic effects than with continued ABVD but not significantly lower efficacy. By contrast authors of the HD18 trial concluded that non-inferiority was demonstrated.[37,38]

Overall, treatment related toxicity was substantially lower when treatment was de-escalated as were mortality and frequency of secondary malignancy (Tables 32 and 33). In the HD18 trial,[37,38] treatment related morbidity and any organ toxicity were significantly less common in those receiving 4 BEACOPP cycles versus 6 to 8 cycles and substantially lower risk of grade 3 or 4 anemia, thrombocytopenia, leukopenia, infection and febrile neutropenia was reported as well. In the RATHL trial,[94] the occurrence of any clinical adverse event was significantly lower in the AVD group compared with the ABVD group as was the risk of febrile neutropenia, pulmonary/respiratory events; however the risk of thrombocytopenia was greater.

Table 29. Advanced-Stage HL (stage IIB-IV): RCTs of PET-adapted therapy evaluating *escalation* of therapy following interim PET: Probability of overall and progression free survival

Outcome	Gallamini 2018 *GITIL/FIL HD 0607 Trial* Initial therapy/PET timing: ABVD/2 cycles					Borchmann 2017a, 2017b *GHSG HD18 trial* Initial therapy/PET timing: eBEACOPP/2 cycles				
	PET + (randomized) (5-PS score of 4 or 5) to 8 cycles (4 baseline, 4 escalated) of:			PET –	PET + vs. PET –	**PET + (randomized) (5-PS score ≥3) to 6 additional cycles* of:**			PET –	PET + vs. PET –
	Intervention: R-BEACOPP (n=72)	Comparator: BEACOPP (n=76)	Comparator vs. Intervention	ABVD (n=630)		Intervention: R-eBEACOPP (n=217)	Comparator: eBEACOPP (n=217)	Comparator vs. Intervention	eBEACOPP (n=1010)	
3-year OS (95% CI)	89% (79%–95%)	90% (78%–95%)	p=0.895	99% (97%–99%)	p<0.001†	95.6% (92.8%–98.4%)	97.1% (94.8%–99.4%)	Absolute RD –1.4% (95% CI –5.1% to 2.2%)	97.3% (96.2%–98.3%)	NR
3-year PFS (95% CI)	63% (50%–74%)	57% (45%–68%)	p=0.534	87% (84%–89%)	p<0.001†	93.2% (89.8%–96.7%)	92.7% (89.1%–96.2%)	Absolute RD 0.6% (95% CI –4.4% to 5.5%)	3.5(95.1%)	NR
5-year OS (95% CI)	-----	-----	-----	-----	-----	93.9% (90.6%–97.3%)	96.4% (93.8%–99.0%)	HR 1.62 (95% CI, 0.70–3.75), p=0.25 Absolute RD –2.5% (95% CI –6.8% to 1.7%)	96.3% (95.0%–97.6%)	p=0.49‡
5-year PFS (95% CI)	-----	-----	-----	-----	-----	88.1% (83.5%–92.7%)	89.7% (85.4%–94.0%)	R 1.25 (95% CI, 0.69–2.26), p=0.46Absolute RD –1.6% (95% CI –7.9% to 4.7%)	91.4% (89.5%–93.4%)	p=0.30‡

ABVD =doxorubicin, bleomycin, vinblastine, and dacarbazine; BEACOPP = bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone; CI = confidence interval; eBEACOPP = escalated (i.e., increased doses of) bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone; HR = hazard ratio; NR = not reported; OS = overall survival; PET = positron emission tomography; PFS = progression-free survival; R-BEACOPP = rituximab + bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone; R-eBEACOPP = rituximab + escalated (i.e., increased doses of) bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone.

*OS and PFS only reported for patients randomized prior to the protocol amendment (i.e., who received 8 total cycles) which introduced a reduction of standard therapy from 8 to 6 cycles of eBEACOPP in total. Because the required sample size for the treatment group comparison in the PET(+) cohort had already been reached, randomization between eBEACOPP and R-eBEACOPP after PET(+) was stopped at this point, and all patients with a PET(+) were subsequently assigned to receive the new standard treatment of 6 × eBEACOPP.

†3-year OS and PFS for the PET + group overall (n=150) were 89% (95% CI, 82% to 93%) and 60% (95% CI, 51% to 68%), respectively.

‡5-year OS and PFS for the PET + group overall (n=948) were 95.5% (95% CI, 93.9% to 97.1%) and 88.3% (95% CI, 85.8% to 90.8%), respectively.

Table 30. Advanced-Stage Hodgkin (stage IIB-IV): RCTs of PET-adapted therapy evaluating *escalation* of therapy: Mortality, Toxicity and Secondary malignancy

Outcome	Borchman 2017a, 2017b *GHSG HD18 trial* *Initial therapy/PET timing: eBEACOPP/2 cycles*		
	PET + (randomized) (N=440)* to 6 additional cycles of:		
	Intervention: R-eBEACOPP (n=220)	*Comparator*: eBEACOPP (n=219)	CIntervention vs. Comparator RR/RD (95% CI)†
Death	6.5% (14/217)	4.1% (9/217)	RR 1.6 (0.7–3.5); RD 2% (-2% to 7%); p=0.29
Toxicity of salvage therapy	1.8% (4/217)	1.4% (3/217)	
Toxicity of study treatment	1.4% (3/217)	0.5% (1/217)	
Second malignancy	0.9% (2/217)	0.9% (2/217)	
Other disease	0.9% (2/217)	0.9% (2/217)	
Accident or suicide	0.9% (2/217)	0%	
Hodgkin lymphoma	0.5% (1/217)	0.5% (1/217)	
Any acute toxicity, grade 3/4 (≥1 event)	97% (213/220)	98% (213/218)	RR 1.0 (0.96–1.0); RD -1% (-4% to 2%); p=0.57
Any acute hematological toxicity, grade 3/4 (≥1 event)	96% (211/220)	97% (211/218)	
Anemia, thrombocytopenia, or infection, grade 3/4	85% (187/220)	81% (177/218)	
Any acute organ toxicity, grade 3/4 (≥1 event)	18% (40/220)	18% (39/218)	
Second malignancies, any event	3.7% (8/217)	4.6% (10/217)	5-year cumulative incidence estimate‡: 3.5% (95% 0.7%-6.3%) vs. 4.0% (95% CI 1.3%-6.7%)
Solid tumor	3.7% (8/217)	0.9% (2/217)	
Acute myeloid leukemia or myelodysplastic syndrome	1.8% (4/217)	2.3% (5/217)	
Non-Hodgkin lymphoma	0.9% (2/217)	1.4% (3/217)	

ABVD =doxorubicin, bleomycin, vinblastine, and dacarbazine; CI = confidence interval; eBEACOPP = escalated bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine and prednisone; HR = hazard ratio; IFRT = involved-field radiotherapy; INRT = involved-node radiotherapy; NR = not report; OS = overall survival; PET = positron emission tomography; PFS = progression-free survival; RD = risk difference; RR = relative risk.

*Only reported for patients randomized prior to the protocol amendment which introduced a reduction of standard therapy from 8 to 6 cycles of eBEACOPP in total. Because the required sample size for the treatment group comparison in the PET(+) cohort had already been reached, randomization between eBEACOPP and R-eBEACOPP after PET(+) was stopped at this point, and all patients with a PET(+) were subsequently assigned to receive the new standard treatment of 6 × eBEACOPP.

†Calculated by AAI.

‡Accounting for death as competing risk.

Table 31. Advanced-Stage Hodgkin (stage IIB-IV): RCTs of PET-adapted therapy evaluating *de-escalation* of therapy following interim PET: Survival outcomes

Outcome	Johnson 2016 *RATHL Trial* Initial therapy/PET timing: ABVD/2 cycles					Borchmann 2017a, 2017b *HD18 trial* Initial therapy/PET timing: eBEACOPP/2 cycles				
	PET – (randomized) (1-3 on 5-point scale) to 4 cycles of:			PET +	PET + vs. PET –	**PET – (randomized) (1-3 on 5-point scale) to:**			PET +	PET + vs. PET –
	Intervention: AVD (n=465)	Comparator: ABVD (n=470)	Comparator vs. Intervention	BEACOPP (n=172)		Intervention: eBEACOPP, x4 cycles (n=501)	Comparator: eBEACOPP, x6-8 cycles (n=504)	Intervention (4 cycles) vs. Comparator (6-8 cycles)	eBEACOPP or R-eBEACOPP (n=948)	
3-year OS (95% CI)	97.6% (95.6%–98.7%)	97.2% (95.1%–98.4%)	HR 0.09 (95% CI, 0.47–1.74), p=0.76	87.8% (81.5%–92.1%)	NR	98.8% (97.8%–99.9%)	95.7% (93.7%–97.6%)	HR 0.32 (95% CI, 0.14–0.72) p=0.0037 Absolute RD: 3.1% (95% CI 0.9%–5.3%)	97.1% (95.9%–98.2%)	NR
3-year PFS (95% CI)	84.4% (80.7%–87.5%)	85.7% (82.1%–88.6%)	HR 1.13 (95% CI, 0.81–1.57), p=0.48 Absolute RD1.6% (95% CI –3.2 to 5.3)	67.5% (59.7%–74.2%)	NR	95.3% (93.3%–97.3%)	91.7% (89.0%–94.4%)	Absolute RD: 3.6% (95% CI –0.2% to 6.9%)	92.5% (90.7%–94.3%)	NR
5-year OS (95% CI)	-----	-----	-----	-----	-----	97.7% (96.2%–99.3%)	95.4% (93.4%–97.4%)	Absolute RD: 2.3% (95% CI –0.2% to 4.9%)	95.5% (93.9%–97.1%)	p=0.49*
5-year PFS (95% CI)	-----	-----	-----	-----	-----	92.2% (89.4%–95.0%)	90.8% (87.9%–93.7%)	HR 0.79 (95% CI, 0.50–1.24), p=NS Absolute RD: 1.4% (95% CI –2.7% to 5.4%)	88.3% (85.8%–90.8%)	p=0.30*

ABVD =doxorubicin, bleomycin, vinblastine, and dacarbazine; AVD = doxorubicin, vinblastine, and dacarbazine; BEACOPP = bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone; CI = confidence interval; eBEACOPP = escalated (i.e., increased doses of) bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone; HR = hazard ratio; NR = not reported; OS = overall survival; PET = positron emission tomography; PFS = progression-free survival; R-BEACOPP = rituximab + bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone; R-eBEACOPP = rituximab + escalated (i.e., increased doses of) bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone.

*5-year OS and PFS for the PET – group overall (n=1010) were 96.3% (95% CI, 95.0% to 97.6%) and 91.4% (95% CI, 89.5% to 93.4%), respectively.

Table 32. Advanced-Stage Hodgkin (stage IIB-IV): RCTs of PET-adapted therapy evaluating *de-escalation* of therapy following interim PET: Mortality, and Secondary malignancy

Outcome	**Borchman 2017a, 2017b** *GHSG HD18 trial* *Initial therapy/PET timing: eBEACOPP/2 cycles*			**Johnson 2016** *RATHL trial* *Initial therapy/PET timing: ABVD/2 cycles*		
	PET – (randomized) (N=1005)*			**PET – (randomized) to 4 cycles of:**		
	Intervention: eBEACOPP, 4 cycles (n=501)	*Comparator*: eBEACOPP, 6-8 cycles (n=504)	Intervention vs. Comparator RR/RD (95% CI)†	Intervention: AVD (n=465)	Comparator: ABVD (n=470)	Intervention vs. Comparator RR/RD (95% CI)†
Death	1.8% (9/501)	5.0% (25/504)	RR 0.4 (0.2–0.8); RD -3% (-5% to -1%); p=0.006	3.7% (17/465)	4.0% (19/470)	RR 0.9 (0.5–1.7); RD -0.3% (-3% to 2%); p=0.76
Second malignancy	0.2% (1/501)	2.2% (11/504)	RR 0.09 (0.01–0.7); RD -2% (-3% to -1%); p=0.004	1.3% (6/465)	0.9% (4/470)	RR 1.5 (0.4–5.3); RD 0.4% (-1% to 2%); p=0.51
Toxicity of study treatment	0%	1.2% (6/504)		0%	0.9% (4/470)	
Toxicity of salvage therapy	0.4% (2/501)	0.4% (2/504)		0.2% (1/465)	0.9% (4/470)	
Hodgkin lymphoma	0.8% (4/501)	0.6% (3/504)		1.7% (8/465)	0.9% (4/470)	
Other disease	0.2% (1/501)	0.2% (1/504)		-----	-----	
Accident or suicide	0.2% (1/501)	0%		-----	-----	
Unclear	0%	0.4% (2/504)		-----	-----	
Cardiac event	-----	-----		0.2% (1/465)	0.2% (1/470)	
Causes unrelated to HL or treatment	-----	-----		0.2% (1/465)	0.4% (2/470)	
Second malignancies, any event	2.6% (13/501)	3.6% (18/504)	5-year cumulative incidence estimate‡: 33.3% (95% CI 1.4%-5.3%) vs. 3.8% (95% CI 1.9%-5.7%)	2.4% (11/465)	2.8% (13/470)	RR 0.9 (0.4–1.9); RD -0.4% (-2% to 2%); p=0.70
Solid tumor	0.6% (3/501)	1.0% (5/504)		-----	-----	
Acute myeloid leukemia or myelodysplastic syndrome	0.4% (2/501)	1.6% (8/504)		-----	-----	

Outcome	**Borchman 2017a, 2017b** *GHSG HD18 trial* *Initial therapy/PET timing: eBEACOPP/2 cycles*			**Johnson 2016** *RATHL trial* *Initial therapy/PET timing: ABVD/2 cycles*		
	PET – (randomized) (N=1005)*			**PET – (randomized) to 4 cycles of:**		
Non-Hodgkin lymphoma	1.6% (8/501)	1.0% (5/504)		-----	-----	

ABVD =doxorubicin, bleomycin, vinblastine, and dacarbazine; AVD = doxorubicin, vinblastine, and dacarbazine (omitting bleomycin); CI = confidence interval; eBEACOPP = escalated bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine and prednisone; G-CSF=granulocyte-colony stimulating factor; HL = Hodgkin lymphoma; NR = not report; PET = positron emission tomography; RD = risk difference; RR = relative risk.

*Patients randomized prior to and after the protocol amendment (which introduced a reduction of standard therapy from 8 to 6 cycles of eBEACOPP in total) were combined into one group for comparative purposes with the group that received only 4 cycles.

†Calculated by AAI.

‡Accounting for death as competing risk.

Table 33. Advanced-Stage Hodgkin (stage IIB-IV): RCTs of PET-adapted therapy evaluating *de-escalation* of therapy following interim PET: Toxicity

Outcome	**Borchman 2017a, 2017b** *GHSG HD18 trial* *Initial therapy/PET timing: eBEACOPP/2 cycles*			**Johnson 2016** *Initial therapy/PET timing: ABVD/2 cycles*		
	PET – (randomized) (N=1005)*			**PET – (randomized) to 4 cycles of:**		
	Intervention: eBEACOPP, 4 cycles (n=501)	*Comparator*: eBEACOPP, 6-8 cycles (n=504)	Intervention vs. Comparator RR/RD (95% CI)†	*Intervention*: AVD (n=457)	*Comparator*: ABVD (n=468)	Intervention vs. Comparator RR/RD (95% CI)†
Any acute toxicity, grade 3/4‡	90.8% (455/501)	96.2% (485/504)	RR 0.94 (0.91–0.98); RD -5% (-8% to -2%); p=0.001	65.4% (299/457)	68.8% (322/468)	RR 0.9 (0.8–1.0); RD -3% (-9% to 3%); p=0.28
Hematological toxicities, grade 3/4	-----	-----	-----	59.7% (273/457)	59.8% (280/468)	RR 1.00 (0.90–1.11); RD -0.1% (-6% to 6%); p=0.98
Anemia, thrombocytopenia, or leukopenia	89.2% (447/501)	94.4% (476/504)	RR 0.94 (0.91–0.98); RD -5% (-9% to -2%); p=0.003	-----	-----	-----
Anemia	38.9% (195/501)	54.4% (274/504)	RR 0.72 (95% CI 0.63, 0.82; RD-15% (95%CI -22% to -9%	-----	-----	-----
Thrombocytopenia	57.1% (286/501)	71.8% (362/504)	RR 0.79 (95%CI 0.72 to 0.87); RD -15% (95%CI -21% to -9%)	3.3% (15/457)	1.3% (6/468)	RR 2.6 (1.0–6.5); RD 2% (0.1%–4%); p=0.04
Leukopenia	87.4% (438/501)	92.7% (467/504)	RR 0.95 (95%CI 0.91 to 0.98); RD -5% (95%CI -9% to -1.5%	-----	-----	-----
Infection	8.0% (40/501)	14.9% (75/504)	RR 0.5 (0.4–0.8); RD -7% (-11% to -3%); p=0.001	10.3% (47/457)	14.5% (68/468)	RR 0.7 (0.5–1.0); RD -4% (-8% to 0%); p=0.05
Treatment-related morbidity	40.7% (204/501)	63.7% (321/504)	RR 0.6 (0.56–0.72); RD -23% (-29% to -17%); p<0.0001	-----	-----	-----
Any organ toxicity, grade 3/4§	7.6% (38/501)	18.1% (91/504)	RR 0.4 (0.3–0.6); RD -10% (-15% to -6%); p<0.0001	-----	-----	-----
Anemia, thrombocytopenia, or infection, grade 4	37.3% (187/501)	56.3% (284/504)		-----	-----	-----
Febrile neutropenia	21.8% (109/501)	28.8% (145/504)	RR 0.8 (0.6–0.9);	2.2% (10/457)	4.7% (22/468)	RR 0.5 (0.2–0.9);

Outcome	**Borchman 2017a, 2017b** *GHSG HD18 trial* *Initial therapy/PET timing: eBEACOPP/2 cycles*			**Johnson 2016** *Initial therapy/PET timing: ABVD/2 cycles*		
	PET – (randomized) (N=1005)*			**PET – (randomized) to 4 cycles of:**		
			RD -7% (-12% to -2%); p=0.01			RD -3% (-5% to -0.2%); p=0.04
Hospitalization due to febrile neutropenia	18.0% (90/501)	21.8% (110/504)		-----	-----	-----
Discontinuation due to toxicity	0.4% (2/501)	2.2% (11/504)	RR 0.2 (0.04–0.8); RD -2% (-3% to -0.4%) p=0.01	-----	-----	-----
Supportive measures						
Use of G-CSF	98.8% (495/501)	98.8% (498/504)		-----	-----	-----
Platelet infusions	24.4% (122/501)	40.5% (204/504)	RR 0.6 (0.5–0.7); RD -16% (-22% to -10%); p<0.0001	-----	-----	-----
Red blood cell infusions	46.5% (233/501)	65.9% (332/504)	RR 0.7 (0.6–0.8); RD -19% (-25% to -13%); p<0.0001	-----	-----	-----
Any clinical adverse event**	-----	-----	-----	21.0% (96/457)	30.6% (143/468)	RR 0.7 (0.5–0.9); RD -10% (-15% to -4%); p=0.0009
Cardiac event	-----	-----	-----	0.4% (2/457)	1.3% (6/468)	RR 0.3 (0.1–1.7); RD -1% (-2% to 0.3%); p=0.17
Fatigue	-----	-----	-----	1.1% (5/457)	3.0% (14/468)	RR 0.4 (0.1–1.0); RD -2% (-4% to -0.1%); p=0.04
Neurologic event	-----	-----	-----	3.1% (14/457)	4.9% (23/468)	RR 0.6 (0.3–1.2); RD -2% (-4% to 1%); p=0.151
Pulmonary or upper respiratory event	-----	-----	-----	0.7% (3/457)	3.2% (15/468)	RR 0.2 (0.1–0.7); RD -3% (-4% to -1%); p=0.005
Vascular event (thrombosis, thrombus, or embolism)	-----	-----	-----	2.6% (12/457)	4.9% (23/468)	RR 0.5 (0.3–1.1); RD -2% (-5% to 0.2%); p=0.07

ABVD =doxorubicin, bleomycin, vinblastine, and dacarbazine; AVD = doxorubicin, vinblastine, and dacarbazine (omitting bleomycin); CI = confidence interval; eBEACOPP = escalated bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine and prednisone; G-CSF=granulocyte-colony stimulating factor; HL = Hodgkin lymphoma; NR = not report; PET = positron emission tomography; RD = risk difference; RR = relative risk.
*Patients randomized prior to and after the protocol amendment (which introduced a reduction of standard therapy from 8 to 6 cycles of eBEACOPP in total) were combined into one group for comparative purposes with the group that received only 4 cycles.
†Calculated by AAI.
‡Also including urogenital tract, cardiac and skin disorders, drug fever, and allergy.
§Includes, nausea or vomiting, mucositis, gastrointestinal tract, respiratory tract, and nervous system disorders, as well as urogenital tract, cardiac and skin disorders, drug fever, and allergy.
**excluding blood or bone marrow events and laboratory events

4.3.2.2.2. PET-adapted treatment: results from observational studies

A total of eight observational studies evaluating PET-adapted treatment strategies in adult patients with advanced-stage HL that met our inclusion criteria were identified.[39,58,60,78,95,124,133,169] Patient characteristics, treatments and methods across studies varied substantially making conclusions across studies difficult. Detailed data on the impact of specific PET-adapted treatments were limited. Small numbers of patients with positive interim PET likely compromised the ability to fully evaluate the potential impact of escalated therapies. The value of interim PET for directing therapy across these observational studies in unclear.

Escalated therapy

Five studies[39,78,124,133,169] focused on the utility of interim PET to guide therapy escalation for patients with a positive PET scan after 2 cycles of ABVD (4 studies) or VABEM (1 study). Patients with a negative interim PET scan after 2 cycles received additional courses of the induction therapy. No study compared escalated therapy to standard therapy for interim PET+ patients.

Only one study[39] reported OS for patients who were PET+ and received escalated therapy versus those who were PET- (continued with baseline treatment); they reported similar 5-year outcomes: 92% versus 88%, respectively. The remaining three studies reported 2-year OS for the whole population only (Table 34).

Progression-free (PFS) or Event-free (EFS) was reported by five studies (Table 34). Across four studies,[78,124,133,169] in patients with interim PET+ findings who received escalated therapy, the probability of 2-year (3 studies) and 5-year (1 study) PFS was lower than those with an interim PET– (these patients continued with baseline treatment). One additional study,[39] which specifically included patients with very poor prognoses than the other studies whose had a more intensive front-line therapy (VABEM) found that therapy escalation to early salvage treatment in patients with an interim PET+ resulted in similar 5-year outcomes (both OS and EFS) as interim PET– patients, which may suggest that PET-adapted escalation in such patients is of benefit. It is unclear to what extent survival would be improved in the absence of escalated treatment.

Only one study provided comparative data related to toxicity.[124] The risk of grade 4-5 toxic events was significantly greater in patients whose treatment was escalated (PET+) versus not-escalated (PET–), as were treatment-related deaths and secondary malignancies.

De-escalated therapy

Two studies reported the use of interim PET to guide the de-escalation of therapy for patients with a negative PET scan after 2 cycles of escalated BEACOPP. Patients with a positive interim PET scan after 2 cycles received additional courses of the induction therapy in one study,[60] while in the other study they received the same treatment as the PET- patients (ABVD). [95]

Five-year overall survival (OS) differed across the studies (Table 34). In the study in which patients with an interim PET+ result continued the baseline escalated therapy,[60] the probability of OS was 100% for both the PET- (de-escalated therapy) and the PET+ groups, however only 9 patients and a positive interim PET. In the study in which patients with an interim PET+ result received the same de-escalated

therapy as the PET- patients, patients who showed early response on interim PET (i.e., were PET-) had a significantly greater probability of 5-year OS (98% vs. 79%)[95]; however small number of PET-positive patients were available.

PFS in patients with interim PET– findings who received de-escalated therapy was more common than those with interim PET+ findings (range, 80%–87% vs. 47%–60%), Table 34.

Toxicity-related adverse events were reported by one study.[95] Grade 3/4 acute hematological toxicity was more frequent during the first two cycles of escalated BEACOPP compared to the subsequent ABVD treatment period (85.7% vs. 36.7%, p<0.001).

Escalated and De-escalated therapy

One observational study reported the use of interim PET to guide both escalation and de-escalation of therapy.[58] Patients were given one of two baseline therapies based on their IPS prognosis score (patients with a worse prognosis got a more intensive baseline treatment compared with those at average/low risk). All patients (regardless of prognostic score) with an interim PET negative result went on to receive the less intensive of the two treatments while those with an interim PET-positive result received the more intensive treatment. Five-year OS did not differ significantly between the two groups, with the interim PET-negative patients showing somewhat better outcomes , however PFS in those receiving de-escalated therapy was slightly better versus those with positive interim PET who received escalated therapy (Table 34).

Toxicity and other treatment-related adverse events were not reported separately for the advanced-stage HL subgroup or for the PET-positive and PET-negative groups.

Table 34. Advanced-Stage HL (stage IIB-IV): Probability of survival from observational studies of PET-adapted therapy following interim PET

		Author	N PET +	PET + Estimate (95%CI)	N PET –	PET – Estimate (95%CI)	Overall or comparative if given (95%CI); notes
Escalated therapy							
OS	2-year	Press 2016	60	NR	271	NR	98% (NR)
		Zinzani 2016	101	NR	409	NR	97% (94%-98%)
		Ganesan 2015	8	NR	41	NR	87.7% (81.6%–93.8%)
	5-year	Carras 2018	11	91.7% (77.3%–100%)	37	88.2% (78%–99.8%)	NS difference PET + vs PET -
PFS or EFS	2-year	Press 2016	60	64% (50%–75%)*	271	82% (77%–86%)	NR
		Zinzani 2016	101	76% (66%–84%)	409	81% (76%–84%)	NR
		Ganesan 2015	8	50% (NR)†	41	82% (NR)	p=0.013
	5-year	Carras 2018	11	81.5% (61.1%–100%)	37	77.8% (65.3%–92.7%)	NS difference PET + vs PET –
		Romano 2018	21	40.1% (NR)	98	84.7% (NR)	PET + vs PET –, aHR 3.8 (1.7%–8.3%);
De-escalated therapy							
OS	5-year	Deau 2015	9	100% (NR)	55	100% (NR)	NA
		Kedmi 2016	17	79% (NR)	52	98% (NR)	PET + vs PET –, p= 0.015
PFS		Deau 2015	9	47% (NR)	55	87% (NR)	PET + vs PET –, p= 0.006
		Kedmi 2016	17	60% (NR)	52	80% (NR)	PET + vs PET –, p= 0.20
Escalated and De-escalated therapy							
OS	5-year	Dann 2017	27	91.4% (79.9%–102.9%)	158	97.8% (95.4%–100.3%)	NS difference PET + vs PET -
PFS			27	68.4% (49.8%–86.9%)	158	80.8% (74%–87.7%)	PET + vs PET –, HR 2.03 (0.92-4.5), p= 0.08

aHR = adjusted hazard ratio; CI = confidence interval; EFS = event-free survival; NR = not report; OS = overall survival; PET = positron emission tomography; PFS = progression-free survival.
*Authors indicate that their reported 2-year PFS for PET2-positive patients is much higher than the expected 2-year PFS of 15% to 30% (authors reference two older studies for this statistic).
†Authors indicate that escalated therapy improved outcomes in PET-2-positive patients compared with historical data.

4.3.2.2.3. Other observational studies

No other nonrandomized comparative observational studies in patients with advanced-stage HL that met our inclusion criteria were identified.

4.3.2.3. Refractory/Relapsed HL

Studies included

Three observational studies across three cohorts and 3 publications[109-111] (from the same author group) were identified. Because the authors describe pre-high dose chemotherapy (HDT)/autologous stem cell transplantation (ASCT) PET scan as an interim PET assessment, we have included these studies in this section. All three studies escalated therapy upon positive interim PET results after two or three cycles of initial therapy and assessed the value of restaging/pre-transplant PET for adapting salvage therapy in autologous stem cell transplant-eligible persons with relapsed or refractory HL. Across all three studies, PET was employed to allocate patients who were PET negative directly to consolidation with HDT and ASCT, with the potential for RT, versus using extended salvage chemotherapy followed later by HDT, RT and ASCT for those with positive PET findings. Initial therapy regimens across these three studies included BV, ICE or augICE, while extended salvage therapy consisted of GVD or augICE. PET positivity varied across studies, ranging from 38% to 71% of patients at interim staging. Ages were similar with median ages between 33 and 35, with study populations of 65 to 97 patients, with between 52% and 58% females. Studies were considered to be at moderately high risk of bias; it is unclear whether PET/CT evaluation was blinded and different lengths of follow up were reported for the cohorts described in the 2015 and 2017 publications.

In two of the studies[109,110], a 5-PS score of ≤3 was initially used as threshold for PET negativity to avoid under treatment, however, authors also report some results using a score ≤2. Authors describe the proportions of patients achieving PET negative status at various stages and at the end as a way of monitoring treatment response but provide limited data on the primary outcomes in answer to the key questions for this HTA.

One publication[110] described a cohort that was later detailed more fully in another publication[109] where the data were combined with a newly enrolled cohort and analyzed further. In brief, patients in the first cohort received BV 1.2 mg/kg on days 1, 8, and 15 every 28 days for 2 cycles followed by evaluation by PET. Those who achieved PET normalization (defined by Deauville score of ≤2) proceeded directly to ASCT; those with persistent abnormalities on PET received 2 cycles of augICE before being considered for ASCT. Patients with localized, nodal-based disease who had not previously received radiation were treated with involved field radiation therapy (IFRT) before conditioning for ASCT. After cohort 1 enrollment was completed, enrollment proceeded for cohort 2 for which the eligibility and treatment were the same except that patients received three cycles of weekly BV rather than two.

No RCTs were identified that evaluated interim PET-adapted therapy for relapsed or refractory HL that met our inclusion criteria.

4.3.2.3.1. PET-adapted treatment: results from observational studies

Overall and progression free survival and toxicity

Evaluation across the three cohorts provide evidence for restaging/pre-treatment PET for adapting salvage therapy in persons with relapsed or refractory HL who were eligible for autologous stem cell transplant is limited. Pre-transplant PET results may provide valuable prognostic information; survival probability was lower in those with a positive pre-transplant PET.

Based on limited data provided for the first cohort (N = 97[111]), PET/CT may help distinguish persons who may be able to progress to HDT/ASCT if results are negative from those who may require extended savage therapy prior to HDT/ASCT; pre-transplant PET results provide prognostic information. Patients with a pre-transplant positive FDG-PET received GVD (extended salvage therapy) prior to HDT/ASCT; those with a negative pre-transplant PET underwent HDT/ASCT. EFS for negative FDG-PET, pre-HDT/ASCT after 1 or 2 cycles of salvage therapy was > 80%, versus 28.6% for patient who had a positive scan (P < .001); no data comparing the impact PET-adapted strategies on toxicity were provided. Authors conclude that goal of salvage therapy in patients with Hodgkin lymphoma should be a negative FDG-PET scan before HDT/ASCT. Across all patients, OS and EFS over all were 80% and 70% respectively. Among persons receiving transplant OS and EFS were 88% and 79%.

Two publications[109,110] provide information on separate populations from the 2012 report. The 2015 publication reports on an initial cohort of patients (Cohort 1, N = 45). The 2017 publication adds an additional cohort of patients (Cohort 2, N = 20, all of which progressed to ASCT). Authors report that there were no significant differences in response for the two cohorts so results were combined for evaluation of prognostic factors and of different methods of PET positivity and quantitation; however, the median follow up for cohort 1 was substantially longer (40 months) than that for cohort 2 (20 months). Authors conclude that the PET-adapted therapy of BV and augICE was associated with a complete remission rate of 75%.

Based on a 5-PS score of ≤3 as a negativity threshold, data derived from author's figure, PFS for those with pre-ASCT PET positive results was significantly less (60%) at 3 years compared with those having a with pre-ASCT PET negative result, estimated at 85% from author's figure (p = 0.05) in persons who had ASCT. Authors conclude that pre-ASCT PET is predictive of PFS. Across patients, 76 % (95% CI 62–89) achieved PET-negative status at the end of all treatments reported. Treatment-related toxicities were described in the 2015 publication but were not reported based on PET-adapted treatments received. Overall, the reported frequency of grade 3 or 4 toxicities was ≤2%.

4.3.2.4. Aggressive NHL

Studies included

Two randomized studies evaluating interim PET/CT in persons with aggressive NHL that met inclusion criteria were identified.[41,63] One randomized trial[63] of PET-adapted therapy (N = 363 randomized) in patients aNHL was identified. PET/CT was done after two cycles of CHOP or R-CHOP (rituximab use was restricted to patients with CD20 positive lymphomas), persons with positive interim PET were randomly

assigned to receive six additional cycles of R-CHOP or six blocks of an intensive Burkitt's lymphoma protocol. Individuals with a negative interim PET with CD20 positive lymphomas were randomly assigned to receive or allocated to receive four additional cycles of R-CHOP or the same treatment with two additional doses rituximab. Patients with B-Cell lymphoma comprised the majority of the population (86.4%) with diffuse large B-cell lymphoma (DLBCL) the most common form (82.7%). PET positivity was assessed based on evaluation of ΔSUVmax method; a positive interim PET was defined as a scan in which the SUVmax was reduced by ≤66% of the baseline SUVmax. Interim scans demonstrating a SUVmax reduction >66% were considered negative. In patients with a low baseline SUV max, interim PET scans without unphysiological FDG uptake, as judged by visual criteria, were regarded as negative, even if the percentage reduction was ≤66%. This modification of the ΔSUVmax method allowed for the fact that, in patients with low baseline SUVmax, a return to physiological activity may require less than a 66% reduction. Recruitment was terminated early because PET-positivity was rare and the recruitment goal was unlikely to be reached within a reasonable time; in addition, the Burkitt protocol was more toxic and possibly less effective than R-CHOP.

The LNH2007-3B trial randomized persons with CD20 positive DLBCL (N = 222)[41] to different induction therapies (R-ACVBP or R-CHOPP). Interim PET was done after 2 and 4 treatment cycles. Results from both interim PET scans were used to adapt treatment (Table 36) but no further randomization was done for treatment based on PET findings. Treatment was escalated in all patients with positive PET at 2 and/or 4 cycles. Individuals with negative PET results at both time frames, received standard immunochemotherapy (SIC) consolidation, where as those with positive PET after cycle 2 received escalated therapy including autologous stem cell transplantation. Treatment of participants with positive PET results after 4 cycles of therapy was at the discretion of the local investigator. Based on the IHP criteria PET positivity was defined as "clearly increased activity relative to the reference background" of at least 25% higher. Authors also explored the impact of using ΔSUVmax, concluding that it provided more accurate evaluation of PET-positivity.

A third RCT (LYSA/GOELAMS trial) of PET-adapted therapy in adults (aged 18 to 75 years) with early-stage/low-risk aNHL that met inclusion criteria was identified.[101] All patients (N=334) had Ann Arbor stage I or II limited DLBCL, with nonbulky mass (tumor <7 cm in diameter). Males comprised 60% of the population. After initial staging, patients were randomized to receive R-CHOP only or R-CHOP plus involved-field radiotherapy (IFRT). Interim PET evaluation was done after all patients had completed four cycles of R-CHOP. Patients with a negative PET result continued with therapy as randomized; those with no adverse risk factors received a total of 4 cycles of R-CHOP (i.e., no further chemotherapy post-PET) while those with at least one risk factor received a total of 6 cycles of R-CHOP (i.e. two additional cycles of chemotherapy post-PET) followed by observation/no IFRT (R-CHOP only group) or IFRT (R-CHOP + IFRT group). Patients with a positive PET result were recommended to receive two additional cycles followed by IFRT, no matter which treatment arm they were allocated to; however only 71% received this treatment with others receiving two additional R-CHOP cycles without IFRT, high-dose chemotherapy, and no further treatment. PET was assessed visually based on ^{18}F-FDG uptake. A negative scan was defined as having no abnormally increased ^{18}F-FDG at any site; a positive scan was defined as having ^{18}F-FDG uptake above the mediastinum or surrounding background in a location incompatible with normal anatomy or physiology. This trial was considered to be moderately low risk of bias since it was unclear if the authors accounted for time at risk between groups.

4.3.2.4.1. PET-adapted treatment: results from RCTs

Survival and treatment-related toxicity

Across the two trials evaluating advanced-stage/high-risk aNHL, conclusions regarding the value of interim PET differ, likely due to differences in patient populations and effectiveness of treatment escalation strategies.

In the PETAL PET-adapted treatment trial,[63] the probability of OS, PFS and EFS at 2-years was lower in those with positive interim PET who received the Burkitt's Lymphoma protocol versus those who received R-CHOP (Table 35). The estimated differences of -7.9/100, - 6.3/100 and -10.5/100 respectively were seen. Differences between groups may be clinically significant but failed to reach statistical significance; authors concluded that switching from R-CHOP to the Burkitt protocol did not improve outcomes but power to detect survival differences may have been compromised by not reaching their recruitment goal. Grade 3 or 4 treatment-related toxicity for patients receiving the Burkitt's protocol was substantially greater compared with those receiving R-CHOP (Table 37). Compared with the R-CHOPP group, the Burkitt protocol group experienced significantly more anemia, leukopenia, thrombocytopenia, infection and mucositis with risk differences ranging from 20% to 38%.

In patients in the PETAL trial who had a negative interim PET, those who received four R-CHOP cycles had somewhat lower OS, PFS and EFS compared with those who received an additional two rituximab doses with the R-CHOP (estimated differences, 1/100, 4.5/100 and 2.9/100) but statistical significance was not reached.[63] Authors concluded that increasing rituximab exposure in patients with a negative interim PET did not impact income.

Based on this trial, the value of interim PET is unclear; there was no evidence of improved outcome in either the PET positive or the PET negative patients based on PET-adapted treatments received. PET positive patients had considerably worse prognosis than PET-negative patients despite receiving additional R-CHOP cycles. In the absence of a head to head comparison demonstrating improved survival with 6 R-CHOP cycles versus 4 cycles in PET- positive patients or no difference in PET-negative patients, the value of knowing a patient's PET status is not clear,

In the LNH-3B trial at 4 years,[41] the probability of OS was similar between those who received escalated therapy and those who had standard immunochemotherapy (SIC) consolidation (difference 0.8/100) (Table 36). PFS was better in those receiving escalated therapy; the estimated difference of 10/100 may be clinically significant even though statistical significance was not reached. Data comparing toxicity of PET-adapted treatment alternatives was not provided. Authors note that using the IHP criteria versus more quantitative Δ SUVmax may influence ability to direct treatment by identifying those who may benefit most from different treatment strategies. They report that the proportion of patients who had negative PET after 2 and 4 cycles was 26% using IHP criteria, and increased to 79% using Δ SUVmax.

Compared with the PETAL trial,[63] the PET positive patients in the LNH-3B trial[41] who underwent escalated treatment had somewhat similar prognosis as well as PET negative patients; whereas the expectation in the absence of escalated treatment was that they would do more poorly. It may be that the treatments in the LNH trial were more effective than those studied in the PETAL trial. Differences in patient populations, use of different methods for determining PET positivity, and length of follow-up may also partially explain differences in results.

Table 35. Aggressive NHL (aNHL): Probabilities of overall and progression-free survival from the RCT by Duhrsen et al. of PET-adapted therapy following interim PET – escalation

Outcome	Duhrsen 2018[63] *PETAL Trial* Initial therapy: R-CHOP, 2 cycles					
	PET + Group (randomized) (n=108) (ΔSUV_{max} method)			PET – Group (randomized) (n=255) (ΔSUV_{max} method)		
	Intervention: 6 cycles of Burkitt protocol† (n=56)	Comparator: 6 cycles of R-CHOP (n=52)	Comparator vs. Intervention	Intervention: 4 cycles of R-CHOP + 2 cycles rituximab (n=126)	Comparator: 4 cycles of R-CHOP (n=129)	Comparator vs. Intervention
All patients						
2-year OS	55.4% (40.7%–67.8%)	63.6% (48.5%–75.3%)	HR 1.3 (95% CI 0.8–2.4), p=0.31	87.2% (79.9%–91.9%)	88.2% (81.2%–92.7%)	HR 0.9 (95% CI 0.5–1.5), p=0.64
2-year PFS	43.1% (29.2%–56.2%)	49.4% (34.7% –62.4%)	NR	77.5% (69.1%–83.9%)	82.0% (74.2%–87.7%)	NR
2-year EFS	31.6% (19.3%–44.6%)	42.0% (28.2%–55.2%)	HR 1.5 (95% CI 0.9–2.4), p=0.09 Adj. HR‡ 1.5 (95% CI 0.9–2.5), p=0.12	73.5% (64.8%–80.4%)	76.4% (68.0%–84.2%)	HR 1.1 (95% CI 0.7–1.6), p=0.82 Adj. HR‡ 1.0 (95% CI 0.7–1.6), p=0.83
DLBCL only	n=31	n=32		n=100	n=97	
2-year OS	47.1% (28.0%–64.2%)	64.8% (45.5%–78.8%)	HR 1.5 (95% CI 0.7–3.2), p=0.24	85.8% (77.3%–91.4%)	88.5% (80.1%–93.4%)	HR 0.96 (95% CI, 0.52 to 1.8), p=0.89
2-year PFS	41.4% (23.4%–58.6%)	55.5% (36.7%–70.8%)	NR	77.7% (68.1%–84.7%)	82.3% (73.0%–88.6%)	NR
2-year EFS	28.3% (13.4%–45.4%)	52.4%§ (33.8%–68.0%)	HR 2.1 (95% CI 1.1–4.0), p=0.02	72.6% (62.7%–80.3%)	78.9% (69.3%–85.9%)	HR 1.2 (95% CI, 0.8 to 2.0), p=0.40
T-cell lymphomas only§	n=10	n=9			n=57, 4 cycles CHOP	
2-year OS	30.0% (7.1%–57.8%)	22.2% (3.4%–51.3%)	NR	NA	77.2% (63.3%–86.3%)	NA
2-year PFS	30.0% (7.1%–57.8%)	12.7% (0.7%–42.7%)	NR	NA	63.3% (48.7%–74.8%)	NA
2-year EFS	30.0% (7.1%–57.8%)	12.7% (0.7%–42.7%)	NR	NA	63.3% (48.7%–74.8%)	NA

ABVD =doxorubicin, bleomycin, vinblastine, and dacarbazine; CI = confidence interval; HR = hazard ratio; IFRT = involved-field radiotherapy; INRT = involved-node radiotherapy; NR = not report; OS = overall survival; PET = positron emission tomography; PFS = progression-free survival;

*Rituximab was restricted to CD20-positive lymphomas

†The authors state that Burkitt protocol is an intensive methotrexate and cytarabine-based regimen that is routinely used in pediatric non-Hodgkin lymphomas and adult Burkitt's lymphoma; no other details are provided.

‡Adjusted for stratification factors; no other details provided.

§For T-cell lymphomas, only patients with positive interim PET findings underwent random assignment. Patients with negative interim PET findings uniformly received CHOP. Patients with CD20-positive T-cell lymphomas also received rituximab

Table 36. Aggressive NHL (aNHL): Probabilities of overall and progression-free survival from the LNH-3B RCT by Casasnovas et al of PET-adapted therapy following interim PET negative results.

Author (year) Study Design	N; Lymphoma subtype; Initial treatment/ cycles prior to interim PET; F/U; PET criteria	Treatment received according to interim PET (+) or PET (-) status	Primary Outcomes	
			OS (95% CI)	PFS (95% CI)
Casasnovas 2017 *GELA/LYSA trial, LNH2007-3B trial*[41] RCT	N=211 High-risk DLBCL *Induction treatment Randomization:* 1. R-ACVBP, 4 cycles (n=109) 2. R-CHOP14, 4 cycles (n=102) Median 45 months PET interpreted according to IHP criteria; PET+ defined as "clearly increased activity relative to the reference background"	**PET (+) cycle 2/PET (–) cycle 4, n=48**: 2 cycles of high-dose Methotrexate, then BEAM or Z-BEAM followed by ASCT	4-year: • *Overall (n=48):* 90.4% (81%-95.1%) • *R-ACVBP group (n=30):* 96.4% (77.2%-99.5%) • *R-CHOP group (n=18):* 88.2% (60.6%-96.9%)	4-year: • *Overall (n=48):* 85% (71.1%-92.6%) • *R-ACVBP group (n=30):* 82% (63.6%-92.5%) • *R-CHOP group (n=18):* 88.5% (61.4%-97%)
		PET (–) cycles 2 and 4, n=52: 1. *R-ACVBP group*: 2 cycles high-dose methotrexate, and then 4 cycles rituximab, ifosfamide, etoposide and 2 cycles cytarabine 2. *R-CHOP group*: 4 additional cycles of R-CHOP	4-year: • *Overall (n=52):* 89.6% (85%-92.2%) • *R-ACVBP group (n=28):* 90% (72.1%-96.7%) • *R-CHOP group (n=24):* 83.3% (61.5%-93.4%)	4-year: • *Overall (n=52):* 75% (60.9%-84.6%) • *R-ACVBP group (n=28):* 78.6% (58.4%-89.8%) • *R-CHOP group (n=24):* 70.8% (48.4%-84.9%)
			PET (+) cycle 2/PET (–) cycle 4 (n=48) vs. PET (–) cycles 2 and 4 (n=52) • *Overall (n=100)*: p=0.21 o *R-ACVBP group (n=58):* p=0.35 o *R-CHOP group (n=42):* p=0.69	**PET (+) cycle 2/PET (–) cycle 4 vs. PET (–) cycles 2 and 4** • *Overall (n=100)*: p=0.26 o *R-ACVBP group (n=58):* p=0.71 o *R-CHOP group (n=42):* p=0.24
		PET (–) cycle 4 (i.e., regardless of cycle 2 PET), n=100: combined analysis of the 2 groups above	4-year: • *Overall (n=100):* 88.9% (82.1%-94.4%) • *R-ACVBP group (n=58):* 93.1% (82.7%-97.4%)	4-year: • *Overall (n=100):* 79.8% (79.4%-86.4%) • *R-ACVBP group (n=58):* 80.8% (68%-88.9%)

Author (year) Study Design	N; Lymphoma subtype; Initial treatment/ cycles prior to interim PET; F/U; PET criteria	Treatment received according to interim PET (+) or PET (-) status	Primary Outcomes	
	should be at least 25% higher than this background		• *R-CHOP group (n=42):* 85.4% (70.3%-93.1%)	• *R-CHOP group (n=42):* 78.2% (62.3%-88%)
		PET (+) cycle 4, n=100: final treatment decision left to local investigator (? Unclear if escalated or standard therapies provided)?	4-year: • *Overall (n=100):* 80% (69%-87.5%) • *R-ACVBP group (n=45):* 79.9% (65%-89%) • *R-CHOP group (n=55):* 80.1% (62.4%-90.1%)	4-year: • *Overall (n=100):* 72.9% (63.1%-80.6%) • *R-ACVBP group (n=45):* 70.8% (55.1%-81.9%) • *R-CHOP group (n=55):* 74.5% (60.8%-84.1%)
			PET (-) cycle 4 (n=100) vs. PET (+) cycle 4 (n=100): • *Overall (n=200)*: p=0.08 o *R-ACVBP group (n=103):* p=0.054 o *R-CHOP group (n=97):* p=0.68	**PET (-) cycle 4 (n=100) vs. PET (+) cycle 4 (n=100):** • *Overall (n=200)*: p=0.16 o *R-ACVBP group (n=103):* p=0.17 o *R-CHOP group (n=97):* p=0.511
		All Patients, n=211 (includes 11 early withdrawals not included above)	4-year: • *Overall (n=211):* 83% (76.8%-87.7%) • *R-ACVBP group (n=109):* 85.3% (77.1%-90.7%) • *R-CHOP group (n=102):* 80.7% (70.4%-87.7%)	4-year: • *Overall (n=211):* 74.6% (68.1%-80%) • *R-ACVBP group (n=109):* 75% (65.6%-82.1%) • *R-CHOP group (n=102):* 74.2% (64.4%-81.6%)

Table 37. Aggressive NHL (aNHL): RCTs of PET-adapted therapy following interim PET – treatment-related mortality and toxicity

Outcome	Duhrsen 2018[63] *PETAL Trial* Initial therapy: R-CHOP, 2 cycles					
	PET + Group (randomized) (n=108) (ΔSUV_{max} method)			PET – Group (randomized) (n=255)* (ΔSUV_{max} method)		
	Intervention: 6 cycles of Burkitt protocol†	Comparator: 6 cycles of R-CHOP	Intervention vs. Comparator RR/RD (95% CI)‡ p-value§	Intervention: 4 cycles of R-CHOP	Comparator: 4 cycles of R-CHOP + 2 cycles rituximab	Intervention vs. Comparator RR/RD (95% CI)‡ p-value§
Treatment-related mortality	5.4% (3/56)	3.9% (2/52)	RR 1.4 (0.2–8.0); RD 2% (-6% to 9%); p=NS	3.9% (5/129)	1.6% (2/126)	RR 2.4 (0.5–12.4); RD 2% (-2% to 6%); p=NS
Grade 3 or 4 Toxicity-related Adverse Events						
Anemia	44.6% (25/56)	25.0% (13/52)	RR 1.8 (1.0–3.1); RD 20% (2%–37%); p=0.04	16.3% (21/129)	11.1% (14/126)	RR 1.5 (0.8–2.8); RD 5% (-3% to 14%); p=NS
Leukopenia	80.4% (45/56)	59.6% (31/52)	RR 1.3 (1.0–1.7); RD 21% (4%–38%); p=0.02	54.3% (70/129)	60.3% (76/126)	RR 0.9 (0.7–1.1); RD -6% (-18% to 6%); p=NS
Neutropenia	33.9% (19/56)	23.1% (12/52)	RR 1.5 (0.8–2.7); RD 11% (-6% to 28%); p=NS	15.5% (20/129)	23.8% (30/126)	RR 0.7 (0.4–1.1); RD -8% (-18% to 1%); p=NS
Thrombocytopenia	58.9% (33/56)	21.2% (11/52)	RR 2.8 (1.6–4.9); RD 38% (21%–55%); p<0.001	14.7% (19/129)	7.9% (10/126)	RR 1.9 (0.9–3.8); RD 7% (-1% to 15%); p=NS
Infection	50.0% (28/56)	21.2% (11/52)	RR 2.4 (1.3–4.2); RD 29% (12%–46%); p=0.002	14.7% (19/129)	11.1% (14/126)	RR 1.3 (0.7–32.5); RD 4% (-5% to 12%); p=NS
Mucositis	37.5% (21/56)	11.5% (6/52)	RR 3.3 (1.4–7.4); RD 26% (11%–41%); p=0.003	1.6% (2/129)	2.4% (3/126)	RR 0.7 (0.1–3.8); RD -1% (-4% to 3%); p=NS
Diarrhea	5.4% (3/56)	9.6% (5/52)	RR 0.6 (0.1–2.2); RD -4% (-14% to 6%); p=NS	3.1% (4/129)	0.8% (1/126)	RR 3.9 (0.4–34.5); RD 2% (-1% to 6%); p=NS

Outcome	Duhrsen 2018[63] *PETAL Trial* Initial therapy: R-CHOP, 2 cycles					
	PET + Group (randomized) (n=108) (ΔSUV_{max} method)			PET – Group (randomized) (n=255)* (ΔSUV_{max} method)		
	Intervention: 6 cycles of Burkitt protocol†	Comparator: 6 cycles of R-CHOP	Intervention vs. Comparator RR/RD (95% CI)‡ p-value§	Intervention: 4 cycles of R-CHOP	Comparator: 4 cycles of R-CHOP + 2 cycles rituximab	Intervention vs. Comparator RR/RD (95% CI)‡ p-value§
Creatinine	1.8% (1/56)	5.8% (3/52)	RR 0.3 (0.03–2.9); RD -4% (-11% to 3%); p=NS	3.1% (4/129)	0.8% (1/126)	RR 3.9 (0.4–34.5); RD 2% (-1% to 6%); p=NS

ABVD =doxorubicin, bleomycin, vinblastine, and dacarbazine; CI = confidence interval; eBEACOPP = escalated bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine and prednisone; HR = hazard ratio; IFRT = involved-field radiotherapy; INRT = involved-node radiotherapy; NC = not calculable; NR = not report; OS = overall survival; PET = positron emission tomography; PFS = progression-free survival; RD = risk difference; RR = relative risk.

*Rituximab was restricted to CD20-positive lymphomas

†The authors state that Burkitt protocol is an intensive methotrexate and cytarabine-based regimen that is routinely used in pediatric non-Hodgkin lymphomas and adult Burkitt's lymphoma; no other details are provided.

‡Calculated by AAI.

§Per study authors.

A third trial evaluated the use of interim PET to guide therapy de-escalation in patients with early-stage/low risk DLBCL, specifically whether involved-field radiotherapy (INFT) could be safely omitted in patients with a negative PET result (in complete remission) after four cycles of R-CHOP. In the LYSA/GOELAMS PET-adapted treatment trial,[101] the probability of OS at 5 years was similar between those who did not and those who did receive INFT (90% vs. 92%, respectively), as was the frequency of relapse (7.3% vs. 6.9%; RD 1%, 95% CI -5% to 7%). These results are consistent with the intention-to-treat (ITT) analyses that included all randomized patients regardless of their interim PET status (Table 38). Event-free survival (EFS) at 5 years and mortality were not reported separately for those with a negative interim PET but the ITT analyses showed no difference between groups. It is unclear whether these difference are clinically important. Treatment-related toxicity was not reported separately by group. The authors conclude that patients who reach complete remission (i.e., are PET-negative) after induction therapy could be spared additional RT, avoiding radiation-related toxicity.

Table 38. Aggressive NHL (aNHL): Results from the RCT by Lamy et al. of PET-adapted therapy following interim PET – *de-escalation*

Outcome	Lamy 2018 *LYSA/GOELAMS Trial* *PET assessed visually**					
	PET – after 4 cycles of R-CHOP (n=281), randomized to:			ITT results (all randomized patients) (n=319)†		
	Intervention: R-CHOP only/no RT (n=137)‡	Comparator: R-CHOP + RT (n=144)§	Intervention vs. Comparator**:	Intervention: R-CHOP only/no RT (n=159)	Comparator: R-CHOP + RT (n=160)	Intervention vs. Comparator**:
5-year OS (±SD) (95% CI)	90% (NR)	92% (NR)	NS	92% ± 2.5%	96% ± 1.7%	HR 0.62 (95% CI 0.3–1.5), p=0.28
5-year EFS (±SD) (95% CI)	NR	NR	NR	89% ± 2.9%	92% ± 2.4%	HR 0.61 (95% CI 0.3–1.2), p=0.18
Mortality	NR	NR	NR	7.5% (12/159)	5.6% (9/160)	RR 1.3 (95% CI 0.6–3.1) RD 2% (95% CI -4% to 7%)
Relapses	7.3% (10/137)	6.9% (10/144)	RR 1.2 (95% CI 0.5–2.7) RD 1% (95% CI -5% to 7%)	8.2% (13/159)	6.3% (10/160)	RR 1.3 (95% CI 0.6–2.9) RD 2% (95% CI -4% to 8%)
Secondary malignancy	NR	NR	NR	1.3% (2/159)	1.3% (2/160)	NS

CI = confidence interval; CR = complete remission; EFS = event-free survival; HR = hazard ratio; RT = radiotherapy; NR = not report; OS = overall survival; PET = positron emission tomography; PR = partial remission.

*A positive PET scan was defined visually as ^{18}F-fluorodeoxyglucose (^{18}F-FDG) uptake above the mediastinum or surrounding background in a location incompatible with normal anatomy or physiology. A negative PET scan was defined as having no abnormally increased ^{18}F-FDG at any site.

†ITT results include 38 patients who were PET + after 4 cycles of R-CHOP and were not treated according to initial randomization (they went on to receive a variety of treatments).

‡Patients either underwent observation (n=76) or received 2 more cycles of R-CHOP (n=61) (for a total of 4 or 6 cycles, respectively).

§Patients either proceeded straight to RT (n=82) or received 2 more cycles of R-CHOP before receiving RT (n=62) (for a total of 4 or 6 cycles prior to RT).

**As reported by authors unless otherwise indicated

4.3.2.4.2. PET-adapted treatment: results from observational studies

A total of three observational studies[116,139,143] evaluating interim PET-adapted treatment strategies in adult patients with aggressive non-Hodgkin (aNHL) that met our inclusion criteria were identified. No studies specifically evaluating children with aNHL that met our inclusion criteria were identified. All three studies evaluated the utility of interim PET to guide therapy escalation for patients with a positive PET scan after 2-4 cycles of R-CHOP (2 studies)[139,143] or 3 cycles of R-MegaCHOP (1 study)[116]. Patients with a negative interim PET scan after 2 cycles received additional courses of the induction therapy. No observational studies evaluating the role of interim PET to guide therapy de-escalation in patients with aNHL were identified.

Overall survival (OS) was reported by all three studies (Table 38). In patients with interim PET+ findings who received escalated therapy, the probability of 2-year (1 study), 3-year (3 studies), and 4-year (1 study) OS was lower than those with an interim PET– (these patients continued with baseline treatment). At 3-years across studies, OS ranged from 68% to 73% for the PET+ group versus 71% to 93% for the PET– group.

Progression-free (PFS) or Event-free (EFS) was reported by all three studies (Table 39). Two of the three trials[116,143] reported lower probabilities of PFS or EFS across 2- to 4-years for patients with interim PET+ findings who received escalated therapy compared with those with an interim PET– result (continued with baseline treatment): 33%–57% versus 71%–81%. In one study[139], the probability of 3-year PFS was greater in the PET+ group (escalated therapy) versus PET- group (65% vs. 53%). In the latter study, the escalated therapy given to the PET+ patients was very different from therapies employed in the other studies in that RBEAM/autologous stem cell transplant was a component.

The impact of early escalated therapy in these patient populations is unclear. No study compared escalated therapy to standard therapy for interim PET+ patients and therapy regimens differed substantially across studies as did patients characteristics.

Two studies[139,143] provided comparative data related to toxicity. The incidence of toxicity-related adverse events was greater in patients whose treatment was presumably escalated (i.e. in PET+) versus continuation of induction therapy (PET–), as were treatment-related deaths

Table 39. Aggressive NHL: Survival probabilities from observational studies of PET-adapted therapy following interim PET

		Author	N PET +	PET + (escalated therapy) Estimate (95%CI)	N PET –	PET – Estimate (95%CI)	Overall or comparative if given (95%CI); notes
Escalated therapy							
OS	2-year	Swinnen 2015[143]	13	77% (90% CI 51%–90%)	63	93% (90% CI 86%–97%)	p=NR for PET(+) vs. PET(-); overall: 90% (90% CI, 83%–94%)
	3-year	Pardal 2014[116]	30	73% (NR)	36	89% (NR)	PET(+) vs. PET(-): p=0.11
		Stewart 2014[139]	36	68.4% (NR)	34	70.5% (NR)	p=NR for PET(+) vs. PET(-); overall: 69.8% (56.9%–79.5%)
		Swinnen 2015[143]	13	69% (90% CI 43%–85%)	63	93% (90% CI 86%–97%)	p=NR for PET(+) vs. PET(-); overall: 89% (90% CI, 81%–93%)
				Range, 68% to 73%		Range, 71% to 93%	
	4-year	Swinnen 2015[143]	13	69% (90% CI 43%–85%)	63	90% (90% CI 81%–95%)	p=NR for PET(+) vs. PET(-); overall: 86% (90% CI, 78%–91%)
PFS or EFS	2-year	Swinnen 2015[143]	13	42% (90% CI 19%–63%)	63	76% (90% CI 65%–84%)	p=NR for PET(+) vs. PET(-); overall: 70% (90% CI, 60–78%)
	3-year	Pardal 2014[116]	30	57% (NR)	36	81% (NR)	p=0.023; adjusted RR 2.6 (1.02–6.65), p=0.044
		Stewart 2014[139]	36	65.2% (NR)	34	52.7% (NR)	p=NR for PET(+) vs. PET(-); overall: 59.3% (46.2%–70.2%)
		Swinnen 2015[143]	13	33% (90% CI 13%–55%)	63	71% (90% CI 59%–80%)	p=NR for PET(+) vs. PET(-); overall: 69% (90% CI, 59–77%)
				Range, 33% to 65%		Range, 53% to 81%	
	4-year	Swinnen 2015[143]	13	33% (90% CI 13%–55%)	63	71% (90% CI 59%–80%)	p=NR for PET(+) vs. PET(-); overall: 64% (90% CI, 53–73%)

aHR = adjusted hazard ratio; CI = confidence interval; EFS = event-free survival; NR = not report; OS = overall survival; PET = positron emission tomography; PFS = progression-free survival.

*Authors indicate that their reported 2-year PFS for PET2-positive patients is much higher than the expected 2-year PFS of 15% to 30% (authors reference two older studies for this statistic).

†Authors indicate that escalated therapy improved outcomes in PET-2-positive patients compared with historical data.

4.3.2.5. Indolent NHL

Studies included

Only one comparative observational study that met our inclusion criteria was identified that reported survival outcomes following interim assessment with PET/CT compared with CT alone in patients with indolent NHL (iNHL).[31] The population was comprised of children (median age 10.5 years) with primarily advanced-stage nonlymphoblastic lymphoma (82% B-cell; 18% T-cell). Interim response assessment was performed after two cycle of chemotherapy (1.5 to 2 months after beginning treatment). No significant differences were seen for the probability of overall (OS) and progression-free survival (PFS) according to PET/CT versus CT alone when patients with and without complete remission according to both modalities were compared, Table 40.

No interim PET-adapted trials or trials that compared treatment strategies based on use of PET with strategies that did not use PET in patients with iNHL were identified. No studies in adults met the inclusion criteria.

Table 40: Interim PET imaging of indolent non-Hodgkin lymphoma (iNHL): Survival outcomes from a clinical observational study

Study, year Study design	Population	Comparison	OS (95% CI)	p-value	PFS (95% CI)	p-value
Bakhshi 2012	Pediatric: median age 10.5 years Subtype: non-lymphoblastic (B-cell 82%; T-cell 18%) Stage: early (9%); advanced (91%)	**patients *with* complete remission on:**				
		PET/CT	78.6% (47.2%–92.5%); SE 11.0	P=0.47*	85.7% (53.9%–96.2%); SE 9.3	P=0.47
		CT	83.3% (48.2%–95.5%); SE 10.8		82.5% (46.1%–95.3%); SE 11.3	
		patients *without* complete remission on				
		PET/CT	50.4% (24.9%–71.4%); SE 12.5	P=0.47	57.0% (30.1%–76.9%); SE 12.4	P=0.47
		CT	49.7% (25.4%–70.0%); SE 11.9		62.3% (36.7%–80.0%); SE 11.3	

CI = confidence interval; CT = computed tomography; NR = not reported; PET = positron emission tomography; PFS = progression-free survival; SE = standard error of the estimate.
*For both OS and PFS, on the basis of the exact symmetry test, there was no significant discordance between interim contrast enhanced CT and PET/CT results.

4.3.2.6. Mixed Hodgkin (HL) and non-Hodgkin lymphoma (NHL)

Studies included

One comparative observational study that met our inclusion criteria was identified that reported the impact of interim PET/CT on radiation-treatment (RT) planning.[150] This study (N=95) included adult patients (median age 61 years) with NHL (74%) or HL (11%) (15% had other hematological malignancies) scheduled to receive RT; 53% of the patients had completed chemotherapy prior to enrollment while for the other 47% RT was a monotherapy.[150] Compared with planning using CT only, co-registration of PET resulted in the following modifications to RT therapy: six patients (6%) had RT cancelled, two patients

(2%) had additional RT sites added, and in 79 patients (89%) a clinically meaningful change in RT volumes (≥5% increased or decreased) was made.

No interim PET-adapted trials or trials that compared treatment strategies based on use of PET with strategies that did not use PET in mixed HL and NHL populations were identified. No studies in children met the inclusion criteria.

4.3.3. End-of-Treatment

Summary of results

Advanced-Stage Hodgkin Lymphoma (HL)

PET-Adapted treatment for advanced-stage HL: RCT

- One RCT[77] in adults with advanced-stage HL evaluated end-of-treatment (EOT) PET for direction of further therapy. No studies compared treatment strategies based on use of PET with strategies that did not use PET. All studies were in adult populations.
- The PET-adapted treatment trial randomized patients who had large nodal masses at baseline but whose interim and post-treatment PET/CTs were negative to consolidation radiation therapy versus no further treatment[77] (no radiation therapy). Similar probabilities for OS and PFS were seen for the treatment groups. PFS for the 260 patients with no LNM at baseline and not randomly assigned to RT were only slightly lower (92%, 95% CI 88% to 95%) with an estimated difference of 5/100 between those randomized to RT and those who were not randomized. In patients with large nodal masses at baseline, PET may help identify patients in whom consolidation RT may safely be eliminated based on negative results for both interim and end of treatment PET.
- Treatment-related toxicity for this set of treatments was not reported.

PET-Adapted treatment for advanced-stage HL: Observational

- PFS from a poor quality retrospective study[65,66] was similar between patients in partial remission with a negative EOT PET and those considered to be in complete remission after induction treatment: authors concluded that radiotherapy could be omitted in those with a negative PET in partial remission.

Aggressive Non-Hodgkin Lymphoma (aNHL)

PET-Adapted treatment for advanced-stage HL: Observational

- In a subset of 50 patients in the PETAL trial[63] who had positive interim PET findings, those with persistent abnormalities on EOT PET who had no clinical evidence of disease were assigned to observation and those and those who had clinical evidence of disease as well received additional immediate treatment. Both OS and PFS were better in the asymptomatic observation group

versus those with clinical evidence of disease. Prognosis in those with negative EOT PET was better compared with the full group of patients with persistent PET abnormalities. This may suggest the importance of considering clinical evidence of disease together with findings of persistent PET abnormalities at the end of treatment. The role of false positives and false negatives is unclear. Data on toxicity specific to EOT PET decisions were not provided.

- One retrospective cohort study[168] provided reported that disease-free survival at 10-years was similar for those with positive EOT PET (received RT) and negative EOT PET (observation) groups (91% versus 90%, respectively). Another observational study[116] reported that EOT PET was a strong predictor of outcome; compared with a PET+ result after completion of treatment, a PET- result was associated with a significantly greater probability of both OS (98% vs. 67%) and PFS (81% vs. 57%).

4.3.3.1. Advanced-Stage HL

Studies included

One PET-adapted RCT[77] and one observational study[65,66](2 publications) provided data on omission of radiotherapy based on PET-findings at the end of induction treatment.

4.3.3.1.1. PET-adapted treatment: results from RCTs

In the HD 0607 trial of PET-adapted therapy[77], PET/CT was performed after 2 ABVD induction treatment cycles and was described previously. Patients with positive interim PET results were randomly escalated to eBEACOPP with or without rituximab to complete first-line treatment. Patients with negative interim PET continued first-line treatment of ABVD for an additional four cycles. The results related to interim PET are described in the section on interim PET. PET/CT was then performed at the end of first line treatment. Patients who had large nodal mass at baseline and whose interim and end of treatment PET were both negative were randomized to receive additional consolidation radiation therapy (n = 148) or no further treatment (no radiation therapy, n= 148).

Similar 3-year probabilities for OS and PFS were seen for the treatment groups indicating that consolidation RT did not confer a substantial clinical benefit (Table 40). OS for those who did and did not receive RT were 100% and 99% (95%-100%) respectively. PFS at three years, probabilities for the RT 97% (92%-99%) and no RT 93% (87%-96%) groups were not statistically different and the difference was small (4/100). In addition the PFS for the 260 patients with no LNM at baseline and not randomly assigned to RT only slightly lower (92%, 95% CI, 88% to 95%) with an estimated difference of 5/100 between those randomized to RT and those who were not randomized. Authors conclude that consolidation therapy in those with baseline large nodal masses can be omitted based on negative results from interim and end of treatment PET (Table 41).

4.3.3.1.2. PET-adapted treatment: results from observational studies

A retrospective observational study (reported across two publications) evaluated the utility of end-of-treatment (EOT) PET to guide the use of radiotherapy (RT) following completion of six to eight cycles of chemotherapy in adults (median age 33 years) with advanced-stage HL. [65,66] The PET-adapted approach was evaluated in a subset of patients only (N=711 of 2126 who were in partial remission and had ≥1 persistent mass measuring ≥2.5 cm and at least 12 months of follow-up) and treatment (RT or no RT) was not randomized based on EOT PET result; therefore, this trial is considered an observational study at high risk of bias based on inability to determine attrition, lack of blinded review of PET images, unclear similarity of groups at baseline and lack control for possible confounding. Demographic data for this specific subset of the overall study was not reported. Patients with EOT PET-negative results (n=529) received de-escalated treatment omitting post-chemotherapy RT, while patients with positive EOT-PET (n=182) patients went on to receive RT. Overall survival was not reported for the PET-adapted RT subset.

The probability of progression-free survival (PFS) at 4-years was significantly greater in those with a negative EOT PET who did not receive RT compared with those who had a positive EOT PET and received RT (95% CI for difference 0.9 to 12.0, p=0.02) (Table 41). When the 4-year PFS of this subset of PET-negative and PET-positive patients in partial remission (with 2.5 cm or larger residual disease) was compared with that of patients in complete remission on CT after completing chemotherapy, the analysis indicated that PET-negative patients had a similar prognosis to those in complete remission (92.6% vs. 92.1%). Authors suggest that radiotherapy can be omitted in patients in persons with PET (–) results who are in partial remission without loss of efficacy. They further conclude that with the EOT PET-adapted treatment approach reduced the number of patient actually receiving additional radiotherapy relative to their previous study.

4.3.3.2. Aggressive NHL

A total of three studies provide limited information on the utility of end-of-treatment PET for guiding management in patients with aNHL: a sub-analysis of the PETAL Trial,[63] which is considered observational for this application of PET, and two additional observational studies.[116,168] Detailed descriptions of the PETAL trial and others are found in the section on interim PET. No RCTs evaluating EOT PET in this population were identified.

4.3.3.2.1. PET-adapted treatment: results from observational studies

In the PETAL trial,[63] EOT PET was restricted to patients with a positive interim scan who did not experience progression on therapy, and were evaluated using International Harmonization Project criteria; 19 patients had negative EOT PET and 31 had positive PET scans showing persistent abnormalities on EOT PET; residual masses (CT) or residual FDG uptake were not considered lymphoma unless there was clinical evidence of disease. Patients with persistent abnormalities on EOT PET who had no clinical evidence of disease were assigned to observation and those and those who had clinical evidence of disease as well received additional immediate treatment.

Data from the PETAL trial suggests that in patients who had a positive interim PET, the prognosis for those with persistent abnormalities on EOT PET (n = 31) is worse compared with those with no such abnormalities (n=19) (Table 41). Both OS and PFS among EOT PET positive patients was lower (65% and 26.4% respectively) compared with those with negative EOT PET (100% for OS, 84.2% for PFS). Among the 31 patients with EOT PET positive results, asymptomatic patients without clinically evident disease who underwent observation had better OS (88.9%) and PFS (66.7%) versus those with clinical evidence of disease who received additional, immediate treatment (OS, 33%; PFS 50.5%) suggesting that EOT PET findings together with consideration of clinical symptoms are important for clinical decision making and that a positive EOT PET may suggest a poor prognosis. Data comparing treatment-related toxicity were not provided.

One retrospective observational study[168] evaluating end-of-treatment (EOT) PET to guide radiation therapy (RT) in adult patients (N=74) with aggressive non-Hodgkin (aNHL) (specifically primary mediastinal large B-cell lymphoma) that met our inclusion criteria was identified. No studies specifically evaluating children with aNHL that met our inclusion criteria were identified. Patients received six cycles of MACOP-B plus rituximab after which those with a positive PET scan received mediastinal RT while those with a negative EOT PET scan were simply observed. Projected 10-year OS and PFS were not reported separately for the two groups; for the entire cohort, OS and PFS were 82% and 88%, respectively (Table 41). Disease-free survival (DFS) at 10-years, estimated from the date of documented complete remission to the last follow-up or to the date of disease recurrence or death as result of lymphoma or acute toxicity of treatment, was 91% for the entire cohort with similar outcomes for those with positive EOT PET (received RT) and negative EOT PET (observation) groups (91% versus 90%, respectively). Only toxicity related to the induction therapy was reported. Similar DSF between those with positive and negative EOT may be interpreted in a number of ways; similarity may suggest that the addition of mediastinal RT improved the prognosis among those with positive interim PET beyond what might have occurred without it. The extent to which EOT PET adequately distinguished patients who may benefit from further therapy from those who may not is not clear from this study.

One additional observational study[116] that evaluated interim PET for guiding therapy escalation for patients with aNHL also reported that EOT PET was a strong predictor of outcome but provided no information treatment-related decisions. Compared with a PET+ result after completion of treatment, a PET – result was associated with a significantly greater probability of both OS (98% vs. 67%) and PFS (81% vs. 57%). Treatment was not modified based on the EOT PET result in this study

Table 41. Advanced-Stage Hodgkin (stage IIB-IV) and Aggressive NHL (stage IIB-IV): PET-adapted therapy evaluating following end-of-treatment PET

	Author (year) Study Design N F/U	Lymphoma subtype Treatment prior to EOT PET PET Criteria	Treatment received according to EOT PET (–) or PET (+) status	Primary Outcomes (95% CI)	p-value
Advanced-Stage HL					
Probability of Overall Survival (1 study[77])	Gallamini 2018 *GITIL/FIL HD 0607 Trial* RCT N=296* Median 3.6 years	Advanced-stage HL 6 cycles of ABVD PET interpreted via the 5-point scale: PET(-), 1-3 PET(+), 4-5	**PET (+), n=148:** Radiotherapy	3-year: 100%	p=NS
			PET (–), n=148: No radiotherapy	3-year: 99% (95%–100%)	
Probability of Progression Free Survival (2 studies, 3 publications[65,66,77])	Gallamini 2018 *GITIL/FIL HD 0607 Trial* RCT N=296* Median 3.6 years	Advanced-stage HL 6 cycles of ABVD PET interpreted via the 5-point scale: PET(-), 1-3 PET(+), 4-5	**PET (+), n=148:** Radiotherapy	3-year: 97% (92%-99%)	p=0.29
			PET (–), n=148: No radiotherapy	3-year: 93% (87%–96%)	
	Engert 2012/2017 *HD15 trial* Prospective cohort (subset analysis of PET-adapted therapy)† N=711† Median 4 years	Advanced-stage HL 6-8 cycles of eBEACOPP‡ PET interpreted via IHP criteria (no further information provided)	**PET (+) in partial remission, n=182:** Radiotherapy	4-year: 86.2% (NR)	p=0.022; 95% CI for difference 0.9–12.0
			PET (–) in partial remission, n=529: No radiotherapy	4-year: 92.6% (NR)	

	Author (year) Study Design N F/U	Lymphoma subtype Treatment prior to EOT PET PET Criteria	Treatment received according to EOT PET (–) or PET (+) status	Primary Outcomes (95% CI)	p-value
			Patients in **complete remission (n=881)** at End-of-Treatment PET	4-year: 92.1% (NR)	NR
aNHL					
Probability of Overall Survival (2 studies)[63,168]	Dührsen 2018 *PETAL Trial* RCT N=50§ Median 44.1 months	B-cell and T-cell lymphomas 6 cycles R-CHOP or 6 cycles of an intensive methotrexate and cytarabine-based regimen (Burkitt's protocol)** IHP criteria	**PET(+), n=31:** Observation (n=18) or immediate therapy (n=13)	**PET(+) observation** 2-year: 88.9% (62.4%-97.1%)	NR
				PET(+) immediate therapy 2-year: 33.3% (10.3%-58.8%)	
				All PET(+) 2-year: 65.9% (95% CI, 45.8%-80.0%)	p=0.0003
			PET(-), n=19: NR	**All PET(-)** 2-year: 100% (NR)	
	Zinzani 2015 Retrospective cohort N=74 Median 62 months	PMLBCL, grade II/IIE (68%) or III/IV (32%) MACOP-B + rituximab/6 cycles PET assessment based on comparison with mediastinal blood pool	Patients with **complete response (n=61)** at follow-up	projected 10 year: 82% (NR)	NR

	Author (year) Study Design N F/U	Lymphoma subtype Treatment prior to EOT PET PET Criteria	Treatment received according to EOT PET (–) or PET (+) status	Primary Outcomes (95% CI)	p-value
Probability of Progression Free Survival (2 studies[63,168])	Dührsen 2018 *PETAL Trial* RCT N=50§ Median 44.1 months	B-cell and T-cell lymphomas 6 cycles R-CHOP or 6 cycles of an intensive methotrexate and cytarabine-based regimen (Burkitt's protocol)** IHP criteria	**PET(+), n=31:** Observation (n=18) or immediate therapy (n=13)	**PET(+) observation** 2-year: 66.7% (40.4%-83.4%)	NR
				PET(+) immediate therapy 2-year: 26.4% (0.7%-52.2%)	
				All PET (+) 2-year: 50.5% (31.8%-66.6%)	p=0.005
			PET(-), n=19: NR	**All PET(-)** 2-year: 84.2% (58.7%-94.6%)	
	Zinzani 2015 Retrospective cohort N=74 Median 62 months	PMLBCL, grade II/IIE (68%) or III/IV (32%) MACOP-B + rituximab/6 cycles PET assessment based on comparison with mediastinal blood pool	**PET(+), n=51:** mediastinal RT	10-year DFS: 90.7% (NR)	p=0.85
			PET(–), n=23: Observation	10-year DFS: 90% (NR)	
			Patients with **complete response (n=61)** at follow-up	projected 10 year DFS: 90.5% (NR) projected 10 year PFS: 87.6% (NR)	NR

ABVD =doxorubicin, bleomycin, vinblastine, and dacarbazine; ASCT= autologous stem cell transplant; CI = confidence interval; CI = confidence interval; eBEACOPP = escalated (i.e., increased doses of) bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone; F/U = follow-up; HDT = high-dose therapy; HR = hazard ratio; IHP = International Harmonization Project; IPS = International Prognostic Score; MINE-R = mesna, ifosfamide, mitoxantrone, etoposide, and rituximab; NA = not applicable; NR = not reported; NS = not statistically significant; OS = overall survival; PDG = cisplatin, gemcitabine, and dexamethasone; PET = positron emission tomography; PFS = progression-free survival; PMLBCL = Primary mediastinal large B-cell lymphoma; VABEM = vindesine, doxorubicin, carmustine, etoposide, and methylprednisolone.

*Patients in complete remission after 6 ABVD cycles with a large (≥5 cm) nodal mass at baseline (n=320) were considered for radiotherapy; 296 patients were actually randomized to treatment (24 not assigned based on patient or medical decision).

†Entry criteria for the PET-adapted RT assessment in this trial were partial response after six to eight cycles of BEACOPP and at least one involved nodal site of 2.5 cm or larger in the transverse or longitudinal diameter as measured by CT. Of 2126 patients from the ITT-analysis, 739 were eligible but 28 did not have documented f/u of at least 12 months and were therefore excluded.

‡Patients were initially randomized to 3 induction treatment groups; of those included in the subset analysis, 226 patients were randomized to 8 cycles of escalated BEACOPP, 249 to 6 cycles of escalated BEACOPP, and 264 patients to 8 cycles of BEACOPP-14.
§End-of-treatment PET scans were restricted to patients with a positive interim scan who did not experience progression on therapy, and were evaluated using International Harmonization Project criteria.
**Patients were previously randomly assigned to receive either six additional cycles of R-CHOP or six cycles of an intensive methotrexate and cytarabine-based regimen (Burkitt's protocol)

4.3.3.2.2. Other observational studies

Two comparative observational studies that met our inclusion criteria were identified that reported survival outcomes following end-of-treatment PET/CT in different populations of non-Hodgkin lymphoma (NHL): one evaluated patients with aggressive NHL (aNHL)[138] and the other indolent NHL (iNHL).[31]

One study (N=41) reported survival outcomes in patients with aggressive T-cell NHLs following end-of-treatment (EOT) assessment with PET/CT compared with CT alone,[138] (Table 41). Median patient age was not reported. EOT assessment was performed after autologous stem cell transplantation (ASCT); prior to ASCT 39% of patients had undergone first-line chemotherapy and 61% second-line chemotherapy. Patients with a negative EOT CT scan confirmed by a negative EOT PET scan showed a roughly 10% increase in the probability of both overall survival (OS) and event-free survival (EFS) compared with patients who had a negative CT scan only; for patents with a positive EOT CT scan confirmed by a positive EOT PET scan, projected survivals decreased by about 5% compared with those patients who had a positive CT scan only (Table 42). In this patient population, a positive post-ASCT PET result was associated with worse prognosis than a negative post-ASCT PET. Furthermore, excluding one ambiguous result, the PET result was predictive in six (60%) of ten patients who showed discordant response between CT and PET images.

One study reported survival outcomes in patients with iNHL following end-of-treatment (EOT) assessment with PET/CT compared with CT alone.[31] The population was comprised of children (median age 10.5 years) with primarily advanced-stage nonlymphoblastic lymphoma (82% B-cell; 18% T-cell). EOT assessment was performed 4 to 6 weeks after the completion of chemotherapy (5.5 to 6 months after beginning treatment). Median length of follow-up was 20.3 months. No significant differences were seen for the probability of overall (OS) and progression-free survival (PFS) according to PET/CT versus CT alone when patients with and without complete remission according to both modalities were compared, Table 42.

Table 42: End-of-treatment PET imaging of aggressive (aNHL) and indolent non-Hodgkin lymphoma (iNHL): Survival outcomes from a clinical observational studies

Study, year Study design	Population	Comparison	2-year OS (± SD)	p-value	6-month EFS (± SD)	p-value
aNHL						
Sohn 2013	Likely adults: median age NR (98% age ≤60 years) Subtype: peripheral T-cell (32%), extranodal NT/T-cell (24%), anaplastic large cell (20%), T-lymphoblastic (12%), angioimmunoblastic (5%) and other (7%) Stage: early (68%); advanced (32%)	**Patients with negative EOT scan**				
		PET-/CT-	53.0% ± 13.2%	NR	72.5% ± 11.8%	NR
		CT- only	44.9% ± 12.0%		62.4% ± 11.3%,	
		Patients with positive EOT scan				
		PET+/CT+	26.8% ± 12.3%	NR	16.3% ± 10.5%	NR
		CT+ only	32.7% ± 10.4%		20.4% ± 9.1%	
Study, year Study design	**Population**	**Comparison**	**OS (95% CI)**	**p-value**	**PFS (95% CI)**	**p-value**
iNHL						
Bakhshi 2012	Pediatric: median age 10.5 years Subtype: non-lymphoblastic (B-cell 82%; T-cell 18%) Stage: early (9%); advanced (91%)	**patients *with* complete remission on EOT:**				
		PET/CT	84.2% (58.6%–94.6%); SE 8.4	>0.99	94.7% (68.1%–99.2%); SE 5.1	>0.99
		CT	84.2% (58.6%–94.6%); SE 8.4		94.7% (68.1%–99.2%); SE 5.2	
		patients *without* complete remission on EOT:				
		PET/CT	55.6% (20.4%–80.4%); SE 16.6	>0.99	66.6% (28.2%–87.8%); SE 15.7	>0.99
		CT	50.0% (20.8%–73.6%); SE 14.4		50% (20.8%– 73.6%); SE 14.4	

CT = computed tomography; EFS = event-free survival; EOT = end-of-treatment; NR = not reported; NT = natural killer; PET = positron emission tomography; PFS = progression-free survival; SD = standard deviation; SE = standard error of the estimate

4.3.4. PET/CT for surveillance and assessment of relapse

Summary of results

- One RCT (Moderately low risk of bias) compared PET/CT with ultrasound and chest radiography for post-remission surveillance in apparently asymptomatic patients following treatment for advanced-stage HL was identified. Four retrospective cohort studies (moderately high to high risk of bias) compared use of PET/CT in patients in asymptomatic patients with its use only in those with clinical suspicion of disease. .
- In the RCT, relapse was detected equally in both PET/CT and US/chest radiography groups (40 in each); clinical findings were found in 36% (29/80) who had relapse. False positive findings were more common in the PET/CT group (13.7%) versus the US/radiography group (3.7%).
- Across three of the cohort studies[59,67,121] in persons with different types of lymphoma, OS and PFS were similar for asymptomatic groups who received follow-up PET/CT versus those who received routine CT or those who received imaging for clinical indication only. (SOE INSUFFICIENT)
- Across three cohort studies, PET/CT may not be superior to clinical follow-up alone (i.e., symptomatic detection) at identifying relapses.[45,67,121]
- In a cohort study of NHL, PET/CT done in those with a clinical suspicion of disease yielded more true positive results that scans done in asymptomatic patients. False positive scans in asymptomatic patients ranged from 2% to 7.5%.[144,145] PET/CT led to changes in clinical management less than 10% of the time in asymptomatic patients compared with >30% in patients receiving scans based on clinical suspicion of disease.
- In general, authors of these studies and those included for the contextual question for diagnostic accuracy conclude that routine surveillance of asymptomatic persons should not be done.

Studies included

One RCT (N = 300)[120] comparing PET/CT with a combination of chest radiography and ultrasound in patients with advanced-stage HL, including those with unfavorable prognostic factors at presentation, who had completely responded to first-line treatment was identified. The total number of exams was 1235 and 1255 respectively for the imaging groups. First-line treatment consisted of six courses of an anthracycline-containing regimen followed by radiation of bulky tumor sites. Authors primarily reported on diagnostic accuracy parameters using histologic examination as a reference standard (see contextual question). None of the patients had focal FDG uptake at post-treatment assessment and 33% in each group had residual masses on CT. The population appears to be asymptomatic at randomization, however, authors do not state whether patients had any clinical suspicion of recurrence at randomization. Patients were followed for a median of 60 months. Authors also provided costing data but did not do a full economic comparison of the imaging modalities. Risk of bias was considered moderately low; methodological concerns included local of independent or blinded assessment of PET/CT.

A total of five retrospective comparative observational studies that met our inclusion criteria were identified that evaluated the utility of imaging with PET/CT for lymphoma in patients who had achieved complete remission following treatment and compared outcomes for surveillance in asymptomatic patients with those who received PET/CT based on clinical suspicion of disease [45,59,67,121,144,145]. Various types of lymphomas were evaluated across the studies: mixed populations of early- and advanced-stage HL (2 studies),[59,121] a mixed population of aNHL and iNHL (primarily DLBCL and follicular) (1 study),[144,145] relapsed/refractory DLBCL (subtype of aNHL) (1 study),[67] and transformed NHL (1 study).[45] Median patient ages ranged from 31 to 59 years across all studies. No studies were identified that evaluated surveillance PET/CT imaging in children.

Survival outcomes

Three cohort studies reported survival outcomes following surveillance with PET/CT (plus clinical follow-up) compared with clinical follow-up alone (imaging done only if indicated by signs or symptoms).[59,67,121] Two studies evaluated patients with HL (mixed early- and advanced-stage)[59,121] and the third evaluated patients with relapsed/refractory DLBCL (subtype of aNHL).[67] Across all three studies, no significant differences between follow-up groups were seen for the probability of overall survival or for progression-free survival in one study,[59] Table 43.

Relapse assessment and surveillance

In the RCT, relapse was detected in 80 patients (40 with PET/CT, 40 with ultrasound/chest radiography) for a projected 5-year recurrence rate of 27% with no difference between groups with regard to time from baseline until recurrence detection (Table 44). Among those with relapse 36% (29/80) had clinical findings of recurrence; relapse was identified by imaging alone in the remaining 64% of patients and authors suggest that information from imaging be combined with clinical evaluations. Sixteen of the relapses were in deep residual masses which were PET negative at time of restaging. False positive findings in those with detected relapse based on imagine were reported in 3.7% (4/40) and 13.7% (15/40) of patients in the ultrasound/chest radiography and PET/CT groups, respectively, using biopsy as a referent. Authors conclude that US/chest radiography for routine surveillance is effective and safe.

Two of the three cohort studies that provided data on survival, one in patients with HL[121] and one in patients with aNHL (DLBCL),[67] also reported the utility of PET/CT compared with clinical follow-up or CT alone for detecting relapse during surveillance. An additional study in adults (median age 59 years) with mostly advanced-stage transformed NHL (all patients had a diagnosis of iNHL with a synchronous or subsequent diagnosis of DLBCL) was also identified that provided this data.[45] Across all studies, results suggest that PET/CT may not be superior to clinical follow-up alone (i.e., symptomatic detection) at identifying relapses; all the studies concluded that PET/CT is of limited utility for surveillance of asymptomatic patients of these lymphomas (Table 44).

An additional retrospective cohort study evaluated 204 patients with biopsy-proven NHL who had one or more follow-up PET/CT scans 6 months following completion of initial therapy,[145] (Table 45). Of the 560 scans performed, 69.7% (n=388) were done without clinical suspicion of recurrence and 30.7% (n=172) were done in those with clinical suspicion of recurrence. In those without clinical suspicion of recurrence, false positive results occurred in 7.5% with true positive results in 12.4% and management

changed after 8.3% of scans. By contrast, in those *with* clinical suspicion of disease, 1.7% of results were false positive with 73.8% being true positives and disease could be ruled out in 16.3%; management changed after 37.8% of scans. In a follow-up subanalysis of 77 patients who had >3 follow-up PET/CT scans following completion of initial therapy,[144] authors reported a higher yield of true positive results among patients who had clinical suspicion of disease compared with those without. In the first 3 post therapy scans in those without clinical suspicion of disease, false positives results occurred in 5.7% with 13.4% true positive scans versus 5.4% false positive and 74.3% true positives scans in those with clinical suspicion of disease in whom tumor was excluded in 18.8%. No management changes were made in 91.7% and 66.2% of patients without and with suspicion of disease respectively based on the scans. For the 4th and subsequent scans in those without clinical suspicion of disease false positive and true positive results occurred in an equal proportion (5.1%) and in those with clinical suspicion of disease, true positive results occurred in 57.6% with and disease was ruled out in 27.3%. No management changes were made based on the scans in 79.4% and 42.4% of patients without and with clinical suspicion of disease respectively.

Information from the contextual question

Tables 9-19 in for the contextual question provide additional information regarding diagnostic accuracy parameters from systematic reviews and included primary studies for use of PET surveillance in patients with lymphoma. Across studies included for the contextual question for diagnostic accuracy authors conclude that routine surveillance of asymptomatic persons should not be done.

Table 43: Surveillance PET imaging of adults with Hodgkin lymphoma (HL) or aggressive non-Hodgkin lymphoma (aNHL): Survival outcomes from clinical observational studies

Study, year Study design	Population	Comparison*	OS (95% CI)	p-value	PFS (95% CI)	p-value
Hodgkin lymphoma						
Dann 2014 Retrospective cohort	Adults: median 31 years Stage: early (36%) and advanced (64%)	Clinical follow-up + dedicated PET/CT (n=305)	5-year overall: 94% (not reported by group)	NS between groups	3 year PFS: 93% (NR) 5 year PFS: 92% (NR)	NS at 3- or 5- year
		Clinical follow-up + CT (only in case of relapse suspicion) (n=63)			3 year PFS: 86% (NR) 5 year PFS: 86% (NR)	
Pingali 2014 Retrospective cohort	Adults: median 35 years Stage: early (47%), advanced (53%)	Routine CT and/or PET (n=174)	5 year: 97% (92% to 99%)	0.41	NR	-----
		Clinical follow-up + imaging only if indicated (n=67)	5 year: 96% (86% to 99%)		NR	
Aggressive non-Hodgkin						
Epperla 2016 Retrospective cohort	Relapsed/ refractory DLBCL Adults: median 55 years Stage: early (24%), advanced (76%)	***Survival from time of transplantation, in patients with relapse†:***				
		Routine CT and/or PET (n=32)	median 1.8 (range 0.4–12.4) years	0.68	NR	-----
		Clinical follow-up (signs/symptoms, or laboratory findings) (n=13)	median 1.6 (range 0.4–8.4) years		NR	
		Survival from time of relapse (post-transplant)‡:				
		Routine CT and/or PET (n=32)	median 1.0 (range 0.02–11.6) years	0.36	NR	-----
		Clinical follow-up (signs/symptoms, or laboratory findings) (n=13)	median 0.34 (range 0.01–2.6) years		NR	

CI = confidence interval; CT = computed tomography; DLBCL = diffuse large B-cell lymphoma; EFS = event-free survival; NR = not reported; PET = positron emission tomography; PFS = progression-free survival.

*In all studies, clinical follow-up entailed medical history, physical examination for signs and symptoms, and laboratory/blood tests.

†Overall survival (OS) was defined as the time from autologous hematopoietic cell transplantation to date of death or last follow-up.

‡Postrelapse survival was defined as time from relapse (after auto-HCT) to date of death or last follow-up.

Table 44: Proportion of Patients with Relapse detected during surveillance imaging by Comparator

Study, year Study design	Population Definition of relapse detection	Method of identification	Relapse FP/FN/TP/TN*	p-value	Author's conclusion
Hodgkin lymphoma					
Picardi 2014 RCT	Adults: median 29 years Stage: advanced (100%) Definitions: NR	All patients	26.7% (80/300)		US/chest radiography are diagnostic tools that enable effective, safe, and low-cost routine surveillance imaging for patients at high risk of Hodgkin lymphoma relapse.
		Routine clinical and PET/CT (n=150)	26.7% (40/150) **FP**: 13.7% (15/40) **FN**: 0% **TP**: 100% (40/40)	P=NR	
		Routine clinical and US/chest radiography (n=150)	26.7% (40/150) **FP**: 3.7% (4/40) **FN**: 2.5% (1/40) **TP**: 97.5% (39/40)		
Pingali 2014 Retrospective cohort	Adults: median 35 years Stage: early (47%), advanced (53%) Definitions: • Surveillance imaging detection = asymptomatic • Clinical detection = in response of symptoms or examination findings	All patients	4.6% (11/241)		Given the lack of apparent benefit from the extra imaging associated with routine surveillance imaging, combined with the significant additional costs and potential risks, clinical surveillance is not inferior to routine surveillance imaging.
		Routine CT and/or PET scans (n=174) (surveillance imaging)	3.4% (6/174)	p=0.39	
		Clinical follow-up only (n=67)	7.5% (5/67)		
Aggressive NHL					
Epperla 2016 Retrospective cohort	Relapsed/refractory DLBCL Adults: median 55 years Stage: early (24%), advanced (76%) Definitions: • Surveillance imaging detection = asymptomatic • Clinical detection = based on signs, symptoms, or laboratory findings	All patients	28% (45/160)		For patients with relapsed and refractory DLBCL who undergo auto-HCT, a majority (71%) of relapses are detected by imaging, unlike the front-line setting, where a majority of relapses are detected clinically. Overall, there appears to be limited utility for routine imaging after auto-HCT except in select cases where earlier detection and salvage with allogeneic HCT is an option.
		PET or PET/CT (surveillance imaging)	15% (24/160) overall; 53% (24/45) of relapses	p=NR	
		CT scan (surveillance imaging)	5% (8/160) overall; 18% (8/45) of relapses		
		Clinical follow-up	8% (13/160) overall;		

Study, year Study design	Population Definition of relapse detection	Method of identification	Relapse FP/FN/TP/TN*	p-value	Author's conclusion
			28% (13/45) of relapses		
Transformed NHL					
Cheah 2014 Retrospective cohort	Transformed iNHL† Adults: median 59 years Stage: early (20%), advanced (80%) <u>Definitions</u>: • Subclinical relapse = detected without any of the below features, on the basis of imaging findings • Suspected relapse = preceded by signs, symptoms or other clinical features such as rising serum LDH	All patients	29.1% (16/55)		Surveillance imaging of transformed NHL achieving CMR is not indicated. PET-CT should be reserved for evaluation of suspected relapse.
		PET (subclinical detection)	12.7% (7/55) **FP‡:** 5% (9/180 scans) **FN‡:** 2% (3/180 scans) **TP:** 8% (15/180 scans) **TN:** 85% (153/180 scans)	p=NR	
		Clinical follow-up (suspected/symptomatic detection)	16.4% (9/55); all DLBCL **FP:** 11% (1/9) **TP:** 89% (8/9)		

Auto-HCT = autologous hematopoietic cell transplantation; CI = confidence interval; CT = computed tomography; DLBCL = diffuse large B-cell lymphoma; iNHL = indolent non-Hodgkin lymphoma; FN = false negative results; FP = false positive results; NR = not reported; PET = positron emission tomography; TN = true negative results; TP = true positive results;

*Pingali 2014 and Epperla 2016 do not report these rates in their populations; this outcome is only reported by Cheah 2014.

†iNHL with a synchronous or subsequent diagnosis of DLBCL.

‡Five indeterminate scans which were not followed by relapse were considered false positives; two indeterminate scans which were followed by relapse were considered false negatives.

§For the group who had PET/CT without prior clinical suspicion, 18 of the 302 negative scans (6%) and 9 of the 86 positive scans (10%) were without confirmation result. For the group who had PET/CT with prior clinical suspicion, 1 of the 30 negative scans (3%) and 12 of the 142 positive scans (8%) were without confirmation result.

Table 45: Summary of accuracy and management changes during surveillance imaging in a mixed population of aNHL and iNHL.

Study year Study design	Population	Method of identification	TP % scans	FP % scans	TN % scans	FN % scans	Change in management (% of scans)
Taghipour 2016a Retrospective cohort (database)	**N**=204 patients (388 scans) **Subtype**: DLBCL 42%; follicular 23%; mantel cell 9.3%; other types 26% **Adults**: mean 56 years Stage: early 5.4%; advanced 30%; unknown 64%	PET/CT **without** prior clinical suspicion (n=388 scans)	12.4% (48/388)	7.5% (29/388)	71.1% (276/388)	2.1% (8/388)	**Yes:** 7.5% (29/388 scans) **No:** 85.6% (332/388 scans) **Unknown:** 6.2% (24/388 scans)
		PET/CT **with** prior clinical suspicion (n=172 scans)	73.8% (127/172)	1.7% (3/172)	16.3% (28/172)	0.6% (1/172)	**Yes:** 34.3% (59/172 scans) **No:** 50.6% (87/172 scans) **Unknown:** 11.6% (20/172 scans)
Taghipour 2016b* Retrospective cohort (database)	**N**=77 patients* (231 first 3 follow-up scans; 208 fourth and subsequent scans) **Subtype**: DLBCL 48%; follicular 12%; mantel cell 7.8%; marginal zone5 5.2% other types 27% **Adults**: mean 56 years Stage: early 12%; advanced 35%; unknown 53%	**In patients with first 3 follow-up scans:**					
		PET/CT **without** prior clinical suspicion (n=157 scans)	13.4% (21/157)	5.7% (9/157)	80.3% (126/157)	0.6% (1/157)	**Yes:** 7.0% (11/157 scans) **No:** 91.7% (144/157 scans) **Unknown:** 1.3% (2/157 scans)
		PET/CT **with** prior clinical suspicion (n=74 scans)	74.3% (55/74)	5.4% (4/74)	18.9% (14/74)	0% (0/74)	**Yes:** 32.4% (24/74 scans) **No:** 66.2% (49/74 scans) **Unknown:** 1.4% (1/74 scans)
		In patients with 4th and subsequent follow-up scans:					
		PET/CT **without** prior clinical suspicion (n=175 scans)	5.1% (9/175)	5.1% (9/175)	72.6% (127/175)	3.4% (6/175)	**Yes:** 9.2% (16/175 scans) **No:** 79.4% (139/175 scans) **Unknown:** 11.4% (20/175 scans)
		PET/CT **with** prior clinical suspicion (n=33 scans)	57.6% (19/33)	3.0% (1/33)	27.3% (9/33)	3.0% (1/33)	**Yes:** 36.4% (12/33 scans) **No:** 42.4% (14/33 scans) **Unknown:** 21.2% (7/33 scans)

Auto-HCT = autologous hematopoietic cell transplantation; CI = confidence interval; CT = computed tomography; DLBCL = diffuse large B-cell lymphoma; iNHL = indolent non-Hodgkin lymphoma; FN = false negative results; FP = false positive results; NR = not reported; PET = positron emission tomography; TN = true negative results; TP = true positive results;
*Subset of Taghipour 2016a; includes 77 patients who had more than 3 follow-up PET/CT scans after the completion of initial therapy. These 77 patients had 231 first three follow-up scans and 208 fourth and subsequent follow-up PET/CT scan

4.4. Key Question 2: Harms and Complications

4.4.1. Number of studies retained

All included comparative studies identified were evaluated for harms and complications. In addition, studies designed specifically to address safety concerns related to the use of PET/CT imaging in patients with lymphoma were sought. Overall nine studies, to include one RCT,[120] seven comparative observational studies,[59,83,90,116,118,121,130] and one case series[56] were included. Of the studies that provided data for efficacy or effectiveness, only three comparative observational studies reported safety outcomes[59,116,121]

Safety outcomes related to PET/CT imaging that were the focus of this report included radiation exposure, additional testing (e.g. to follow-up on false positive results), incidental findings and reaction or allergy to the contrast agent.

The overall strength of evidence was considered low for all safety outcomes; most studies were moderately high or high risk of bias. Details of specific adverse events reported by each study is available in below and in Appendices F–I. Section 5 of the report provides details of strength of evidence determination for each outcome assessed.

Summary of results:

- Mean effective radiation doses from PET/CT varied across studies based on number of scans performed, length of follow-up and other factors such as patient's age. Reported doses for adults ranged from 14.5 mSv to 26.1 mSv for surveillance/relapse assessment (3 studies) and 26.1 mSv during treatment (1 study); for children, doses were estimated to be 6.4 mSv and 8.6 mSv for a 10- and 15-year-old child, respectively from one study.
- Across two studies, high false positive rates with PET/CT reported during surveillance led to additional perhaps unnecessary diagnostic testing: 50% (vs. 15% for US/chest radiography) (1 RCT) and 14% (vs. 6% for CT alone) (1 observational).
- Three studies reported relapse identification during surveillance/relapse assessment. One RCT found that US/radiography was as effective at identifying relapse as PET/CT (96% vs. 100%, respectively) with significantly lower radiation exposure (10.2 mSv vs. 449.5 mSv). Two observational studies found that fewer imaging tests were needed to identify a relapse when done as clinically indicated (5 and 15 scans) versus as a routine practice with PET/CT (48 and 127 scans).

Findings from the previous report

The previous report found limited evidence on safety and noted that although there is a moderate radiation dose associated with each PET and PET/CT scan performed, lymphoma is a potentially lethal disease. Concern for the effects of radiation may be more important for younger patients and for repeated PET and CT studies during follow-up. They rated the overall strength of evidence as low.

Findings from this updated report

4.4.2. Radiation exposure

Oncology patients generally receive multiple imaging tests that involve ionizing radiation for diagnosis, stage, evaluation of treatment response and post-treatment surveillance. In addition some cancer treatment regimens involve the use of radiation. While debate remains regarding the accuracy of the linear no-threshold model's ability to correctly quantify low doses of radiation exposure and its effects, radiation is cumulative and may increase life-time risk of radiation-induced malignancies.[81] The impact of radiation exposure on older adults with shorter lifespans in whom NHL is most common may be different than for HL patients who are generally younger patients who will be most likely be cured and have a long life span. Children are potentially at higher life-time risk for radiation-related malignancies; they are younger when exposed and have more years for malignancies to manifest. Potential benefits and risks, including any related to not performing the test, should be carefully considered before ordering tests that will expose patients to ionizing radiation. The results of shared decision making between patients and their providers around the long-term relative risks and benefits from the radiation associated with PET/CT will differ based on the patient's circumstance, preferences and values. Data on radiation doses may inform such decisions.

Radiation dose from PET/CT is significant. Data from the Radiological Society of North America (RSNA) and American College of Radiology (ACR) Radiologyinfo.org website (updated 2018) indicates that the approximate effective radiation dose for an average size adult from combined PET/CT is 25 mSv (comparable to approximately 250 chest x-ray equivalents or 8 years natural background radiation).[127] Most of the radiation come from the CT scan. Dose from CT varies depending on whether the CT is a low-dose CT performed to anatomical correlation only or a standard CT. Depending on location of procedure, dosage from CT can range from 2 mSv (comparable to 20 chest x-ray equivalents or 8 months of natural background radiation) to 20 mSv (comparable to 200 chest x-ray equivalents or 7 years of natural background radiation) for adults, Table 46 below. Effective radiation doses differ for children (Table 47). The cumulative amount of radiation a patient receives depends on the number and type of scans (PET and others) performed. Best practices suggest that minimizing radiation and attending to the ALARA – "as low as reasonably achievable" – tenet whenever possible is prudent.

Table 46: Radiation Dose from Common X-ray and CT Examinations

Procedure	Approximate effective radiation dose for an average-sized adult:	Comparable to natural background radiation for:
Nuclear Medicine		
Positron Emission Tomography-Computed Tomography (PET/CT)	25 mSv	8 years
Abdominal Region		
Computed Tomography (CT)–Abdomen and Pelvis	10 mSv	3 years
Computed Tomography (CT)–Abdomen and Pelvis, repeated with and without contrast material	20 mSv	7 years
Computed Tomography (CT)–Colonography	6 mSv	2 years
Intravenous Pyelogram (IVP)	3 mSv	1 year
Barium Enema (Lower GI X-ray)	8 mSv	3 years
Upper GI Study with Barium	6 mSv	2 years
Bone		
Spine X-ray	1.5 mSv	6 months
Extremity (hand, foot, etc.) X-ray	0.001 mSv	3 hours
Central Nervous System		
Computed Tomography (CT)–Head	2 mSv	8 months
Computed Tomography (CT)–Head, repeated with and without contrast	4 mSv	16 months
Computed Tomography (CT)–Spine	6 mSv	2 years
Chest		
Computed Tomography (CT)-Chest	7 mSv	2 years
Computed Tomography (CT)-Lung Cancer Screening	1.5 mSv	6 months
Chest X-ray	0.1 mSv	10 days
Dental		
Dental X-ray	12 mSv	1 day
Heart		
Coronary Computed Tomography Angiography (CTA)	12 mSv	4 years
Cardiac CT for Calcium Scoring	3 mSv	1 year
Men's Imaging		
Bone Densitometry (DEXA)	0.001 msv	3 hours
Women's Imaging		
Bone Densitometry (DEXA)	0.001 msv	3 hours
Mammography	0.4 mSv	7 weeks

Procedure	Approximate effective radiation dose for an average-sized adult:	Comparable to natural background radiation for:
Note for pediatric patients: Pediatric patients vary in size. Doses given to pediatric patients will vary significantly from those given to adults.		
The information in this table has been abstracted from a Patient Safety publication for clinicians and patients from the Radiologic Society of North America and the American College of Radiology.[127] The following table is excerpted from this publication which can be accessed in full at: https://www.radiologyinfo.org/en/info.cfm?pg=safety-xray		

Table 47: Effective Radiation Dose from 18F-FDG PET scan in Pediatric Patients

		Age (Nominal Weight, kg)			
		1 year (10)	5 years (19)	10 years (32)	15 years (55)
2007 European Association of Nuclear Medicine dosage card	Effective Radiation Dose (mSv)	6.7	6.7	7.0	7.2
	Comparable to natural background radiation for:	2.25 years	2.25 years	2.33 years	2.4 years
2010 North American consensus guidelines	Effective Radiation Dose (mSv)	4.8	5.5	6.2	6.9
	Comparable to natural background radiation for:	1.5 years	1.8 years	2.1 years	2.33 years
The information in this table regarding the effective radiation dose from PET for pediatric patients has been abstracted from an article published succeeding a meeting regarding the harmonization of the European Association of Nuclear Medicine dosage card and the North American consensus guidelines.[102]					

Results

A total of five studies provided data related to radiation exposure in patients undergoing diagnostic PET/CT for management of lymphoma[56,83,118,120,130]; only one study included pediatric patients.[130]

Adults

Across the four studies that provided data for adults (Table 48), radiation exposure from PET/CT was evaluated during the surveillance period in three studies[83,118,120]; and also during the treatment period in one of these studies.[83] The timing of the evaluation was unclear in the fourth study (the case series).[56] Two studies evaluated patients with HL (early-stage[118] and late-stage[120]) with median ages of 29 and 33 while the other two studies evaluated mixed populations of primarily NHL (71%-84%) and HL who were somewhat older than the HL populations (median age 54 years in one study; 76% aged 40-79 in the other).[56,83] Follow-up periods varied across the studies.

Effective radiation dose

One RCT (N=300)[120] and two comparative observational studies (N=78 to 108)[83,118] reported the mean effective radiation dose from PET/CT (dose per scan) which ranged from 14.5 mSv to 26.6 mSv; for the

surveillance period only estimates ranged from 14.5 mSv to 26.1 mSv (3 studies) and for the treatment period, 26.6 mSv (1 study), Table 48. The estimated radiation dose of 25 mSv for a PET/CT scan as specified by the RSNA and ACR lies at the top (and slightly above) this range (Table 46). One study reported mean effective radiation dose for PET alone (14.1 mSv) which was 12 mSv less than the dose for the combined PET/CT.[118] For comparison, estimated mean effective radiation dose for other imaging modalities evaluated in these studies include: CT of the chest, abdomen and pelvis (range for surveillance, 22.0-44.1 mSv across two studies[83,118]; for treatment period 26.8 mSv in one study[83]); CT of the head and neck (range for surveillance 4.3-6.0 mSv across two studies[83,118]; for treatment period 4.7 mSv in one study[83]); and ultrasound (US)/chest radiography (0.06 mSv in one study[120]). Of note, this latter study was a RCT which found that PET/CT and US/chest radiography were similarly effective at detecting relapse (100% vs. 97.5%, respectively) in patients with advanced-stage HL who were in complete remission following primary treatment; however, the estimated radiation dose needed to detect one relapse was significantly greater with PET/CT (449.5 mSv vs. 10.2 mSv, p=0.0002). Authors do not describe the methodology for this determination; Each imaging group had 150 patients and presumably the estimate is across all exams (n = 1235 for PET/CT and 1255 for US/chest radiography).

Cumulative effective radiation dose

One case series (N=486)[56] reported an estimated cumulative effective dose from all ionizing radiation imaging received over a 1 year follow-up period (range 1-3 years) of a median 69 mSv per patient (interquartile range [IQR] 42-118); of the total cumulative dose, PET/CT contributed 8% (~5.5 mSv) and CT with contrast contributed 89% (~61.4 mSv) (Table 47). Per year of follow-up, the median cumulative dose was 48 mSv. The median number of scans per patients was nine, with a median of 0 (IQR 0-1) for PET and 4 (IQR 2-6) for CT. The authors also report that the following variables were associated with a greater likelihood of exposure to ≥75 mSv of ionizing radiation: younger age (p=0.03), longer the duration of follow-up (p=0.04), more advanced stage (p=0.04).

Pediatric

One study in 99 pediatric patients estimated cumulative radiation exposure from various diagnostic imaging modalities used for surveillance over a 2-year period following completion of primary treatment; however, cumulative estimates for PET or PET/CT specifically were not provided.[130] Overall estimated cumulative radiation doses by disease stage can be found in Table 48. The authors do state that the radiation effective dose of FDG-PET/CT scans from FDG alone (not considering the additional radiation from the concomitant CT scan) are approximately 6.4 and 8.6 mSv for a 10- and 15-year-old child, respectively. In this pediatric population the median number of surveillance scans received for 2 years following the termination of therapy was 11 (range 1–26), the majority of which were CT scans. The effective doses of various CT scans as reported by the study are available in Table 48. Authors note that surveillance with such tests, to include PET and PET/CT, may result in substantial cumulative radiation exposure and risks of second malignancy from ionizing radiation but does not help to identify relapse.

Table 48. Summary of reported effective radiation dose across included studies

Author (year) Population RoB	Timing; Frequency of imaging	Measure of Radiation and Results
Adults		
Picardi 2014 N = 300 patients PET/CT exams n = 1235 US/Chest x-ray exams n = 1255 Median age 29 years HL (stage: 100% advanced) RCT *Moderately Low*	Surveillance (median f/u 5 years) Years 1-2: every 4 months Year 3: every 6 months Years 4-5: every 12 months Years 6-9: bi-yearly	**Estimated mean effective radiation dose per scan (SD)** **PET/CT**: 14.5 mSv (NR) **Chest x-ray**: 0.1 mSv (NR) Estimated dose from additional imaging: **PET/CT**: NR **Ultrasound/Chest x-ray**: 7 mSv (diagnostic CT of thorax done) **Estimated dose needed to detect one relapse* (determination method NR)**: PET/CT: 449.5 mSv, US/Chest radiography: 10.2 mSv; p=0.0002
Patel 2013 N = 78 (2418 scans) Median age 33 HL (stage: 100% early) Retrospective cohort *Moderately High*	Surveillance (median f/u 4.6 years) <u>Type of Imaging Modalities (2418 scans)</u> Chest CT: 33% Abdominopelvic CT: 26% X-ray: 20% PET and PET/CT: 8% Neck CT: 7% Single Photon Emission Tomography: 3% MRI: 2% Head CT: 1%	**Estimated mean effective radiation dose (SD)** **PET:** 14.1 mSv (NR) **PET/CT:** 26.1 mSv (NR) **CT** of the: • Chest: 8 mSv (NR) • Abdominopelvic region: 14 mSv (NR) • Neck: 4 mSv (NR) • Head: 2 mSv (NR) **X-ray:** 0.06 mSv (NR)
Guttikonda 2014 N = 108 Median age 54 years Mixed NHL (71%) and HL (29%) Retrospective cohort *Moderately High*	During treatment period (n=84) and/or Surveillance (n=85)† <u>Treatment</u> **PET/CT (n=39 patients)** • 73 scans total • mean 1.4 scans/patient **CT (n=45 patients)** • 161 scans • mean 0.24 head CTs/patient	**Estimated mean effective radiation dose (SD)** <u>Treatment</u> **PET/CT** (whole body)**:** 26.6 mSv (10.2) (range 9.1-45.2) **CT of the:** • Chest, abdomen, and pelvis: 26.8 mSv (18.1) (range 2.9-78.2) • Head and neck: 4.7 mSv (7.2) (range 0.8-24.2) <u>Surveillance</u>

Author (year) Population RoB	Timing; Frequency of imaging	Measure of Radiation and Results
Adults		
	• mean 0.2 neck CTs/patient • mean 1.8 chest CTs/patient • mean 1.4 abdominal/pelvic CTs/patient Treatment period: 8 months <u>Surveillance</u> **PET/CT (n=25 patients)** • 39 scans • mean 0.58 scan/patient) **CT (n=60 patients)** • 378 scans total • mean 0.28 head CTs/patient • mean 0.57 neck CTs/patient • mean 2.8 chest CTs/patient • mean 2.4 abdominal/pelvic CTs/patient Surveillance period: 23 months Frequency of Imaging: NR	**PET/CT** (whole body)**:** 18.0 mSv (8.1) (range 7.9-36.3) **CT of the** • Chest, abdomen, and pelvis: 44.1 mSv (34.5) (range 2.6-136.8) • Head and neck: 4.3 mSv (2.7) (range 0.7-10.7)
Crowley 2016 N = 486 Median age NR (76% 40-79 years) Mixed NHL (84%) and HL (16%) Case series *High*	Timing NR (median follow-up: 1 year, range 1-3) <u>Median number of scans per patient</u>: All ionizing radiation imaging: 9 (IQR, 5-15) PET/CT: median 0 (IQR, 0-1) CT: median 4 (IQR 2-6)	Estimated median ***cumulative*** **effective dose per patient** (from all ionizing radiation imaging): 69 mSv (IQR 42–118) • CT contributed 89% (61.4 mSv) of the radiation dose and PET/CT contributed 8% (5.5 mSv) • Age <40 years: 89 mSv (IQR 55-124) (75% with curable disease) • Stem cell transplantation: 162 mSv (IQR 135-225) (59% had RT) **Median cumulative effective dose per year of follow-up**: 48 mSv (IQR, 24-70) **Cumulative effective dose** >150 mSv: 14% of patients The following variables were associated with a greater likelihood of exposure to ≥75 mSv of ionizing radiation:

Author (year) Population RoB	Timing; Frequency of imaging	Measure of Radiation and Results
Adults		
		• younger age (p=0.027), • longer the duration of follow-up (p=0.042) • more advanced stage (p=0.042)
Pediatric Patients		
Rathore 2012 N = 99 (1455 scans†) Median age 13 years HL (stage: early 67%; advanced 23%) Retrospective cohort *Moderately High*	Surveillance (all ionizing radiation imaging)‡ Median 11 (0-26) scans in first 2 years after therapy completion; mean of 15 scans per patient 3.9% (58/1455) scans were PET or PET/CT scans; authors do not report radiation dose specific to these.	Stage at diagnosis (all ionizing radiation imaging) Stage I: 31.97 mSv‡ Stage II: 37.76 mSv‡ Stage III: 48.08 mSv‡ Stage IV: 51.35 mSv‡ Effective radiation dose of FDG PET alone: • 6.4 mSv for a 10-year-old • 8.6 mSv for a 15-year-old Effective radiation dose for a 10-year-old for a CT of the: • Neck: 3 mSv • Chest: 3 mSv • Abdomen/pelvis: 5 mSv • Chest/abdomen/pelvis: 8 mSv • Neck/chest/abdomen/ pelvis: 11 mSv Only 1.3% (17/1358) scans† detected relapse; 13% (13/99) patients relapsed; PET/CT detected relapse in 2.6%

CT = computed tomography; HL = Hodgkin lymphoma; IQR = interquartile range; MRI = magnetic resonance imaging; NHL = non-Hodgkin lymphoma; NR = not reported; mSv = millisievert; PET/CT = positron emission tomography; RCT = randomized controlled trial; RoB = risk of bias; RT = radiation therapy; SD = standard deviation; US = ultrasound.

*Authors do not specify the methodology for determining this; presumably it is across all scans for the 150 patients in each imaging group

†Patients could have had scans performed during either treatment or surveillance periods, or during both periods.

‡Imaging involving ionizing radiation only (e.g. CT, radiographs, PET, PET/CT); surveillance scanning among high-risk patients typically consisted of CT scans every 3 to 6 months for the first 12 to 24 months following treatment completion. CT scans are required every 3 months for the first 18 months for patients treated according to the intermediate-risk protocol and every 4 to 6 months for the first 12 to 24 months according to the low-risk protocols

4.4.3. Additional testing

False positive PET scans may result in additional biopsies or in mistaken upstaging of HL and NHL patients which could impact treatment decisions.

A total of four comparative observational studies (N=106 to 368) provided data related to additional procedures or tests acquired during the post-treatment surveillance period in adult patients undergoing diagnostic PET/CT for management of lymphoma.[59,90,120,121] No studies in pediatric populations were identified. Three studies[59,120,121] included patients with HL (primarily advanced-stage) with median ages ranging from 29 to 35 years and one study included patients with primarily early-stage (37%) DLBCL (subtype of aNHL) with a median age of 54 years.[90] Follow-up periods varied in length across the studies (Table 49).

The two studies,[59,121] in addition to usual clinical follow-up, compared surveillance by means of routine PET/CT imaging versus imaging (usually contrast enhanced CT) only when indicated by signs or symptoms and reported that patients in the latter group had significantly more additional testing over the course of follow-up (Table 49). In one study, the use of routine imaging required 3.9 scans per patient compared to 0.6 scans per patient in the symptom-driven imaging group.[59] In a second study, patients in the routine surveillance imaging group underwent 4.5 times more imaging over the course of a year than those in the clinical surveillance group (average 0.98 versus 0.21 scans per year, respectively).[121] Additionally, in both studies, the number of imaging tests needed to diagnose or detect a single relapse was significantly higher in the routine imaging versus clinical surveillance groups: 47.5 versus 4.7 in one study[59] and 127 versus 14.6 in the other.[121]

A third study[90] in patients with DLBCL reported that, as a result of false positives, imaging with PET/CT resulted in additional procedures (e.g., biopsy, ENT or gynecological exam) or imaging tests (e.g., CT or MRI) in 15 (of 21) patients (14% overall) compared with six of seven patients (6% overall) imaged with CT scan (6 and 1 patient with false positives on imaging, respectively, underwent observation), Table 49. Additionally, relapse was detected by routine imaging in only one out of eight patients who experienced relapse within 24 months after completing the treatment.

The fourth study was a RCT comparing PET/CT and US/chest radiography for relapse detection in patients with advanced-stage HL who were in complete remission following primary treatment.[120] The authors report that due to a higher number of false positive results in the deep compartments of the mediastinum obtained with PET/CT compared with chest radiography, a significantly greater proportion of patients required mediastinotomies in the PET/CT group (50% vs. 15%, p=0.04), Table 49.

4.4.4. Over-treatment

One prospective observational study in 66 adult patients with DLBCL suggested that 57% (17/30) of patients who were positive on interim PET (after 3 cycles of R-Mega CHOP) may have been over treated as a result of using visual evaluation to determine a positive PET result.[116] Upon retrospective re-review of the scans using a semi-quantitative evaluation method, it was determined that those 17 PET-positive patients may have actually had a negative PET scan. Based on negative interim PET results, these patients would have been considered to be in CR and therefore would have only received an additional 3 courses of R-MegaCHOP. However, they received two courses of R-IFE (rituximab, ifosfamide,

etoposide) followed by BEAM (BCNU, etoposide, cytarabine, melphalan) and autologous stem cell transplantation which may have been unnecessary treatment.

Table 49. Summary of additional tests performed across included studies

Author (year) Population Lymphoma Type RoB	Timing of PET/CT	Type of additional test or proportion over-treated Frequency % (n/N)
Dann 2014 N=368 Median age 31 years HL (stage: early 36% and advanced 64%) Retrospective cohort *High*	Surveillance (median f/u 5 years) Routine imaging every 6 mos. for 2 years and once in year 3	<u>No. of imaging tests required per patients</u> Clinical follow-up and: • dedicated PET/CT (n=305): 3.9 • CT/imaging only if symptomatic: 0.6 P<0.001 <u>No. of imaging tests needed to diagnose/detect a single relapse:</u> Clinical follow-up and: • dedicated PET/CT (n=305): 47.5 • CT/imaging only if symptomatic: 4.7
Hong 2014 N=106 (501 scans) Median age 54 years DLBCL (stage: 63% early, 37% advanced) Retrospective cohort *Moderately High*	Surveillance (median f/u 3 years) Number of planned visits with routine imaging: median 4, range 0–15, mean 4.8 Patients could have both a CT and PET/CT scan	<u>False positives</u> PET/CT: 21 patients [20%] CT scan: 7 patients [7%] <u>Additional tests as a result of false positives</u> **PET/CT (23 scans among 21 patients;** 20% [21/106 patients] required further treatment**)** Biopsy (various): 34.8% (6/21) ENT exam: 4.3% (1/21) Pap with gynecological exam: 8.7% (2/21) Correlation with CT: 2.2% (5/21) Correlation with MRI: 4.3% (1/21) Observation: 26.1% (6/21) **CT (7 scans among 7 patients)** Biopsy (various): 28.6% (2/7) Correlation with MRI: 28.6% (2/7) Short-term CT follow-up: 28.6% (2/7) Observation: 14.3% (1/7)
Picardi 2014 N = 300 Median age 29 years HL (stage: 100% advanced) RCT	Surveillance (median f/u 5 years) Years 1-2: every 4 months Year 3: every 6 months Years 4-5: every 12 months Years 6-9: bi-yearly	<u>Proportion of patients requiring surgical biopsies (mediastinotomy)</u>: • PET/CT: 50% (20/40) • US/Chest radiography: 15% (6/40) p=0.04 For PET/CT, 11/20 were FP; for US/chest radiography, 1/6 was FN

Author (year) Population Lymphoma Type RoB	Timing of PET/CT	Type of additional test or proportion over-treated Frequency % (n/N)
Moderately Low		
Pingali 2014 N=241 Median age 35 years HL (stage: early 49%, advanced 51%) Retrospective cohort *Moderately High*	Surveillance (median f/u 4.2 years)	Average number of scans per year: • Routine PET/CT imaging: 0.98 (95% CI 0.89-1.08) • Clinical follow-up (imaging only if indicated, CT or PET): 0.21 (95% CI 0.18-0.31) • Relative risk: 4.5 (95% CI 3.1-5.5); p<0.0001 Number of scans performed per relapse detected: • Routine PET/CT imaging: 127 • Clinical follow-up (imaging only if indicated): 14.6

CT = computed tomography; DLBCL = diffuse large B-cell lymphoma; ENT = ear, nose and throat; FN = false negative; FP = false positive; f/u = follow-up; HL = Hodgkin lymphoma; MRI = magnetic resonance imaging; PET/CT = positron emission tomography; RCT = randomized controlled trial; RoB = risk of bias; US = ultrasound

4.4.5. Incidental findings

When performing a diagnostic imaging test there is a chance that the scan could identify additional previously undiagnosed conditions, which would likely require further medical attention and potentially psychological distress. These are called incidental findings. No study meeting our inclusion criteria was identified that reported on incidental findings.

4.4.6. Other adverse events or harms (e.g. contrast reactions, reactions to radiopharmaceutical or PET procedure)

As stated in the prior report, the radiopharmaceutical ^{18}FDG used for PET scanning is an analog of glucose. Intuitively, ^{18}FDG should be well tolerated as a glucose analog. The authors reported that no contrast reactions had been noted to date for ^{18}FDG. Similarly, no studies that met the inclusion criteria for this re-review that reported on reactions to ^{18}FDG or to other contrast or radiopharmaceutical mediums or to the PET procedure were identified.

4.5. Key Question 3: Differential Efficacy and Harms in Subpopulations

4.5.1. Number of studies retained

For this key question, RCTs that stratified on baseline patient characteristics and evaluated effect modification were sought. Subgroups of interest included (but were not limited to): age, sex, race, ethnicity, socioeconomic status, payer, and worker's compensation. All RCTs included to evaluate the efficacy or safety of PET/CT were assessed. No trials meeting the inclusion criteria were identified.

4.6. Key Question 4: Cost effectiveness

4.6.1. Number of studies retained

One cost-effectiveness analysis (CEA)[43] and one cost-utility analysis (CUA)[93] were identified that met the inclusion criteria assessing the cost effectiveness of PET for initial staging and for routine surveillance post-treatment, respectively, in adults with lymphoma. The CEA included individuals with Hodgkin Lymphoma (HL) aged 13 years and older, however mean age suggests the population was predominately adult. The CUA focused on an adult population age 55 years with diffuse large B-cell lymphoma (DLBCL). Both studies adopted payer perspectives, although the CUA claimed a societal perspective. The CEA for initial staging did not provide a description of the model used while the CUA implemented a Markov Model to forecast long-term costs of health outcomes for surveillance of asymptomatic patients at time of first remission. The main outcome measured in both of the studies was the incremental cost effectiveness ratio (ICER). The ICER for the CEA reflected the cost per modified treatment case divided by the percent of patients with appropriate treatment using PET/CT compared to conventional staging

methods. The CUA reported the ICER in terms of the marginal cost of adding one quality adjusted life year (QALY) to a patient's life when using PET/CT compared to routine follow-up.

Each of the studies retained was scored using the Quality of Health Economic Studies (QHES) instrument, a standardized metric to assess the soundness of economic evaluations. The CEA score was 17/100 points. Primary concerns were lack of model specification, absence of sensitivity analyses, unclear sources of economic parameters, a poorly defined effectiveness outcome, and limited applicability to the American healthcare environment due to the international nature of the analysis. The CUA score was 78/100 points and the study was reasonably well conducted. The primary concern was that it was unclear where the RCT-derived model parameters were reported in the RCT. In addition, there was a lack of transparency regarding other components of the model (e.g. set up of transitions and source of health state utility parameters). See Appendix E for details of the quality assessments.

No full economic studies evaluating the cost-effectiveness of interim PET/CT (including restaging, treatment response or post-treatment evaluation) or for other types of lymphoma were identified.

Summary of Economic Studies

- One poor-quality (QHES 17/100 points) CEA and one moderate quality (QHES 78/100 points) CUA evaluated the cost effectiveness of PET for evaluation of lymphoma in adult patients. One study did not report the source of funding and the other study was funded by the Brazilian Health Ministry. No full economic studies related to use of PET/CT for evaluation of lymphoma in patients under 13 years of age or of use of interim PET/CT were identified.
- Pre-treatment staging, HL: one poor-quality CEA was conducted in Brazil and took a payer perspective; however it only included some direct costs. It concluded that both combined PET/CT and PET with CT separately are cost effective compared with conventional staging.
    - ICER for PET/CT: $16,215 per modified treatment case. ICER for PET with separate CT: $35,490 per modified treatment case. Authors concluded that both strategies are cost-effective in Brazil, using a threshold of 3 times GDP per capita.
    - Limitations:
        - No model specification, unclear methodology, short time horizon, no sensitivity analyses
        - Physician fees not included, source of cost data was vague
        - Denominator of ICER derived using author-created "reference standard" which was poorly defined and not well validated
        - Because this is not a standardized way of reporting cost-effectiveness, results are not comparable across studies or treatments
        - Results not generalizable to US
- Surveillance at first remission (asymptomatic patients) with DBLCL: one moderate quality CUA was conducted in the US and claims to have taken a societal perspective; however it did not include indirect costs. It concluded that neither PET/CT nor CT alone is cost effective compared with routine surveillance without imaging.
    - ICER for PET/CT: $168,750 per QALY. ICER for CT alone: $164,960 per QALY. Authors concluded that neither strategy is cost effective using a threshold of $150,000.

- Probabilistic sensitivity analysis suggests that the PET/CT is not cost-effective. The likelihood of the intervention being cost-effective is 9% at the willingness-to-pay threshold of $150,000 per QALY over a lifetime time horizon based modeling of a hypothetical cohort.
- Limitations:
  - Used a lifetime horizon but RCT data is limited to 6 years after autologous SCT
  - Not generalizable to patients at high risk of relapse
  - Source of utility weights not specified

4.6.2. PET/CT for Pre-Treatment Staging of Hodgkin Lymphoma

Study characteristics and framework

One poor quality CEA evaluated the cost effectiveness of conventional staging with PET/CT or PET with separate CT compared to conventional staging alone for HL in patients 13 years of age and older.[43] Conventional staging included biopsy and immunohistochemistry, bilateral bone marrow biopsy, laboratory tests, echocardiogram, and CT. The study was conducted in Brazil. The costing year was 2009. Although both Italian and Brazilian cohorts were included in the study, only the Brazilian cohort was used for the cost-effectiveness analysis. The median age was 34.6 ± 16.2 years, suggesting that the population was primarily adults, and a slightly higher proportion of patients were male compared to female. Only short-term costs were included and the time horizon was not specified.

The study assumed PET/CT and PET with separate CT in addition to conventional staging would result in a higher percentage of patients receiving appropriate treatment compared to conventional staging alone. This was determined by comparing results of each staging method to determine if modification in clinical stage would lead to a change in treatment.

The study adopted a payer perspective but only incorporated a portion of direct costs; physician fees were not included. Costs included physical exams, biopsy and immunohistochemistry panels, laboratory tests, bilateral bone marrow biopsy (BMB), CT scans, PET scans, PET/CT scans, and echocardiograms. The costs for PET/CT, PET, and CT were reported to be $2,037, $1,330, and $1,200 per scan, respectively.

Study characteristics, results, and conclusions are summarized in Table 50.

Results

The direct costs (excluding physician fees) associated with adding scans to conventional staging were $37,411 for PET/CT and $38,397 for PET and separate CT. The direct cost (excluding physician fees) associated with conventional staging alone was $36,584.

Based on clinical follow-up and actual treatment regimens, PET/CT or PET with separate CT correctly defined treatment in 94.2% of patients. In contrast, conventional staging alone correctly defined treatment in 89.1% of patients. However, the authors do not describe specific treatment strategies or if appropriate treatment was received, but provide only information on whether disease status was upstaged or downstaged. The corresponding cost per correctly defined treatment case was $39,714 for PET/CT, $40,761 for PET with separate CT, and $41,059 for conventional staging alone. Importantly,

there was no verification of whether patients received appropriate treatment or even if treatment was altered on the basis of imaging results.

The ICER summarized the differences in costs and effectiveness by expressing the incremental cost of using imaging in addition to conventional staging versus conventional staging alone for each modified treatment case. However, given that it was not stated whether treatment was modified, this definition of the ICER is not accurate. The denominator used for the ICER is the difference in the proportion of correctly defined treatment regimens comparing PET/CT or PET with separate CT to conventional staging. Thus, the ICER for PET/CT was $16,215 per percent of patients with a correctly defined treatment regimen and the ICER for PET with separate CT was $35,490 per percent of patients with a correctly defined treatment regimen.

Sensitivity Analyses
This study did not conduct any sensitivity analyses.

Conclusions and Limitations
With a threshold of 3 x GDP per capita (Brazil's GDP per capita was $19,016 in 2010), the CEA found both PET/CT and PET with separate CT to be cost effective interventions compared to conventional staging alone.

The major limitations of this study are the lack of model specification, questionable methodological validity, and the lack of any sensitivity analysis. There was no indication of how decisions about treatment modification were made nor how treatment would have been modified. As a result, the meaning of the reported ICER differs from the interpretation set forth by the authors. It is also unclear to what extent comparison of staging and post-treatment scans was impacted by actual disease progression. The authors also did not address how they dealt with subjectivity in evaluating the scans. The sourcing of model parameters was also concerning; the calculations and methodology used to derive cost data were not transparent. In addition, only the Brazilian cohort was evaluated and limited data were provided on the Italian cohort. Overall, the methods and results were poorly reported. Another consideration is this study has limited relatability to healthcare in the US because pricing and costs of care in the US differ from those in Brazil as does the system of health care provision. The QHES score was 17/100 points.

4.6.3. Surveillance at first remission of asymptomatic patients with diffuse large B-cell lymphoma

Study characteristics and framework
One moderate-quality CUA evaluated the cost effectiveness of adding routine PET/CT or CT scans to clinical follow-up compared to routine surveillance (without imaging) for patients with DLBCL.[93] The cohort included 55 year-old patients in first complete remission (not defined). It was conducted in the US and the costing year was 2013. A six health state Markov model was used to model outcomes for patients' remaining lifetimes.

The study reported the cost-effectiveness in terms of quality-adjusted life years. Most model inputs were derived from the randomized-controlled CORAL study (Collaborative Trial in Relapsed Aggressive Lymphoma),[87] which compared R-ICE (rituximab, ifosfamide, carboplatin, and etoposide) to R-DHAP (rituximab, dexamethasone, ara-C, and cisplatin) for treatment of DLBCL. The baseline model assumed improved clinical outcomes (lower secondary age-adjusted International Prognostic Index [sAA-IPI] and improved survival after autologous stem cell transplant [SCT]) in patients with imaging-detected

asymptomatic relapse. The rationale behind this assumption was based on a retrospective study that found aggressive non-HL patients with asymptomatic relapse detected through imaging were more likely to have lower sAA-IPI scores[104] and long-term studies that found improved survival after autologous SCT among patients with lower sAA-IPI scores.[79,80]

The authors stated they adopted a societal perspective, but no indirect costs were included. Costs were broken down into 11 categories: routine follow-up visit, CT imaging, PET/CT imaging, work-up for positive CT scan, work-up for positive PET/CT, symptomatic relapse evaluation, salvage immunochemotherapy, autologous SCT, palliative immunotherapy, lymphoma-related death, and non-lymphoma death. The cost for PET/CT and CT were reported to be $1,209 and $928 per scan, respectively.

Both one-way sensitivity and probabilistic sensitivity analyses were conducted.

Study characteristics, results, and conclusions are summarized in Table 50.

Base Case Results
Adding PET/CT to routine follow-up was found to cost an additional $4,270 (time frame undefined) and resulted in an increase of 0.025 QALYs. Adding CT to routine follow-up was found to cost an additional $3,310 (time frame undefined) and resulted in an increase of 0.020 QALYs. The ICER for PET/CT was $168,750 per QALY while the ICER for CT was $164,960 per QALY.

Sensitivity Analyses
Probabilistic sensitivity analysis (PSA) suggests PET/CT is not cost-effective. The likelihood of the intervention being cost-effective is 9% at the willingness-to-pay threshold of $150,000 per QALY and <1% at a threshold of $100,000. ICERs ranged from $117,440/QALY to $853,550/QALY during Monte Carlo simulations. The most influential parameters in the one-way sensitivity analyses were the probability of favorable sAA-IPI in both imaging-detected relapse and symptomatic relapse. An exploratory multiway sensitivity analysis found that increasing risk of relapse from 20% during 5 years to 35% during the first 2 years of follow-up resulted in an ICER of less than $50,000/QALY. Thus in subpopulations with higher rates of relapse, PET/CT could potentially be cost-effective.

Conclusions and Limitations
The authors concluded that surveillance imaging for asymptomatic DLBCL is unlikely to be cost effective for patients in first remission. They speculate this may be a reflection of lymphoma biology; relapse of less aggressive lymphoma is more likely to be identified using routine surveillance. Authors also suggest that because these lymphomas are less severe, the clinical utility of early detection is minimal. However, the exploratory sensitivity analysis suggests that PET/CT might be cost-effective in DLBCL patients at high risk for relapse, such as those with double-hit lymphoma.

The primary limitation of this study is that the model was not well-defined. It was unclear how the model transitions were set up, the source of the utility weights was not transparent, and it was not evident where the parameters they used were reported in the RCT. For some parameters (false positive results) data were from retrospective studies and thus may not be reflective of current technology or treatment outcomes. In addition, while the model assumed a lifetime horizon, the RCT data only provided information up to 6 years post autologous SCT. Finally, the results of this study are not generalizable to patients at high risk of relapse because the relapse rate was assumed to be 20% in this model. These concerns led to a QHES score of 78/100 points.

Table 50. Overview of formal economic studies for Initial Staging and Surveillance Imaging

	Cerci 2011[43] (Pre-treatment staging)	Huntington 2015[93] (Surveillance)
Population	Brazilian cohort only (full economic evaluation only included Brazilian patients) Newly diagnosed, biopsy proven HL 47.5% female Median age 34.6 (±16.2) Clinical stages: 2.5% stage I, 33.3% stage II, 25.8% stage III, 38.3% stage IV	Cohort mirrored population of CORAL study Diffuse large B-cell lymphoma (DLBCL) 55 years of age, asymptomatic, in first complete remission
Intervention(s)	Pre-treatment staging PET/CT PET + separate CT	Surveillance at first remission PET/CT CT
Comparator(s)	Conventional staging (biopsy and immunohistochemistry, bilateral BMB, lab results, EKG, and CT)	Routine follow-up
Country	Brazil (and Italy)	United States
Funding	Brazilian Health Ministry	NR
Study design	CEA	CUA
Perspective	Payer	Payer (societal stated)
Time horizon	Not stated	Lifetime
Analytic model	Not specified	Markov model with 6 health transition states
Effectiveness outcome	$ per modified treatment case	QALYs
Effectiveness outcome components	% of patients who would have had modified treatment if PET/CT or PET+CT was used in place of conventional staging	Assumed improved clinical outcomes in patients with imaging-detected asymptomatic relapse Key health states/complications: Relapse after achieving complete remission, incidence of detecting asymptomatic relapse, likelihood of sAA-IPI <2 at time of relapse, incidence of false-positive imaging, likelihood of receiving salvage at relapse, likelihood of autologous SCT post-salvage, outcomes following autologous SCT (early TRM and durable second remission)
Source for effectiveness data	Author's own study	Prior literature
Costing year	2009	2013
Currency	1 USD = 1 Real	USD
Discounting	Not specified	Outcomes and costs, 3%

	Cerci 2011[43] (Pre-treatment staging)	Huntington 2015[93] (Surveillance)
Components of cost data	Physical exams, biopsy and immunohistochemistry panel, laboratory tests, bilateral BMB, CT, PET, PET/CT, EKG Physician fees not included	Routine follow-up clinic visit, CT imaging, PET/CT imaging, work-up for positive CT scan, work-up for positive PET/CT scan, symptomatic relapse evaluation, salvage immunochemotherapy, autologous SCT, palliative immunotherapy, lymphoma-related death, non-lymphoma death
Cost sources	Microcost building methods to estimate procedure and test unit costs (not specified)	2013 Medicare fee schedule Medical literature
Sensitivity analysis	None	One-way: all model parameters over relevant clinical ranges and costs PSA: probabilities and utilities represented by β distributions and costs represented by γ distributions, Monte-Carlo simulation of PSA using 100,000 iterations for each strategy Exploratory multiway sensitivity analysis
QHES	17/100	78/100
Results:		
Cost / effectiveness outcome of intervention	PET/CT: $37,411/94.2% = $39,714 PET + CT: $38,397/94.2% = $40,761	PET/CT: $41,550/14.865 = $2,795 CT: $41,590/14.860 = $2,799
Cost / effectiveness outcome of comparator(s)	$36,584/89.1% = $41,059	$38,280/14.840 = $2,580
ICER	PET/CT: $16,215 per modified treatment case PET + CT: $35,490 per modified treatment case	PET/CT: $168,750 per QALY CT: $164,960 per QALY
One-way sensitivity analysis	NR	Likelihood of low sAA-IPI if imaging detected relapse – decrease to 50% caused ICER to increase to $705,000/QALY for CT and $788,000/QALY for PET/CT Most changes in baseline parameters resulted in increased ICERs – all remained >$110,000/QALY Other influential parameters: likelihood of low sAA-IPI if symptomatic relapse and PET/CT cost

	Cerci 2011[43] (Pre-treatment staging)	Huntington 2015[93] (Surveillance)
Other sensitivity analysis	NR	Results from Monte Carlo simulations PET/CT: $168,750 ($117,440 - $853,550) CT: $164,960 ($116,510 - $766,930) 9% of simulations were under $150,000/QALY while <1% were under $100,000/QALY No simulations were under $50,000/QALY Exploratory multiway sensitivity analysis: increasing relapse risk to 35% during first 2 years up follow-up resulted in ICERs below $50,000/QALY
Author's Conclusion	Using a cost-effectiveness threshold of 3 x GDP per capita, both PET/CT and PET + CT are cost effective for Brazil	Clinical utility of imaging-based surveillance was small Surveillance imaging for asymptomatic DLBCL patients in first remission is unlikely to be cost effective
Limitations	• Source of cost data is vague; determination of benefit is unclear • Use of reference standard instead of gold standard • Calculations/model not clearly specified • Physician fees not included • No sensitivity analyses • No discounting • Short time horizon	• Model design not transparent; parameters not readily evident in RCT; transitions not clearly defined • Parameters based on retrospective cohort studies may not reflect current technology • Lifetime horizon but RCT data is limited to 6 years post autologous SCT • No citations for utility weights • Not generalizable to patients at high risk of relapse

BMB = bone marrow biopsy; CCS = conventional clinical staging; CEA = cost-effectiveness analysis; CT = computed tomography; CUA = cost-utility analysis; DLBCL = diffuse large B-cell lymphoma; EKG: echocardiogram; GDP = gross domestic product; ICER = incremental cost-effectiveness ratio; PET = positron emission tomography; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life years RCT = randomized controlled trial; sAA-IPI = secondary age-adjusted International Prognostic Index; SCT = stem cell transplant; TRM = transplant-related mortality; USD = United States dollar.

5. Strength of Evidence (SoE) Summary Tables

The following summaries of evidence have been based on the highest quality of studies available. Additional information on lower quality studies is available in the report. A summary of the primary outcomes for each key question are provided in the tables below and are sorted by timing of PET and type of lymphoma. Details of other outcomes are available in the report.

5.1. Strength of Evidence Summary: Efficacy Results

5.1.1. Strength of evidence summary: Initial staging PET in adults (Survival outcomes)

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect Estimate (95% CI)	Quality (SOE)
Hodgkin Lymphoma (early-stage)								
Probability of Relapse- Free Survival	4 years	1 prospective cohort (N=37) *Figura 2017*	Yes[1] (-1)	Unknown[2]	No	Yes[1] (-1)	**Pre-treatment PET vs. no PET** 97% (NR) vs. 67% (NR) p = 0.001 **Conclusions:** 4-year RFS was greater in patients who underwent a PET prior to starting chemotherapy compared with those who did not have a PET.	⊕◯◯◯ INSUFFICIENT
Indolent NHL (follicular lymphoma)								
Probability of Overall and Progression Free Survival	5 years	1 retrospective cohort (N=206) *Friedberg 2012*	Yes[1] (-1)	Unknown[2]	No	No	**Initial staging using PET vs. conventional methods** OS: data NR, p=NS PFS: aHR 0.87 (95% CI 0.47 to 1.62) **Conclusion**: Authors report no statistical difference in OS or PFS at 5 years when initial staging was done with PET compared with conventional methods.	⊕◯◯◯ INSUFFICIENT

CI = Confidence Interval; aHR = adjusted Hazard Ratio; NR = Not Reported; NS = Not significant; OS = Overall Survival; PET = Positron Emission Tomography; PFS = Progression Free Survival; RFS = Relapse Free Survival.

Reasons for downgrading:

1. Serious risk of bias: the majority of studies violated one or more of the criteria for good quality RCT (or cohort study) related to the outcome reported (see Appendix for details).
2. Inconsistency: differing estimates of effects across trials; If point estimates across trials are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies may be downgraded. Consistency is also unknown across PET-adapted studies given the differences between studies in initial as well as randomized treatment protocols based on PET results.

3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.

4. Indirect, intermediate or surrogate outcomes may be downgraded.

5.1.2. Strength of evidence summary: Interim PET in adults with early-stage Hodgkin Lymphoma (Randomized PET-Adapted studies)

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect Estimate (95% CI)	Quality (SOE)
Escalation of therapy (Randomization of PET + to escalated versus standard treatment)								
Probability of Overall Survival	5 years	1 Trial (N = 361 PET + persons) Andre 2017, Raemaekers 2014 *H10 Trial*	No	Unknown[2]	No	No	**Escalated treatment vs. standard*, PET +** 96.0% (91.1%–98.2%) vs. 89.3% (83.4%–93.2%); Estimated difference 6.7/100 HR 0.45 (0.19–1.07), p=0.062 **Conclusions:** OS was slightly higher in those receiving escalated treatment compared with standard therapy; results may be of clinical significance but did not reach statistical significance.	⊕⊕⊕◯ MODERATE
Probability of Progression Free Survival			No	Unknown[2]	No	No	**Escalated treatment vs. standard*, PET +** 90.6% (84.7%–94.3%) vs. 77.4% (70.4%–82.9%); Estimated difference 12.9/100 HR 0.42 (0.23–0.74), p=0.002 **Conclusion**: PFS was better in patients receiving escalated therapy compare with those who continued with standard therapy.	⊕⊕⊕◯ MODERATE
Treatment Toxicity (GRADE 3 or 4)§			No	Unknown[2]	No	No (for all except death)	**Escalated treatment vs. standard*, PET +** **Hematologic grade 3/4** *Neutropenia:* 53.5% vs. 30.3%; RR 1.8 (1.4–2.3); RD 23% (13% to 33%); *Febrile neutropenia:* 23.9% vs. 1.1%; RR 22.7 (5.6–92.6); RD 23% (16% to 29%) *Thrombocytopenia:* 19.7% vs. 0%; RR NC; RD 20%; p<0.0001 *Anemia:* 4.9% vs. 0%; RR NC; RD 5%; p=0.002 **Infections, grade 3/4 (without neutropenia)** 5.6% vs. 1.1%; RR 5.1 (1.1–23.3); RD 4% (1% to 8%); p=0.02	⊕⊕⊕◯ MODERATE

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect Estimate (95% CI)	Quality (SOE)
							Toxicity-related death: 1st line protocol: 0.6% vs. 0.5% ; RD 0.1% (-1% to 2%); NS 2nd line treatment: 0% vs. 1.1% ; RD -1%; NS **Conclusion:** Patients receiving escalated treatment had significantly higher risk of grade 3 or4 hematologic toxicities and infections compared with those receiving standard treatment. There may have not been sufficient statistical power to compare toxicity related deaths in the two groups.	
De-escalation of therapy (Randomization of PET (-) to de-escalated treatments vs. continuation)								
Probability of Overall Survival	3 years	1 (N= 420) RAPID trial (Radford, 2015)	No	Unknown[2]	No	No	**No further treatment vs. INRT†, PET *negative-*** 99.0% (97.6%–100%) vs. 97.1% (, 94.8%–99.4%); Estimated difference 1.9/100 HR 0.51 (0.15%–1.68%), p=0.27 **Conclusion:** In persons whose interim PET was negative, the difference in OS between groups was small (1.9/100) and did not reach statistical significance. The clinical significance of this difference is unclear.	⊕⊕⊕◯ MODERATE
	5 years	1 (N = 515, PET - individuals) Andre 2017, Raemaekers 2014 *H10 Trial*	No	Unknown[2]	No	No	**De-escalated (INRT omission) vs. standard†, PET *negative-*** Initial status Favorable ‡ 99.6% (97.0%–99.9%) vs. 100% Estimated difference -0.4/100 Unfavorable‡ 98.3% (96.0%–99.3%) vs. 96.7% (93.7%– 98.3%); Estimated difference 1.6/100 Across favorable and unfavorable (estimated): 99% vs. 98%, difference 1/100 **Conclusion:** In persons whose interim PET was negative, for 5-year OS there were no statistical differences between those in whom INRT was omitted	⊕⊕⊕◯ MODERATE

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect Estimate (95% CI)	Quality (SOE)
							vs. those who continued with standard treatment. The clinical significance of the small differences is unclear.	
Probability of Progression Free Survival	3 years	1 (N= 420) RAPID trial (Radford, 2015)	No	Unknown[2]	No	No	**No further treatment vs. INRT†, PET *negative*** 90.8% (86.9%–94.8%) vs. 94.6% (91.5%–97.7%); HR 1.57 (0.84%–2.97%) p=0.16; Absolute RD 3.8% (-1.3 to 8.8) **Conclusion:** In persons whose interim PET was negative, 3.8/100 fewer patients who received no further treatment had PFS vs. those who had had the standard treatment with INRT. The difference may be clinically important.	⊕⊕⊕◯ MODERATE
	5years	1 (N = 515, PET - individuals) Andre 2017, Raemaekers 2014 *H10 Trial*	No	Unknown[2]	No	Yes[3] (-1) (for favorable)	**De-escalated (INRT omission) vs. standard†, PET *negative*** Favorable‡ 87.1% (82.1%–90.8%) vs. 99.0% (95.9% –99.7%); HR, 15.8; 95% CI, 3.8 to 66.1) in favor of ABVD + INRT; Estimated difference : -11.9/100 Unfavorable‡ 89.6% (85.5% –92.6%) vs. 92.1% (88.0% –94.8%); HR 1.45 (0.8 to 2.5); estimated difference: - 2.5/100 Across favorable and unfavorable (estimated): 88% vs. 95%; difference -7/100 **Conclusion:** In persons with initially favorable prognosis whose interim PET was negative, PFS at 5 years was better in those who received standard therapy (ABVD + INRT) compared with the those receiving the de-escalated regimen of ABVD alone, (difference 11.9/100) however, wide confidence interval suggests a lack of precision for the estimate. In the unfavorable group, the difference between treatment strategies was smaller (2.5/100). Across prognostic groups, PFS was better with standard versus de-escalated therapy (difference -7/100). The differences may be clinically important.	⊕⊕⊕◯ MODERATE

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect Estimate (95% CI)	Quality (SOE)
Treatment Toxicity (Grade 3 or 4)§		2 trials (N = 515, PET - individuals) Andre 2017, Raemaekers 2014 *H10 Trial* (N= 420) RAPID trial (Radford, 2015)	No	Unknown[2]	No	Yes[3] (-2)	**RAPID trial: No further treatment vs. INRT†; PET negative** *Death due to secondary malignancy:* 1.4% vs. 0.5%; RR 3.0 (0.3–28.3); RD 1% (-1% to 3%) **H10 trial: INRT omission vs. standard†, PET negative** Initial status Favorable‡ *Frequency of second malignancy:* 2.9% vs. 1.3% RR 2.2 (0.6–8.5); RD 2% (-1% to 4%), Unfavorable‡ *Frequency of second malignancy:* 3.0% vs. 3.4%; RR 0.9 (0.4–2.1); RD -0.4% (-3% to 2%); **Conclusion:** Evidence on the impact of PET-adapted de-escalation strategies on treatment-related toxicities is limited and power to detect rare events and differences between strategies was likely limited. Overall, there were no statistical differences between de-escalated strategies and standard treatments. Follow-up time may have been insufficient to capture some RT-related toxicities.	⊕⊕◯◯ LOW

CI = Confidence Interval; HR = Hazard Ratio; INRT = Involved Node Radiation Therapy; NC = Not calculable; NR = Not Reported; NS = Not significant; OS = Overall Survival; PET = Positron Emission Tomography; PFS = Progression Free Survival; R-ACVBP = rituximab, doxorubicin, cyclophosphamide, vindesine, bleomycin, and prednisone; R-CHOP = rituxan, cyclophosphamide, doxorubicin, vincristine, and prednisone; RD = Risk Difference; RR = Risk Ratio

*Escalated treatments: H10 trial: escalated treatment, 2 eBEACOPP cycles + involved-node radiotherapy (INRT) vs. standard arm consisted of ABVD followed by INRT)
†De-escalation of treatment: RAPID trial: Intervention was no further treatment, comparator was involved-field radiotherapy; H10 trial: Treatment strategies for "favorable" were omission of INRT (i.e. 2 additional cycles of ABVD alone) vs. standard treatment (1 additional ABVD cycle followed by INRT; for "unfavorable" were 4 additional ABVD cycles vs. 2 additional ABVD cycles followed by INRT
‡ Refers to prognostic status at initial staging: Favorable status indicates age <50 years with ≤3 involved nodal areas, absence of mediastinal bulk (mediastinum-to-thorax ratio <0.35), and erythrocyte sedimentation rate (ESR) <50 mm without B symptoms or ESR <30 mm with B symptoms; Unfavorable status indicates age ≥50 years, >4 involved nodal areas, presence of mediastinal bulk (mediastinum-to- thorax ratio ≥0.35), or ESR ≥50 mm without B symptoms or ESR ≥30 mm with B symptoms.
§Grading of Adverse Events (AE) is according to the Common Terminology Criteria for Adverse Events (CTCAE) v3.0. Grades of 1-5 refer to the severity of the AE. Grade 1 Mild AE, Grade 2 Moderate AE, Grade 3 Severe AE, Grade 4 Life-threatening or disabling AE, Grade 5 Death related to AE.
Reasons for downgrading:

1. Serious risk of bias: the majority of studies violated one or more of the criteria for good quality RCT (or cohort study) related to the outcome reported (see Appendix for details). None of the included randomized controlled trials for this report randomized patients to receive PET/CT versus a strategy that did not include use of PET/CT, but rather, patients were randomized to various treatment options based on results of their PET/CT. There was no downgrade of these studies based on lack of comparison to a no PET/CT strategy. Risk of bias was assessed based on criteria in Appendix D and any down grade for risk of bias was based on these criteria.

2. Inconsistency: differing estimates of effects across trials; If point estimates across trials are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies may be downgraded. Consistency is also unknown across PET-adapted studies given the differences between studies in initial as well as randomized treatment protocols based on PET results.
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

5.1.3. Strength of evidence summary: Interim PET in adults with advanced-stage Hodgkin Lymphoma (Randomized PET-Adapted studies)

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect Estimate (95% CI)	Quality (SOE)
Escalation of therapy (Randomization of PET + to different escalated treatments)*								
Probability of Overall Survival	3 years	2 Trials (N = 2232 Total) GITIL/FIL HD 0607 Trial (Gallamini 2018, N = 788) *GHSG HD18 trial* (Borchmann 2017b< N = 1444)†	No	Unknown[2]	No	No	**Escalated treatments*, PET + (n= 582)** **HD0607 trial:** 89% (79%–95%) vs. 90% (78%–95%); Estimated difference: -1/100 **HD18 trial:** 95.6 (92.8%–98.4%) vs. 97.1% 97.1% (94.8%–99.4%); Absolute RD –1.4% (–5.1% to 2.2%) **PET– (n= 1640)** **HD0607 trial** (n=630): 99% **HD18 trial** (n=1010): 97.3% **PET+ (overall) vs. PET–** **HD0607 trial:** 89% vs. 99%; p<0.001 **HD18 trial:** 95.6% and 97.1% vs. 97.3%; p=NR **Conclusions:** The probability of OS was similar for escalated treatments in each trial for those with a positive interim PET with a small difference between treatments (1/100 to 1.4/100). Prognosis: The HD0607 trial reported lower OS in PET+ vs. PET– groups, less difference between PET + and PET– was seen in HD18 suggesting somewhat similar prognosis for PET + and PET– in that trial. Somewhat higher OS was seen in the HD18 trial. Differences in treatments and patient populations may contribute to this.	⊕⊕⊕◯ MODERATE
	5 years	1 (N = 1444) *GHSG HD18 trial*	No	Unknown[2]	No		**Escalated treatments*, PET + (n=434)** HD18 trial: 93.9% vs. 96.4%; HR 1.62 (0.70–3.75, eBEACOPP vs. escalation with R-eBEACOPP);	⊕⊕⊕◯ MODERATE

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect Estimate (95% CI)	Quality (SOE)
		(Borchmann 2017b)†					Absolute RD –2.5% (–6.8% to 1.7%) **PET– (n=1010)** HD18 trial: 96.3% **PET + vs. PET–** HD18 trial: 95.5% vs. 96.3%; p=0.49 Estimated difference: -0.8/100 **Conclusion**: OS was not significantly lower for the R-BEACOPP group versus the BEACOPP group. OS was similar across escalated treatment groups in those who were PET + and when compared with those with negative interim PET.	
Probability of Progression Free Survival	3 years	2 (N=2232) GITIL/FIL HD 0607 Trial No (Gallamini 2018) *HD18 trial* (Borchmann 2017b)†	No	Unknown[2]	No		**Escalated treatments*, PET + (n = 582)** **HD0607 trial:** 63% (50%–74%) vs. 57% (45%–68%); p 0.534; estimated difference 6/100 **HD18 trial:** 93.2% (89.8%–96.7%) vs. 92.7% (89.1%–96.2%); Absolute RD 0.6% (–4.4% to 5.5%) **PET – (n= 1640)** HD0607 trial: 87%; HD18 trial: 93.5% **PET + vs. PET -** **HD0607 trial:** 60% vs. 87%; p<0.001 **HD18 trial:** 93.2% and 92.7%; vs. 93.5%; p=NR **Conclusion:** PFS was similar across treatment group for PET + group in the HD 18 trial. Although PFS was higher in for escalated treatment with R-BEACOPP in the HD0607 trial vs. BEACOPP, the difference may not be beyond what may be expected by chance. While the HD0607 trial reported lower PFS in PET + vs. PET – groups, no difference between PET + and PET – was seen in HD18.	⊕⊕⊕◯ MODERATE
	5 years	1 (N = 1444) *GHSG HD18 trial* (Borchmann 2017b)†	No	Unknown[2]	No	No	**Escalated treatments*, PET + (n= 434)** HD18 trial: 88.1% (83.5%–92.7%) vs. 89.7% 89.7% (85.4%–94.0%); HR 1.25 (0.69–2.26, eBEACOPP vs. escalation with R-eBEACOPP); Absolute RD –1.6% (95% CI –7.9% to 4.7%) **PET – (n=1010)** HD18 trial: 91.4%	⊕⊕⊕◯ MODERATE

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect Estimate (95% CI)	Quality (SOE)
							PET + vs. PET - HD18 trial: 88.3 vs. 91.4%; p=0.30 **Conclusion:** PFS was similar across treatment groups in those who were PET +.	
Treatment Toxicity (GRADE 3 or 4)††	NR	2 (N=2232, total) FIL HD 0607 Trial No (Gallamini 2018) *HD18 trial* (Borchmann 2017a and b)†	No	Unknown[2]	No	No	**Escalated treatments*, PET + (n= 434)** **HD18 trial:** R-eBEACOPP vs. eBEACOPP • *Any Mortality:* 6.5% (14/217) vs. 4.1% (9/217); RD 2% (-2% to 7%) • *Acute toxicity:* 97% (213/220) vs. 98% (213/218); RD -1% (-4% to 2%) • *Second malignancy (any):* 3.7% (8/217) vs. 4.6% (10/217); 5-year cumulative incidence estimate‡: 3.5% (95% 0.7%-6.3%) vs. 4.0% (95% CI 1.3%-6.7%) **HD 0607 trial**: No adapted-therapy specific data; • *Death (PET+ vs. PET-):* 11% (16/150) vs. 2% (12/630) • *Toxicity (PET+ vs. PET-) (Grade 3 or4)* ▪ Hematologic toxicity: 76% vs. 30% ▪ Infections: 10% vs. NR ▪ Pulmonary toxicity: NR vs. 2% **Conclusion:** The HD18 trial reported similar proportions of patients experiencing grade 3 and 4 treatment-related toxicities for the escalated treatment groups. The HD 0607 trial reports that fewer toxic events occurred in persons with PET negative versus PET positive only.	⊕⊕⊕◯ MODERATE
De-escalation of therapy (Randomization of PET - to de-escalated treatments vs. continuation)‡								
Probability of Overall Survival	3 years	2 (N=2551, total) *HD18 trial* (Borchmann 2017b, N = 1444)†	No	Unknown[2]	No	No	**De-escalated treatments‡, PET *negative* (n = 1940)** **HD18 trial:** 98.8% (97.8%–99.9%) vs. 95.7 (93.7%–97.6%); Absolute RD: 3.1% (0.9% to –5.3%) **RATHL trial:** 97.6% (95.6%– 98.7%) vs. 97.2% (95.1%– 98.4%); HR 0.09 (0.47–1.74); Estimated difference: 0.4/100 **PET + (n = 1120)**	⊕⊕⊕◯ MODERATE

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect Estimate (95% CI)	Quality (SOE)
		RATHL Trial (Johnson 2016, N = 1107)					HD18 trial: 97.1% (95.9%–98.2%) RATHL trial: 87.8% (81.5%–92.1%) **PET + vs. PET -** HD18 trial: 97.1% vs. [98.8% and 95.7]; p=NR RATHL trial: 87.8% vs. [84.45 and 85.7%]; p=NR **Conclusion:** In persons whose interim PET was negative, the absolute difference in OS between the de-escalated group versus those who received additional treatment cycles in the HD18 trial was 3.1%, which may be clinically important; OS was similar between those receiving de-escalated treatment and those who continued with standard therapy in the RATHL trial. De-escalation of treatment did not adversely impact OS.	
	5 years	(N =1444, total) *GHSG HD18 trial* (Borchmann 2017b)†	No	Unknown[2]	No	No	**De-escalated treatments‡, PET *negative* (n=1005)** HD18 trial: 97.7% (4 cycles) vs. 95.4% (6-8 cycles); HR 0.32 (0.14 to 0.72, 6-8 cycles vs. 4 cycles); p= 0.0037; Absolute RD: 2.3% (–0.2% to 4.9%); **PET + (n=948)** HD18 trial: 95.5% (93.9% to 97.1%) **PET + vs. PET -** HD18 trial: 95.5% vs. 96.3%p =0.49 **Conclusion:** OS was similar at 5 years for those whose treatment was de-escalated and those who continued on more intense treatment. OS between those with PET + and PET – at the interim time point were also similar.	⊕⊕⊕◯ MODERATE
Probability of Progression Free Survival	3 years	2 (N=2551) *HD18 trial* (Borchmann 2017b)† RATHL Trial (Johnson 2016)	No	Unknown[2]	No	No	**De-escalated treatments‡, PET *negative* (n= 1940)** **HD18 trial**: 95.3% (93.3% to 97.3%) vs. 91.7% (89.0% to 94.4%); Absolute RD: 3.6% (–0.2% to 6.9%) **RATHL trial:** 84.4% (80.97.5%) vs. 85.7% (82.1%–88.6%), HR 1.13 (0.81, 1.5, ABVD vs. AVD); Absolute RD 1.6% (–3.2 to 5.3); **PET + (n= 1120)**	⊕⊕⊕◯ MODERATE

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect Estimate (95% CI)	Quality (SOE)
							HD18 trial: 92.5% (90.7%–94.3%) RATHL trial: 67.5% (59.7%–74.2%) **PET + vs. PET -** HD18 trial: 92.5% vs. [95.3% and 91.7%] RATHL trial: 67.5% vs. [84.4% and 85.7%] **Conclusion:** In persons whose interim PET was negative, PFS was statistically similar in groups receiving de-escalated treatment versus those with more intensive therapy in both trials. In the HD 18 trial, PFS was slightly greater in those receiving de-escalated treatment. This Thisindicates that PFS was not compromised with de-escalation of treatment.	
	5 years	1 (N=1444) *GHSG HD18 trial* (Borchmann† 2017b)	No	Unknown[2]	No	No	**De-escalated treatments‡, PET *negative* (n=1005)** HD18 trial: 92.2% (4 cycles) vs. 90.8% (6-8 cycles) HR 0.79 (0.50 to 1.24, 6-8 cycles vs. 4 cycles); Absolute RD: –1.4% (-2.7% to 5.4%) (6-8 cycles vs. 4 cycles) **PET + (n=948)** HD18 trial: 88.3% (85.8%–90.8%) **PET + vs. PET -** HD18 trial: 88.3% vs. [92.2% and 90.8%]; p=0.30 **Conclusion:** In persons whose interim PET was negative, the difference in PFS between groups was small (1.4/100), not statistically significant and of unclear clinical significance. No difference in PFS was seen between those whose interim PET was positive and those whose results were negative.	⊕⊕⊕◯ MODERATE
Treatment Toxicity (GRADE 3 or 4) ††		2(N = 2551) *HD18 trial* (Borchmann 2017b) RATHL Trial (Johnson 2016)	No	Unknown[2]	No	No	**De-escalated treatments‡, PET *negative*** **HD18 trial (n=1005):** 4 vs. 6 to 8 eBEACOPP cycles • *Any second malignancy:* 2.6% vs. 3.6%; 5-year cumulative incidence estimate§: 3.3% (1.4%-5.3%) vs. 3.8% (1.9% to 5.7%) • *Discontinuation due to toxicity:* 0.4% vs. 2.2%; RD -2% (-3% to -0.4%) • *Any acute toxicity*: 90.8% vs. 96.2%; RD -5% (-8% to -2%);	⊕⊕⊕◯ MODERATE

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect Estimate (95% CI)	Quality (SOE)
							• *Treatment-related morbidity*: 40.7% vs. 63.7%; RR 0.6 (0.56–0.72); RD -23% (-29% to -17%) • *Any organ toxicity*: 7.6% vs. 18.1%; RD -10% (-15% to -6%); • *Anemia:* 38.9% vs. 54.4%; RD-15% (-22% to -9%) • *Thrombocytopenia*: 57.1% vs. 71.8%; RD -15% (-21% to -9%) • *Leukopenia:* 87.4% vs. 92.7%; RD -5% (-9% to -1.5% • *Infection:* 8.0% vs. 14.9%; RD -7% (-11% to -3%); • *Febrile neutropenia*: 21.8% vs. 28.8% RR 0.8 (0.6–0.9); RD -7% (-12% to -2%); • *Need for supportive measures*: Platelet infusion 24.4% vs. 40.5% (RD -16%; -22% to -10%); RBC infusion 46.5% vs. 65.9% (RD -19%; -25% to -13%) **RATHL trial ((n = 935)** • *Mortality from second malignancy: 1.3% vs. 0.9%; RD 0.4% (-1% to 2%)* • *Any second malignancy: 2.4% vs. 2.8%; RD -0.4% (-2% to 2%);* • *Any clinical adverse event**: 21% vs. 30.6%; RD -10% (-15% to -4%);* • *Any acute toxicity*: 65.4% vs. 68.8; RD -3% (-9% to 3%); • *Hematologic toxicities*: 59.7% vs. 59.8%; RD -0.1% (-6% to 6%) • *Thrombocytopenia*: 3.3% vs. 1.3%; RD 2% (0.1%–4%); • *Infection*: 10.3% vs. 14.5%; RD -4% (-8% to 0%); • *Febrile neutropenia:* 2.2% vs. 4.7%; RD -3% (-5% to -0.2%); • *Neurological event*: 3.1% vs. 4.9%; RD -2% (-4% to 1%) • *Pulmonary/respiratory event*: 0.7% vs. 3.2%; (RD -3% (-4% to -1%);	

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect Estimate (95% CI)	Quality (SOE)
							• *Vascular event*: 2.6% vs. 4.9%; RD -2% (-5% to 0.2%) • *Fatigue*: 1.1% vs. 3.0%; RD -2% (-4% to -0.1%) **Conclusion**: Overall, both trials report that most treatment-related toxicities were less common in patients who received de-escalated therapy compared with those who continued on more intense regimens; some were substantially less common.	

CI = Confidence Interval; eBEACOPP = escalated Bleomycin, Etoposide, Doxorubicin, Cyclophosphamide, Vincristine, Procarbazine, Prednisolone; HR = Hazard Ratio; NR = Not Reported; OS = Overall Survival; PET = Positron Emission Tomography; PFS = Progression Free Survival; RD = Risk Difference

* Escalated treatments: HD 0607 trial R-BEACOPP vs. BEACOPP after 2 ABVD cycles; HD18Trial R-eBEACOPP vs. eBEACOPP after 2 eBEACOPP cyles

†Borchmann: Only OS and PFS data from randomized patients is reported here; this includes patients randomized prior to the protocol amendment (i.e., who received 8 total cycles) which introduced a reduction of standard therapy from 8 to 6 cycles of eBEACOPP in total. Because the required sample size for the treatment group comparison in the PET(+) cohort had already been reached, randomization between eBEACOPP and R-eBEACOPP after PET(+) was stopped at this point, and all patients with a PET(+) were subsequently assigned to receive the new standard treatment of 6 × eBEACOPP.

‡De-escalation of treatment: RAHTL trial bleomycin was omitted vs. included per standard ABVD after 2 ABVD cycles; HD18Trial 2 cycles of eBEACOPP vs.6- 8 cycles of eBEACOPP after initial 2 eBEACOPP cycles

§ Accounting for death as competing risk.

**excluding blood or bone marrow events and laboratory events

††Grading of Adverse Events (AE) is according to the Common Terminology Criteria for Adverse Events (CTCAE) v3.0. Grades of 1-5 refer to the severity of the AE. Grade 1 Mild AE, Grade 2 Moderate AE, Grade 3 Severe AE, Grade 4 Life-threatening or disabling AE, Grade 5 Death related to AE.

Reasons for downgrading:

1. Serious risk of bias: the majority of studies violated one or more of the criteria for good quality RCT (or cohort study) related to the outcome reported (see Appendix for details) None of the included randomized controlled trials for this report randomized patients to receive PET/CT versus a strategy that did not include use of PET/CT, but rather, patients were randomized to various treatment options based on results of their PET/CT. There was no downgrade of these studies based on lack of comparison to a no PET/CT strategy. Risk of bias was assessed based on criteria in Appendix D and any down grade for risk of bias was based on these criteria.
2. Inconsistency: differing estimates of effects across trials; if point estimates across trials are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies were downgraded. Consistency is also unknown across PET-adapted studies given the differences between studies in initial as well as randomized treatment protocols based on PET results.
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from “mild” to “substantial”. If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for “mild/small” effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

5.1.4. Strength of evidence summary: Interim PET in adults with relapsed or refractory Hodgkin Lymphoma (Observational PET-Adapted studies)

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect Estimate (95% CI)	Quality (SOE)
Overall and (based on pre-ASCT PET)	Various	3 Observational Moskowitz 2012 (N = 97) 2015 and 2017 (N = 65)	Yes[1] (-1)	Unknown[2]	No	No	**2012 study (4.25 years)** OS: all patients 80%, ASCT 88%; data by PET/CT results are not provided **2015 (3.3 years)/2017 (2 years) cohorts;** OS all patients: 95% **Conclusions:** No firm conclusions regarding the impact of PET-adapted therapy are drawn.	⊕◯◯◯ Insufficient
Event Free Survival (based on pre-ASCT PET)							**2012 study (4.25 years) patients with ASCT** EFS: All patients 70%, ASCT, 70% EFS: PET + 28.6% vs. PET - >80% **2015 (3.3 years)/2017 (2 years) cohorts;** EFS all patients: 82% EFS: PET + 60% vs. PET – 85% (estimated)† **Conclusions:** Pre-ASCT PET positive status was associated with worse prognosis compared with PET negative status.	⊕◯◯◯ Insufficient
Treatment-related Toxicity‡	NR	2 Observational Moskowitz 2015/2017 (N = 65)	No	Unknown[2]	No		Conclusions: Treatment-related toxicities were not reported based on PET-adapted treatments received. Overall, the reported frequency of grade 3 or 4 toxicities was ≤2%. No conclusions regarding the impact of PET on toxicities are possible.	⊕◯◯◯ Insufficient

ASCT = Autologous Stem Cell Transplantation; EFS= event free survival; NR = Not Reported; OS = Overall Survival; PET = Positron Emission Tomography

* PET Positive: extended salvage chemotherapy of GVD or augICE followed later by HDT, RT and ASCT; PET negative directly to consolidation with high dose chemotherapy (HDT) and ASCT, with the potential for RT

†Estimated from author figures for patients receiving ASCT.

‡ Grading of Adverse Events (AE) is according to the Common Terminology Criteria for Adverse Events (CTCAE) v3.0. Grades of 1-5 refer to the severity of the AE. Grade 1 Mild AE, Grade 2 Moderate AE, Grade 3 Severe AE, Grade 4 Life-threatening or disabling AE, Grade 5 Death related to AE.

Reasons for downgrading:

1. Serious risk of bias: the majority of studies violated one or more of the criteria for good quality RCT (or cohort study) related to the outcome reported (see Appendix for details) None of the included randomized controlled trials for this report randomized patients to receive PET/CT versus a strategy that did not include use of PET/CT, but rather, patients were randomized to various treatment options based on results of their PET/CT. There was no downgrade of these studies based on lack of comparison to a no PET/CT strategy. Risk of bias was assessed based on criteria in Appendix D and any down grade for risk of bias was based on these criteria.
2. Inconsistency: differing estimates of effects across trials; If point estimates across trials are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was downgraded. Consistency is also unknown because of overlap of study populations (2015 and 2017) and use of the similar/the same treatment protocols.

3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

5.1.5. Strength of evidence summary: Interim PET in adults with aggressive Non-Hodgkin Lymphoma (PET-Adapted studies)

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect Estimate (95% CI)	Quality (SOE)
Escalated therapy								
PET-Adapted treatment RCT (Dührsen 2018)								
Probability of Overall Survival	2 years	1 trial (n=363, 60% stage III-IV, 86.4% DLBCL) Dührshen 2018 PETAL trial (ΔSUV_{max} method)	No	Unknown[2] Single study	No	No	**PET +: Burkitt protocol vs. R-CHOP (n=108)** 55.4% vs. 63.6%; HR 1.3 (0.8 to 2.4); Estimated difference: 7.9/100 **PET – : 4 R-CHOPP cycles vs. 4 R-CHOPP + 2 additional rituximab (n=255)** 87.2% vs. 88.2%; HR 0.9 (0.5–1.5); Estimated difference: 1/100 **Conclusion:** OS probability was lower in those with positive interim PET receiving the Burkitt's protocol versus additional R-CHOP. Results are not statistically significant but may be clinically important. Failure to reach the recruitment goal for PET+ patients may have impacted statistical power to detect survival differences. In those with negative interim PET, OS was similar between treatments; the additional doses of rituximab did not substantially improve OS.	⊕⊕⊕◯ **Moderate**
Probability of Progression Free Survival and EFS			No	Unknown[2] Single study	No	No	**PET +: Burkitt protocol vs. R-CHOP (n=108)** **PFS:** 43.1% vs. 49.4% Estimated difference: -6.3/100 **EFS:** 31.5% vs. 42.0%; Adj. HR 1.5 (0.9 to 2.5); Estimated Difference: 10.5/100 **PET – 4 R-CHOPP cycles vs. 4 R-CHOPP + 2 additional rituximab (n=255)**	⊕⊕⊕◯ **Moderate**

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect Estimate (95% CI)	Quality (SOE)
							PFS: 77.5% vs. 82.0%; Estimated difference: -4.5/100 **EFS**: 73.5 vs. 76.4; Adj. HR 1.0 (0.7 to 1.6), Estimated difference: -2.9/100 **Conclusion:** PFS and EFS were lower in PET-positive patients receiving the escalated treatment of Burkitt's protocol compared with those receiving additional R-CHOP; these were not statistically different but may be clinically important. In PET-negative patients, PFS and EFS were somewhat lower in those who received versus those who did not receive the additional rituximab doses. These differences may be clinically important.	
Treatment Toxicity (Grade 3 or4)†			No	Unknown[2] Single study	No	Yes[3] (-1)	**PET + (escalated) vs. R-CHOP (n = 108)** *Treatment-related mortality*: 5.4% vs. 3.9%; RD 2% (-6% to 9%) *Anemia:* 44.6% vs. 25.0%; RD 20% (2% to 37%); *Leukopenia: 80.4% vs. 59.6%;* RD 21% (4% to 38%); *Neutropenia: 33.9% vs. 23.1%;* RD 11% (-6% to 28%); *Thrombocytopenia: 58.9% vs. 21.2%;* RD 38% (21% to 55%) *Infection: 50% vs. 21.2%:* RD 29% (12% to 46%); *Mucositis: 37.5% vs. 11.5%;* RD 26% (11% to 41%) *Diarrhea:* 5.4% vs. 9.6%; RD -4% (-14% to 6%) *Creatinine:* 1.8% vs. 5.8%; RD -4% (-11% to 3%) **PET – 4 R-CHOPP cycles vs. 4 R-CHOPP + 2 additional rituximab (n = 255)** *Treatment-related mortality:* 3.9% vs. 1.6%; RD 2% (-2% to 6%); *Anemia:* 16.3% vs. 11.1%; RD 5% (-3% to 14%) *Leukopenia:* 54.3% vs. 60.3%; RD -6% (-18% to 6%) *Neutropenia:* 15.5% vs. 23.8%; RD -8% (-18% to 1%) *Thrombocytopenia:* 14.7% vs. 7.9%; RD 7% (-1% to 15%) *Infection: 14.7% vs. 11.1%;* RD 4% (-5% to 12%) *Mucositis:* 1.6% vs. 2.4%; RD -1% (-4% to 3%); *Diarrhea:* 3.1% vs. 0.8%; RD 2% (-1% to 6%)	⊕⊕◯◯ **Low**

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect Estimate (95% CI)	Quality (SOE)
							Creatinine: 3.1% vs. 0.8%; RD 2% (-1% to 6%) **Conclusions:** **PET +:** Escalated treatment with the Burkitt protocol was associated with an increased risk of a number of treatment related toxicities. This arm was discontinued early and there may not have been sufficient power to detect differences between treatments for rare events (e.g. treatment-related death). **PET –:** Compared with those who had R-CHOP only, those receiving additional doses of rituximab to R-CHOP may have had clinically important higher risk of some treatment-related toxicities (leukopenia, neutropenia), but lower risk of others (thrombocytopenia, infection). Associations may not be beyond what might be expected by chance.	
RCT of induction treatment (Casasnovas 2017)								
Probability of Overall Survival	4 years	1 trial (N = 211, 97% stage III-IV) Casanovas 2017 LNH2007-3B trial *Initial therapy randomized:* R-ACVBP vs. R-CHOP14 **Interim PET at the end of 2 (PET 2) and 4 cycles (PET4);**	Yes[1] (-1)	Unknown[2] Single study	No	No	**Escalated (PET 2+, PET4 -) vs. SIC (PET 2- and 4 -)** (n=48): 90.4% (81%-95.1%) vs. (n=52): 89.6% (85%-92.2%); p = 0.21 Estimated difference: 0.8/100 **PET 4 + (overall) vs. PET 4 - (overall)** (n=100): 80% (69%-87.5%) vs. (n=100): 88.9% (82.1%-94.4%); p=0.08 Estimated difference: 8.9/100 **Conclusions:** OS was similar between those who received escalated therapy and those who had standard immunochemotherapy (SIC) consolidation suggesting that SCI is sufficient in patients with a negative interim PET after 4 cycles, even if PET after 2 cycles was positive.	⊕⊕◯◯ LOW

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect Estimate (95% CI)	Quality (SOE)
		IHP criteria for positivity					With regard to prognosis, lower OS (estimated difference of ~9/100) in patients with a positive PET 4 scan compared with those whose PET 4 scan was negative was seen.	
Probability of Progression Free Survival		)	Yes[1] (-1)	Unknown[2] Single study	No	Yes[3] (-1)	**Escalated (PET 2+, PET4 -) vs. SIC (PET 2- and 4 -)** (n=48): 85% (71.1%-92.6%) vs. (n=52) 75% (60.9%-84.6%); Estimated difference: 10/100 **PET 4 + (overall) vs. PET 4 - (overall)** (n=100): 72.9% (63.1%-80.6%) vs. 79.8 (79.4-86.4); Estimated difference 6.9/100 **Conclusion**: The probability of PFS was higher in those who received escalated therapy versus those who had standard immunochemotherapy (SIC) consolidation and may suggest improved PFS with escalated treatment. Prognosis for those who had PET 4 + results was slightly lower vs. PET 4 -; the estimated difference was 6.9/100.	⊕⊕◯◯ LOW
Treatment Toxicity	NR		Yes[1] (-1)	Unknown[2] Single study	Yes (-1) (outcomes by PET-adapted strategy not reported)	Yes[3] (-1)	No toxicity data relevant to impact of PET on treatment decisions was provided in the randomized trial of initial treatment. Authors compare toxicities for R-ACVBP and R-CHOP noting that the former was associated with more frequent infections (64% vs. 27%) and need for transfusions versus R-CHOP. Given that patients in both PET + and PET – groups received each of these treatments initially, specific treatment regimens for those with a positive PET 4 scan were not provide and no toxicity data on use of escalated treatment strategies were provided, conclusions related to the impact of PET on reducing toxicity from this trial are not possible.	⊕◯◯◯ Insufficient
RCT Pet-adapted treatment; initial randomization continuation (Lamy 2018)								
Probability of Overall Survival	5 years	1 trial (n = 281,negative interim PET)	No	Unknown[2] Single study	No	No	**R-CHOP vs. R-CHOP + RT** **PET negative (n = 281)**	⊕⊕⊕◯ **Moderate**

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect Estimate (95% CI)	Quality (SOE)
		Lamy 2018 (*LYSA/ GOELAMS);* (low risk, Stage I or II w/o bulky mass)					90% vs. 92% (estimated difference 2/100) (NS, no data provided) **Conclusion:** In persons with a negative interim PET, OS was similar between groups; it is unclear if the estimated 2/100 difference is clinically important. Deletion of RT may not adversely impact OS.	

CI = Confidence Interval; DLBCL = Diffuse large B-cell lymphoma; eBEACOPP = escalated Bleomycin, Etoposide, Doxorubicin, Cyclophosphamide, Vincristine, Procarbazine, Prednisolone; EFS= event free survival; HR = Hazard Ratio; IHP = International Harmonization Project; NR = Not Reported; OS = Overall Survival; PET = Positron Emission Tomography; PFS = Progression Free Survival; R-ACVBP = rituximab, doxorubicin, cyclophosphamide, vindesine, bleomycin, and prednisone; R-CHOP = rituxan, cyclophosphamide, doxorubicin, vincristine, and prednisone; RD = Risk Difference; SIC = Standard Immunochemotherapy

*PETAL Randomized treatments based in interim PET: PET + persons six blocks of an intensive Burkitt's lymphoma or six additional cycles of R-CHOP; PET – persons with CD20 positive lymphomas, four additional cycles of R-CHOP or the same treatment with two additional doses rituximab.

LNH2007-3B trial: Treatment escalated in all patients with positive PET at 2 and/or 4 cycles. PET (+) cycle 2/PET (–) cycle 4, received 2 cycles of high-dose methotrexate, then BEAM or Z-BEAM followed by ASCT; for patients with positive PET at cycle 4, final treatment decision left to local investigator. Individuals with negative PET results at both time frames, received standard immunochemotherapy (SIC) consolidation.

LYSA/GOELAMS Trial: Patients were randomized after initial staging to receive R-CHOP or R-CHOP with RT; those with negative interim PET (complete remission) continued in these random assignments; Authors ITT analysis includes all patients as randomized regardless of interim PET results (38 patients had positive interim PET) are incl0075 vded in the full report.

†Grading of Adverse Events (AE) is according to the Common Terminology Criteria for Adverse Events (CTCAE) v3.0. Grades of 1-5 refer to the severity of the AE. Grade 1 Mild AE, Grade 2 Moderate AE, Grade 3 Severe AE, Grade 4 Life-threatening or disabling AE, Grade 5 Death related to AE.

Reasons for downgrading:

1. Serious risk of bias: the majority of studies violated one or more of the criteria for good quality RCT (or cohort study) related to the outcome reported (see Appendix for details). None of the included randomized controlled trials for this report randomized patients to receive PET/CT versus a strategy that did not include use of PET/CT. This trial randomized patients to different induction treatments. There was no downgrade of these studies based on lack of comparison to a no PET/CT strategy. Risk of bias was assessed based on criteria in Appendix D and any down grade for risk of bias was based on these criteria.
2. Inconsistency: differing estimates of effects across trials; If point estimates across trials are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies were downgraded. Consistency is also unknown across PET-adapted studies given the differences between studies in initial as well as randomized treatment protocols based on PET results.
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded.

5.1.6. Strength of evidence summary: End of Treatment PET in adults with advanced-stage HL and aggressive NHL

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect estimate (95% CI)	Quality (SOE)
Advanced-Stage HL								
Probability of Overall Survival	3 years	1 trial (n=296)* GITIL/FIL HD 0607 Trial (Gallamini 2018)	Yes1 (-1)	Unknown[2]	No	No	**Consolidation RT or no RT (PET -, with baseline large nodal mass)*** HD0607: 100% vs. 99% (95%-100%), Estimated difference: 1/100 **Conclusions:** Consolidation RT did not provide substantial clinical benefit in terms of OS.	⊕⊕◯◯ LOW
Probability of Progression Free Survival			Yes1 (-1)	Unknown[2]	No	No	**Consolidation RT or no RT (PET -, with baseline large nodal mass, n=263)*** HD0607: 97% (92%-99%) vs.93% (87%-96%) Estimated difference: 4/100 Patients without LNM at baseline not randomized to RT (n=260): 92% (88 to 85%) **Conclusion:** Although PFS was slight better in those receiving RT, the difference was not beyond what might be expected by chance. PFS was not substantially different in those has no baseline LNM.	⊕⊕◯◯ LOW
Aggressive NHL (DBLCL)								
Probability of Overall Survival	2 years	1 study (n=50 from positive interim PET group)† PETAL Trial (Duhrsen 2018	Yes[1] (-1)	Unknown[2]	No	Yes[3](-1)	**EOT PET positive (n=31):** Observation (n=18 asymptomatic vs. additional treatment (n=13 clinical evidence of disease)† 88.9% (62.4% to 97.1) vs. 33.3% (10.3% to 58.8%) **EOT PET (-) n = 18 vs. PET(+), persistent abnormalities, n=31)** 100% vs. 65.9% (95% CI, 45.8% to 80.0%) **Conclusion:** Prognosis with respect to OS was better in patients with a negative EOT PET versus those with persistent abnormalities; Among those with positive	⊕◯◯◯ INSUFFICIENT

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect estimate (95% CI)	Quality (SOE)
							EOT PET, asymptomatic patients who were observed had better prognosis than those who had persistent PET findings in addition to clinical evidence of disease.	
Probability of Progression Free Survival	2 years	1 Study (n =50 from postiive interim PET group)† PETAL Trial (Duhrsen 2018 (IHP criteria used)	Yes1 (-1)	Unknown[2]	No	Yes[3](-1)	**EOT PET positive (n =31):** Observation (n=18 asymptomatic vs. additional treatment (n=13 clinical evidence of disease)† 66.7% (40.4% to 83.4%) vs. 26.4% (0.7% to 52.2%) **EOT PET (-) n = 19 vs. PET(+, persistent abnormalities, n=31)** 84.2% (58.7% to 94.6%) vs. 50.5% (31.8% to 66.6%) **Conclusion:** PFS was better in patients with a negative EOT PET versus those with persistent abnormalities; Among those with positive EOT PET, asymptomatic patients who were observed had better prognosis than those who had persistent PET findings in addition to clinical evidence of disease.	⊕◯◯◯ INSUFFICIENT

LNM = large nodal mass RT = radiation therapy; EOT= end-of-treament;

*HD 0607 Treatments: 296 patients who had large nodal masses at baseline and whose interim and end of treatment PET were both negative were randomized to radiotherapy or no further treatment base on PET/CT done after completion of first line therapy (6 cycles of ABVD)

† PETAL Trial treatments were not randomly assigned: End-of-treatment PET was restricted to patients with a positive interim scan who did **not** experience progression on therapy; 19 were PET(-) and 31 were PET(+) and had persistent abnormalities on EOT PET; Residual masses (CT) or residual FDG uptake were NOT considered lymphoma unless there was clinical evidence of disease. Of the 31 PET(+) patients, 18 (with no other evidence of active disease) received observation while 13 (with active disease) received additional treatment

Reasons for downgrading:

1. Serious risk of bias: the majority of studies violated one or more of the criteria for good quality RCT (or cohort study) related to the outcome reported (see Appendix for details) None of the included randomized controlled trials for this report randomized patients to receive PET/CT versus a strategy that did not include use of PET/CT, but rather, patients were randomized to various treatment options based on results of their PET/CT. There was no downgrade of these studies based on lack of comparison to a no PET/CT strategy. Risk of bias was assessed based on criteria in Appendix D and any down grade for risk of bias was based on these criteria. In the PETAL trial EOT PET-adapted treatments were not randomized. Risk of bias for a given study may differ by outcome.

2. Inconsistency: differing estimates of effects across trials; If point estimates across trials are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies were downgraded. Consistency is also unknown across PET-adapted studies given the differences between studies in initial as well as randomized treatment protocols based on PET results.

3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.

4. Indirect, intermediate or surrogate outcomes may be downgraded.

5.1.7. Strength of evidence summary: PET for Surveillance and Evaluation of Relapse in adults (Survival Outcomes)

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion Effect Estimate (95% CI)	Quality (SOE)
Hodgkin Lymphoma								
Probability of Overall Survival	5 years	1 retrospective cohort (N=241) *Pingali 2014*	Yes[1] (-1)	Unknown[2]	No	No	**Routine/surveillance vs. imaging if indicated** 97% (92% to 99%) vs. 96% (86% to 99%), p=0.41 **Conclusions:** OS was similar for asymptomatic patients in remission who received surveillance PET/CT and those who received it only if clinically indicated	⊕◯◯◯ INSUFFICIENT
Probability of Progression Free Survival	3, 5 years	1 retrospective cohort (N=368) *Dann 2014*	Yes[1] (-1)	Unknown[2]	No	No	**Routine/surveillance vs. imaging if indicated** 3 year PFS: 93% (NR) vs. 86% (NR), p=NS 5 year PFS: 92% (NR) vs. 85% (NR), p= NS **Conclusion**: Authors report no statistical difference in PFS at 3 or 6 years when imaging was routinely done vs. for clinical suspicion of relapse	⊕◯◯◯ INSUFFICIENT
Relapsed, Refractory Aggressive NHL (DLBCL)								
Probability of Overall Survival From time of ASCT	Time from ASCT	1 retrospective cohort (N=45) *Epperla 2016*	Yes[1] (-1)	Unknown[2]	No	Yes[3](-1)	**Routine/surveillance* vs. clinical follow-up/lab** Median years: from time of transplantation† 1.8 (range 0.4–12.4) vs. 1.6 (range 0.4–8.4) Median years: from time of relapse post ASCT† 1.0 (range 0.02–11.6) vs. 0.34 (range 0.01–2.6) **Conclusion**: Authors report no statistical difference in median survival; sample size is small.	⊕◯◯◯ INSUFFICIENT

ASCT = autologous stem cell transfusion; CI = confidence interval; OS = overall survival, PET/CT = positron emission tomography/computed tomography
* Routine CT and/or PET

†Overall survival (OS) was defined as the time from autologous hematopoietic cell transplantation to date of death or last follow-up. Post-relapse survival was defined as time from relapse (after auto-HCT) to date of death or last follow-up.

Reasons for downgrading:

1. Serious risk of bias: the majority of studies violated one or more of the criteria for good quality RCT (or cohort study) related to the outcome reported (see Appendix for details) None of the included randomized controlled trials for this report randomized patients to receive PET/CT versus a strategy that did not include use of PET/CT, but rather, patients were randomized to various treatment options based on results of their PET/CT. In the PETAL trial EOT PET-adapted treatments were not randomized. Risk of bias for a given study may differ by outcome.
2. Inconsistency: differing estimates of effects across trials; If point estimates across trials are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies were downgraded. Consistency is also unknown across PET-adapted studies given the differences between studies in initial as well as randomized treatment protocols based on PET results.
3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial". If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.
4. Indirect, intermediate or surrogate outcomes may be downgraded

5.2. Strength of Evidence Summary: PET/CT Safety and Adverse Events Results

5.2.1. Strength of evidence summary: Safety

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion	Quality (SOE)
Effective and Cumulative Effective Radiation Dose	Adults *Surveillance*: range, median 2–5 years *Treatment*: 8 months Pediatric 2 years	5 studies total N=1071 Adults (4 studies), N=972 Pediatric (1 study), N=99 Adults 1 RCT, N=300 (Picardi 2014) 2 comparative observational studies, N=186	Yes[1] (-1)	No	No	No	**Adults** *Effective radiation dose** Range of mean effective radiation doses for PET/CT during: • Surveillance (1 RCT, 2 observational): 14.5 mSv to 26.1 mSv • Treatment (1 observational): 26.1 mSv *Cumulative effective radiation dose* Median 69 mSv (IQR 42-118) per patient from all ionizing radiation imaging (1 case series): • PET/CT: 8% of total (~5.5 mSv) • CT with contrast: 89% of total (~61.4 mSv)	⊕⊕◯◯ Low

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion	Quality (SOE)
		(Guttikonda 2014, Patel 2013) 1 case series N=486 (Crowley 2016) Pediatric 1 comparative observational, N=99 (Rathore 2012)					**Pediatric** One study reports effective radiation doses from FDG alone are approximately 6.4 and 8.6 mSv for a 10- and 15-year-old child, respectively. Conclusion: Mean effective radiation doses from PET/CT varied across studies based on number of scans performed, length of follow-up and other factors such as patient's age. RSNA and ACR estimate the effective dose for a PET/CT is 25 mSv in adults.	
Additional Testing during relapse assessment or surveillance	3–5 years	2 studies total, N=406 (adults) 1 RCT, N=300 (Picardi 2014) 1 comparative observational, N=106 (Hong 2014)	Yes[1] (-1)	No	No	No	*Additional testing due to false positives* • *1 RCT*: High number of false positives (data not provided) obtained with PET/CT compared with US/chest radiography resulted in greater need for mediastinotomies: 50% (20/40) vs. 15% (6/40), p=0.04 • *1 observational*: False positives for PET/CT vs. CT: 20% (21/106) vs. 7% (7/106); leading to additional testing in 14% (15/106) vs. 6% (6/106) Conclusion: High false positive rates reported led to additional perhaps unnecessary diagnostic testing.	⊕⊕◯◯ Low
Relapse identification during surveillance	4–5 years	3 studies total, N=909 (adults) 1 RCT, N=300 (Picardi 2014)	Yes[1] (-1)	No	No	No	PET/CT vs. US/radiography (1 RCT) • Effectiveness at detecting relapse: 100% vs. 95.7% • Estimated radiation dose to detect one relapse: 449.5 mSv vs. 10.2 mSv (p=0.0002)	⊕⊕◯◯ Low

Outcome	Follow-up	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Conclusion	Quality (SOE)
		2 comparative observational, N=609 (Dann 2014, Pingali 2014)					Routine PET/CT imaging versus clinically-indicated imaging† (2 observational studies): • Number of imaging tests needed to diagnose/detect a single relapse: 48 and 127 vs. 5 and 15 Conclusion: US/radiography was as effective at identifying relapse with significantly lower radiation exposure. Fewer imaging tests are needed to identify a relapse when done as clinically indicated vs. as a routine practice.	

ACR = American College of Radiology; CHOP = cyclophosphamide, doxorubicin, vincristine, prednisolone; FDG = Fludeoxyglucose; PET/CT = positron emission tomography/computed tomography; R = rituximab; RSNA = Radiological Society of North America

*Mean effective radiation dose for FDG PET alone versus combined PET/CT for surveillance was 14.1 mSv vs. 26.1 mSv in one observational study (Patel 2013).

†Imaging done only when indicated by signs or symptoms; imaging was usually contrast enhanced CT.

Reasons for downgrading:

1. Serious risk of bias: the majority of studies violated one or more of the criteria for good quality RCT (or cohort study) related to the outcome reported (see Appendix for details). One RCT was identified that compared PET/CT with ultrasound/chest radiography.

2. Inconsistency: differing estimates of effects across trials; If point estimates across trials are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency. The consistency of single studies is unknown; evidence from single studies was downgraded. Consistency is also unknown because of overlap of study populations (2015 and 2017) and use of the similar/the same treatment protocols.

3. Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention; If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from “mild” to “substantial”. If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for “mild/small” effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.

4. Indirect, intermediate or surrogate outcomes may be downgraded.

5.3. Strength of Evidence Summary: Differential Efficacy and Harms

No studies were identified that met the inclusion criteria.

5.4. Strength of Evidence Summary: Cost-Effectiveness

One poor quality cost-effectiveness analysis (CEA)[43] for initial staging with PET/CT in adult patients with HL and one moderate-quality cost-utility analysis (CUA)[93] for use of PET/CT for surveillance in adults with diffuse large B-cell lymphoma (DLBCL)were identified. There currently is no method for evaluating and reporting strength of evidence across economic studies. Major finding are reported in the previous results section.

References

1. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality. January 2014. Chapters available at: www.effectivehealthcare.ahrq.gov.
2. Surveillance, Epidemiology, and End Results (SEER) Program, National Cancer Institute. Cancer Stat Facts: Hodgkin Lymphoma. at https://seer.cancer.gov/statfacts/html/hodg.html.)
3. Surveillance, Epidemiology, and End Results (SEER) Program, National Cancer Institute. Cancer Stat Facts: non-Hodgkin Lymphoma. at https://seer.cancer.gov/statfacts/html/nhl.html.)
4. Centers for Medicare and Medicaid Services (CMS). Medicare National Coverage Determination (NCD) for Positron Emission Tomography (FDG) for Oncologic Conditions 2014. (Accessed 05/01/18, at https://www.cms.gov/medicare-coverage-database/details/ncd-details.aspx?ncdid=331.)
5. NICE Guidelines: Non-Hodgkin's lymphoma: diagnosis and management: NICE; 2016 07/20/16.
6. BlueCross BlueShield of Minnesota Medical Policy: Positron Emission Tomography (PET), V-27-005. 2017. (Accessed 05/01/18, at https://www.bluecrossmnonline.com/minnesota-medical-policy/external-medical-policy/minnesota-medical-policy/V-27-005.html.)
7. United Health Care. Coverage Summary: Positron Emission Tomography (PET)/Combined PET-CT (Computed Tomography). United HealthCare Services, Inc., 2017. (Accessed 05/01/18, at https://www.uhcprovider.com/content/dam/provider/docs/public/policies/medadv-coverage-sum/positron-emission-tomography-pet-combined-pet-ct-computed-tomography.pdf.)
8. Aetna. Clinical Policy Bulletins: Positron Emission Tomography (PET), No. 0071. Aetna, 2018. (Accessed 07/06/18, at http://www.aetna.com/cpb/medical/data/1_99/0071.html.)
9. AIM Specialty Health Clinical Appropriateness Guidelines: Positron Emission Testing, Other PET Applications, including Oncologic Tumor Imaging 2018 AIM Specialty Health; 2018 03/12/18.
10. Alberta Health Services Clinical Practice Guideline 2018 - Adult Lymphomas. Edmonton, AB: Alberta Health Services; 2018 01/2018.
11. Blue Cross Blue Shield Louisiana Medical Policies: Positron Emission Tomography (PET) Oncology Applications. 2018. (Accessed 05/01/18, at https://www.bcbsla.com/-/media/Medical-Policies/2017/04/27/20/07/PositronEmissionTomographyPETOncologyApplications.pdf.)
12. Abou-Nassar KE, Vanderplas A, Friedberg JW, et al. Patterns of use of 18-fluoro-2-deoxy-D-glucose positron emission tomography for initial staging of grade 1-2 follicular lymphoma and its impact on initial treatment strategy in the National Comprehensive Cancer Network Non-Hodgkin Lymphoma Outcomes database. Leukemia & lymphoma 2013;54:2155-62.
13. Adams HJ, Kwee TC. Prognostic value of pretransplant FDG-PET in refractory/relapsed Hodgkin lymphoma treated with autologous stem cell transplantation: systematic review and meta-analysis. Annals of hematology 2016b;95:695-706.
14. Adams HJ, Kwee TC. Prognostic value of interim FDG-PET in R-CHOP-treated diffuse large B-cell lymphoma: systematic review and meta-analysis. Critical reviews in oncology/hematology 2016c;106:55-63.
15. Adams HJ, Kwee TC. Pretransplant FDG-PET in aggressive non-Hodgkin lymphoma: systematic review and meta-analysis. European journal of haematology 2017b;98:337-47.

16. Adams HJ, Kwee TC, de Keizer B, Fijnheer R, de Klerk JM, Nievelstein RA. FDG PET/CT for the detection of bone marrow involvement in diffuse large B-cell lymphoma: systematic review and meta-analysis. European journal of nuclear medicine and molecular imaging 2014;41:565-74.

17. Adams HJ, Nievelstein RA, Kwee TC. Outcome of Hodgkin Lymphoma Patients With a Posttreatment 18F-Fluoro-2-Deoxy-d-Glucose Positron Emission Tomography (FDG-PET)-Negative Residual Mass: Systematic Review and Meta-analysis. Pediatric hematology and oncology 2015a;32:515-24.

18. Adams HJ, Nievelstein RA, Kwee TC. Prognostic value of complete remission status at end-of-treatment FDG-PET in R-CHOP-treated diffuse large B-cell lymphoma: systematic review and meta-analysis. British journal of haematology 2015b;170:185-91.

19. Adams HJ, Nievelstein RA, Kwee TC. Prognostic value of interim and end-of-treatment FDG-PET in follicular lymphoma: a systematic review. Annals of hematology 2016d;95:11-8.

20. Adams HJ, Nievelstein RA, Kwee TC. Systematic Review on the Additional Value of 18F-Fluoro-2-Deoxy-D-Glucose Positron Emission Tomography in Staging Follicular Lymphoma. J Comput Assist Tomogr 2017a;41:98-103. doi: 10.1097/RCT.0000000000000485.

21. Adams HJA, Kwee TC. Proportion of false-positive lesions at interim and end-of-treatment FDG-PET in lymphoma as determined by histology: Systematic review and meta-analysis. European journal of radiology 2016;85:1963-70.

22. Adams HJA, Kwee TC. Proportion of false-positive lesions at interim and end-of-treatment FDG-PET in lymphoma as determined by histology: Systematic review and meta-analysis. European journal of radiology 2016a;85:1963-70.

23. Adams HJA, Kwee TC. Fact Sheet About Interim and End-of-Treatment (18)F-FDG PET/CT in Lymphoma. Journal of nuclear medicine : official publication, Society of Nuclear Medicine 2017;58:1178-9.

24. Adams HJA, Kwee TC. Interim FDG-PET in lymphoma, a questionable practice in hematology. European journal of nuclear medicine and molecular imaging 2017;44:2014-7.

25. Amitai I, Gurion R, Vidal L, Dann EJ, Raanani P, Gafter-Gvili A. PET-adapted therapy for advanced Hodgkin lymphoma - systematic review. Acta Oncol 2018;18:1-8.

26. Andre MPE, Girinsky T, Federico M, et al. Early Positron Emission Tomography Response-Adapted Treatment in Stage I and II Hodgkin Lymphoma: Final Results of the Randomized EORTC/LYSA/FIL H10 Trial. J Clin Oncol 2017;35:1786-94. doi: 10.200/JCO.2016.68.6394. Epub 2017 Mar 14.

27. Angelopoulou MK, Mosa E, Pangalis GA, et al. The Significance of PET/CT in the Initial Staging of Hodgkin Lymphoma: Experience Outside Clinical Trials. Anticancer Res 2017;37:5727-36.

28. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. Bmj 2004;328:1490.

29. Awan UE, Siddiqui N, SaadUllah M, et al. FDG-PET scan in assessing lymphomas and the application of Deauville Criteria. JPMA The Journal of the Pakistan Medical Association 2013;63:725-30.

30. Bakhshi S, Bhethanabhotla S, Kumar R, et al. Posttreatment PET/CT Rather Than Interim PET/CT Using Deauville Criteria Predicts Outcome in Pediatric Hodgkin Lymphoma: A Prospective Study Comparing PET/CT with Conventional Imaging. J Nucl Med 2017;58:577-83. doi: 10.2967/jnumed.116.176511. Epub 2016 Oct 6.

31. Bakhshi S, Radhakrishnan V, Sharma P, et al. Pediatric nonlymphoblastic non-Hodgkin lymphoma: baseline, interim, and posttreatment PET/CT versus contrast-enhanced CT for evaluation--a prospective study. Radiology 2012;262:956-68.

32. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011;64:401-6.

33. Barrington SF, Kirkwood AA, Franceschetto A, et al. PET-CT for staging and early response: results from the Response-Adapted Therapy in Advanced Hodgkin Lymphoma study. Blood 2016;127:1531-8. doi: 10.182/blood-2015-11-679407. Epub 2016 Jan 8.

34. Barrington SF, Mikhaeel NG, Kostakoglu L, et al. Role of imaging in the staging and response assessment of lymphoma: consensus of the International Conference on Malignant Lymphomas Imaging Working Group. Journal of clinical oncology 2014;32:3048.

35. Bednaruk-Mlynski E, Pienkowska J, Skorzak A, et al. Comparison of positron emission tomography/computed tomography with classical contrast-enhanced computed tomography in the initial staging of Hodgkin lymphoma. Leukemia & lymphoma 2015;56:377-82.

36. Biggi A, Gallamini A, Chauvie S, et al. International validation study for interim PET in ABVD-treated, advanced-stage hodgkin lymphoma: interpretation criteria and concordance rate among reviewers. Journal of nuclear medicine : official publication, Society of Nuclear Medicine 2013;54:683-90.

37. Borchmann P, Goergen H, Kobe C, et al. PET-guided treatment in patients with advanced-stage Hodgkin's lymphoma (HD18): final results of an open-label, international, randomised phase 3 trial by the German Hodgkin Study Group. The Lancet 2017b;390:2790-802.

38. Borchmann P, Haverkamp H, Lohri A, et al. Progression-free survival of early interim PET-positive patients with advanced stage Hodgkin's lymphoma treated with BEACOPPescalated alone or in combination with rituximab (HD18): an open-label, international, randomised phase 3 study by the German Hodgkin Study Group. Lancet Oncol 2017a;18:454-63. doi: 10.1016/S470-2045(17)30103-1. Epub 2017 Feb 22.

39. Carras S, Dubois B, Senecal D, et al. Interim PET Response-adapted Strategy in Untreated Advanced Stage Hodgkin Lymphoma: Results of GOELAMS LH 2007 Phase 2 Multicentric Trial. Clin Lymphoma Myeloma Leuk 2018;18:191-8. doi: 10.1016/j.clml.2018.01.003. Epub Jan 31.

40. Carrillo-Cruz E, Marin-Oyaga VA, de la Cruz Vicente F, et al. Role of 18F-FDG-PET/CT in the management of marginal zone B cell lymphoma. Hematol Oncol 2015;33:151-8. doi: 10.1002/hon.2181. Epub 014 Nov 19.

41. Casasnovas RO, Ysebaert L, Thieblemont C, et al. FDG-PET-driven consolidation strategy in diffuse large B-cell lymphoma: final results of a randomized phase 2 study. Blood 2017;130:1315-26. doi: 10.182/blood-2017-02-766691. Epub 2017 Jul 12.

42. Casulo C, Schoder H, Feeney J, et al. 18F-fluorodeoxyglucose positron emission tomography in the staging and prognosis of T cell lymphoma. Leukemia & lymphoma 2013;54:2163-7.

43. Cerci JJ, Trindade E, Buccheri V, et al. Consistency of FDG-PET accuracy and cost-effectiveness in initial staging of patients with Hodgkin lymphoma across jurisdictions. Clinical lymphoma, myeloma & leukemia 2011;11:314-20.

44. Chaganti S, Illidge T, Barrington S, et al. Guidelines for the management of diffuse large B-cell lymphoma. Br J Haematol 2016;174:43-56.

45. Cheah CY, Dickinson M, Hofman MS, et al. Limited clinical benefit for surveillance PET-CT scanning in patients with histologically transformed lymphoma in complete metabolic remission following primary therapy. Annals of hematology 2014;93:1193-200.

46. Chen YK, Yeh CL, Tsui CC, Liang JA, Chen JH, Kao CH. F-18 FDG PET for evaluation of bone marrow involvement in non-Hodgkin lymphoma: a meta-analysis. Clin Nucl Med 2011;36:553-9.

47. Cheng G, Alavi A. Value of 18F-FDG PET versus iliac biopsy in the initial evaluation of bone marrow infiltration in the case of Hodgkin's disease: a meta-analysis. Nuclear medicine communications 2013;34:25-31.

48. Cheng G, Servaes S, Zhuang H. Value of (18)F-fluoro-2-deoxy-D-glucose positron emission tomography/computed tomography scan versus diagnostic contrast computed tomography in initial staging of pediatric patients with lymphoma. Leukemia & lymphoma 2013;54:737-42.

49. Cheson BD. PET/CT in Lymphoma: Current Overview and Future Directions. Semin Nucl Med 2018;48:76-81.

50. Cheson BD, Ansell S, Schwartz L, et al. Refinement of the Lugano Classification lymphoma response criteria in the era of immunomodulatory therapy. Blood 2016;128:2489-96.

51. Cheson BD, Fisher RI, Barrington SF, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2014;32:3059-68.

52. Cohen JB, Behera M, Thompson CA, Flowers CR. Evaluating surveillance imaging for diffuse large B-cell lymphoma and Hodgkin lymphoma. Blood 2017;129:561-4.

53. Commonwealth of Australia Medical Services Advisory Committee (MSAC). Postiron emission tomography for lymphoma. 2010.

54. Commonwealth of Australia Medical Services Advisory Committee (MSAC). 18F-FDG PET for indolent non-Hodgkin’s lymphoma. 2016.

55. Cronin CG, Swords R, Truong MT, et al. Clinical utility of PET/CT in lymphoma. AJR Am J Roentgenol 2010;194:W91-W103.

56. Crowley MP, O'Neill SB, Kevane B, et al. Ionizing radiation exposure as a result of diagnostic imaging in patients with lymphoma. Clin Transl Oncol 2016;18:533-6. doi: 10.1007/s12094-015-1394-8. Epub 2015 Aug 26.

57. Dabaja BS, Advani R, Hodgson DC, et al. ACR appropriateness Criteria® diffuse large b-cell lymphoma. American journal of clinical oncology 2015;38:610-20.

58. Dann EJ, Bairey O, Bar-Shalom R, et al. Modification of initial therapy in early and advanced Hodgkin lymphoma, based on interim PET/CT is beneficial: a prospective multicentre trial of 355 patients. Br J Haematol 2017;178:709-18. doi: 10.1111/bjh.14734. Epub 2017 Jun 7.

59. Dann EJ, Berkahn L, Mashiach T, et al. Hodgkin lymphoma patients in first remission: routine positron emission tomography/computerized tomography imaging is not superior to clinical follow-up for patients with no residual mass. Br J Haematol 2014;164:694-700.

60. Deau B, Franchi P, Briere J, et al. PET2-driven de-escalation therapy in 64 high-risk Hodgkin Lymphoma patients treated with escalated BEACOPP. British journal of haematology 2015;170:278-81.

61. Dhakal S, Advani R, Ballas LK, et al. ACR Appropriateness Criteria® Hodgkin Lymphoma-Favorable Prognosis Stage I and II. American journal of clinical oncology 2016;39:535-44.

62. Dreyling M, Ghielmini M, Rule S, Salles G, Vitolo U, Ladetto M. Newly diagnosed and relapsed follicular lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of oncology : official journal of the European Society for Medical Oncology 2016;27:v83-v90.

63. Duhrsen U, Muller S, Hertenstein B, et al. Positron Emission Tomography-Guided Therapy of Aggressive Non-Hodgkin Lymphomas (PETAL): A Multicenter, Randomized Phase III Trial. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2018;36:2024-34.

64. Eichenauer D, Aleman B, André M, et al. Hodgkin lymphoma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Annals of Oncology 2018.

65. Engert A, Goergen H, Markova J, et al. Reduced-Intensity Chemotherapy in Patients With Advanced-Stage Hodgkin Lymphoma: Updated Results of the Open-Label, International, Randomised Phase 3 HD15 Trial by the German Hodgkin Study Group. HemaSphere 2017;1:e5.

66. Engert A, Haverkamp H, Kobe C, et al. Reduced-intensity chemotherapy and PET-guided radiotherapy in patients with advanced stage Hodgkin's lymphoma (HD15 trial): a randomised, open-label, phase 3 non-inferiority trial. Lancet (London, England) 2012;379:1791-9.

67. Epperla N, Shah N, Hamadani M, et al. Impact of Routine Surveillance Imaging on Outcomes of Patients With Diffuse Large B-Cell Lymphoma After Autologous Hematopoietic Cell Transplantation. Clin Lymphoma Myeloma Leuk 2016;16:672-8. doi: 10.1016/j.clml.2016.08.018. Epub Aug 29.

68. Facey K, Bradbury I, Laking G, Payne E. Overview of the clinical effectiveness of positron emission tomography imaging in selected cancers. Health technology assessment (Winchester, England) 2007;11:iii-iv, xi-267.

69. Figura N, Flampouri S, Mendenhall NP, et al. Importance of baseline PET/CT imaging on radiation field design and relapse rates in patients with Hodgkin lymphoma. Adv Radiat Oncol 2017;2:197-203. doi: 10.1016/j.adro.2017.01.006. eCollection Apr-Jun.

70. Fitzpatrick JJ, Ryan MA, Bruzzi JF. Diagnostic accuracy of diffusion-weighted imaging-magnetic resonance imaging compared to positron emission tomography/computed tomography in evaluating and assessing pathological response to treatment in adult patients with lymphoma: A systematic review. Journal of medical imaging and radiation oncology 2018.

71. Friedberg JW, Byrtek M, Link BK, et al. Effectiveness of first-line management strategies for stage I follicular lymphoma: analysis of the National LymphoCare Study. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2012;30:3368-75.

72. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Med Decis Making 1991;11:88-94.

73. Fujiwara H, Maeda Y, Nawa Y, et al. The utility of positron emission tomography/computed tomography in the staging of extranodal natural killer/T-cell lymphoma. European journal of haematology 2011;87:123-9.

74. Furth C, Amthauer H, Hautzel H, et al. Evaluation of interim PET response criteria in paediatric Hodgkin's lymphoma--results for dedicated assessment criteria in a blinded dual-centre read. Annals of oncology : official journal of the European Society for Medical Oncology 2011;22:1198-203.

75. Furth C, Erdrich AS, Steffen IG, et al. Interim PET response criteria in paediatric non-Hodgkin's lymphoma. Results from a retrospective multicenter reading. Nuklearmedizin Nuclear medicine 2013;52:148-56.

76. Gallamini A, Barrington SF, Biggi A, et al. The predictive role of interim positron emission tomography for Hodgkin lymphoma treatment outcome is confirmed using the interpretation criteria of the Deauville five-point scale. Haematologica 2014;99:1107-13.

77. Gallamini A, Tarella C, Viviani S, et al. Early Chemotherapy Intensification With Escalated BEACOPP in Patients With Advanced-Stage Hodgkin Lymphoma With a Positive Interim Positron Emission Tomography/Computed Tomography Scan After Two ABVD Cycles: Long-Term Results of the GITIL/FIL HD 0607 Trial. J Clin Oncol 2018;36:454-62. doi: 10.1200/JCO.2017.75.543. Epub 8 Jan 23.

78. Ganesan P, Rajendranath R, Kannan K, et al. Phase II study of interim PET-CT-guided response-adapted therapy in advanced Hodgkin's lymphoma. Ann Oncol 2015;26:1170-4. doi: 10.093/annonc/mdv077. Epub 2015 Feb 20.

79. Gisselbrecht C, Glass B, Mounier N, et al. Salvage regimens with autologous transplantation for relapsed large B-cell lymphoma in the rituximab era. Journal of Clinical Oncology 2010;28:4184.

80. Gisselbrecht C, Schmitz N, Mounier N, et al. Rituximab maintenance therapy after autologous stem-cell transplantation in patients with relapsed CD20+ diffuse large B-cell lymphoma: final analysis of the collaborative trial in relapsed aggressive lymphoma. Journal of Clinical Oncology 2012;30:4462.

81. Goske MJ, Frush DP, Brink JA, Kaste SC, Butler PF, Pandharipande PV. Curbing potential radiation-induced cancer risks in oncologic imaging: perspectives from the 'image gently' and 'image wisely' campaigns. Oncology (Williston Park) 2014;28:232-8, 43.

82. Grellier JF, Vercellino L, Leblanc T, et al. Performance of FDG PET/CT at initial diagnosis in a rare lymphoma: nodular lymphocyte-predominant Hodgkin lymphoma. European journal of nuclear medicine and molecular imaging 2014;41:2023-30.

83. Guttikonda R, Herts BR, Dong F, Baker ME, Fenner KB, Pohlman B. Estimated radiation exposure and cancer risk from CT and PET/CT scans in patients with lymphoma. European journal of radiology 2014;83:1011-5.

84. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011;64:383-94.

85. Guyatt GH, Oxman AD, Schunemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. J Clin Epidemiol 2011;64:380-2.

86. Ha CS, Hodgson DC, Advani R, et al. ACR appropriateness criteria follow-up of Hodgkin lymphoma. Journal of the American College of Radiology : JACR 2014;11:1026-33.e3.

87. Hagberg H, Gisselbrecht C. Randomised phase III study of R-ICE versus R-DHAP in relapsed patients with CD20 diffuse large B-cell lymphoma (DLBCL) followed by high-dose therapy and a second randomisation to maintenance treatment with rituximab or not: an update of the CORAL study. Annals of oncology 2006;17:iv31-iv2.

88. Han EJ, O JH, Yoon H, et al. FDG PET/CT response in diffuse large B-cell lymphoma: Reader variability and association with clinical outcome. Medicine (Baltimore) 2016;95:e4983. doi: 10.1097/MD.0000000000004983.

89. Higgins J. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.02011.

90. Hong J, Kim JH, Lee KH, et al. Symptom-oriented clinical detection versus routine imaging as a monitoring policy of relapse in patients with diffuse large B-cell lymphoma. Leukemia & lymphoma 2014;55:2312-8.

91. Hoppe BS, Hodgson DC, Advani R, et al. ACR Appropriateness Criteria® Localized Nodal Indolent Lymphoma. Oncology 2013;27:786-.

92. Hoppe RT, Advani RH, Ai WZ, et al. NCCN Clinical Practice Guidelines in Oncology Hodgkin Lymphoma Version 3.2018, Ft. Washington, PA: NCCN; 2018 04/16/18.

93. Huntington SF, Svoboda J, Doshi JA. Cost-effectiveness analysis of routine surveillance imaging of patients with diffuse large B-cell lymphoma in first remission. J Clin Oncol 2015;33:1467-74. doi: 10.200/JCO.2014.58.5729. Epub 2015 Mar 30.

94. Johnson PW. Response-adapted frontline therapy for Hodgkin lymphoma: are we there yet? Hematology Am Soc Hematol Educ Program 2016;2016:316-22. doi: 10.1182/asheducation-2016.1.316.

95. Kedmi M, Apel A, Davidson T, et al. High-Risk, Advanced-Stage Hodgkin Lymphoma: The Impact of Combined Escalated BEACOPP and ABVD Treatment in Patients Who Rapidly Achieve Metabolic Complete Remission on Interim FDG-PET/CT Scan. Acta haematologica 2016;135:156-61.

96. Kinahan PE, Fletcher JW. Positron emission tomography-computed tomography standardized uptake values in clinical practice and assessing response to therapy. Semin Ultrasound CT MR 2010;31:496-505.

97. Kirby AM, Mikhaeel NG. The role of FDG PET in the management of lymphoma: what is the evidence base? Nuclear medicine communications 2007;28:335-54.

98. Kluge R, Chavdarova L, Hoffmann M, et al. Inter-Reader Reliability of Early FDG-PET/CT Response Assessment Using the Deauville Scale after 2 Cycles of Intensive Chemotherapy (OEPA) in Hodgkin's Lymphoma. PLoS One 2016;11:e0149072. doi: 10.1371/journal.pone.. eCollection 2016.

99. Kostakoglu L, Gallamini A. Interim 18F-FDG PET in Hodgkin lymphoma: would PET-adapted clinical trials lead to a paradigm shift? Journal of nuclear medicine : official publication, Society of Nuclear Medicine 2013;54:1082-93.

100. Kwee TC, Kwee RM, Nievelstein RA. Imaging in staging of malignant lymphoma: a systematic review. Blood 2008;111:504-16.

101. Lamy T, Damaj G, Soubeyran P, et al. R-CHOP 14 with or without radiotherapy in nonbulky limited-stage diffuse large B-cell lymphoma. Blood 2018;131:174-81.

102. Lassmann M, Treves ST, Group ESPDHW. Paediatric radiopharmaceutical administration: harmonization of the 2007 EANM paediatric dosage card (version 1.5. 2008) and the 2010 North American consensus guidelines. European journal of nuclear medicine and molecular imaging 2014;41:1036-41.

103. Liao C-C, Qin Y-Y, Tan X-H, et al. Predictive value of interim PET/CT visual interpretation in the prognosis of patients with aggressive non-Hodgkin's lymphoma. OncoTargets and therapy 2017;10:5727.

104. Liedtke M, Hamlin P, Moskowitz C, Zelenetz A. Surveillance imaging during remission identifies a group of patients with more favorable aggressive NHL at time of relapse: a retrospective analysis of a uniformly-treated patient population. Annals of oncology 2006;17:909-13.

105. Matchar DB. Introduction to the Methods Guide for Medical Test Reviews. In: Chang SM and Matchar DB, eds. Methods guide for medical test reviews. Rockville, MD: Agency for Healthcare Research and Quality; June 2012. p. 1.1-11. AHRQ Publication No. 12-EHC017. Available at www.effectivehealthcare.ahrq.gov/medtestsguide.cfm.

106. McKay P, Fielding P, Gallop-Evans E, et al. Guidelines for the investigation and management of nodular lymphocyte predominant Hodgkin lymphoma. Br J Haematol 2016;172:32-43.

107. Metser U, Dudebout J, Baetz T, et al. [(18) F]-FDG PET/CT in the staging and management of indolent lymphoma: A prospective multicenter PET registry study. Cancer 2017;123:2860-6. doi: 10.1002/cncr.30672. Epub 2017 Mar 13.

108. Moghbel MC, Mittra E, Gallamini A, et al. Response Assessment Criteria and Their Applications in Lymphoma: Part 2. Journal of nuclear medicine : official publication, Society of Nuclear Medicine 2017;58:13-22.

109. Moskowitz AJ, Schoder H, Gavane S, et al. Prognostic significance of baseline metabolic tumor volume in relapsed and refractory Hodgkin lymphoma. Blood 2017;130:2196-203. doi: 10.1182/blood-2017-06-788877. Epub 2017 Sep 5.

110. Moskowitz AJ, Schoder H, Yahalom J, et al. PET-adapted sequential salvage therapy with brentuximab vedotin followed by augmented ifosamide, carboplatin, and etoposide for patients with relapsed and refractory Hodgkin's lymphoma: a non-randomised, open-label, single-centre, phase 2 study. Lancet Oncol 2015;16:284-92. doi: 10.1016/S470-2045(15)70013-6. Epub 2015 Feb 13.

111. Moskowitz CH, Matasar MJ, Zelenetz AD, et al. Normalization of pre-ASCT, FDG-PET imaging with second-line, non-cross-resistant, chemotherapy programs improves event-free survival in patients with Hodgkin lymphoma. Blood 2012;119:1665-70.

112. Cancer Stat Facts: Hodgkin Lymphoma. (Accessed Aug., 2018, at https://seer.cancer.gov/statfacts/html/hodg.html.)

113. Ofman JJ, Sullivan SD, Neumann PJ, et al. Examining the value and quality of health economic analyses: implications of utilizing the QHES. J Manag Care Pharm 2003;9:53-61.

114. Pakos EE, Fotopoulos AD, Ioannidis JP. 18F-FDG PET for evaluation of bone marrow infiltration in staging of lymphoma: a meta-analysis. Journal of nuclear medicine : official publication, Society of Nuclear Medicine 2005;46:958-63.

115. Papajik T, Myslivecek M, Skopalova M, et al. Determining the extent and stage of disease in patients with newly diagnosed non-Hodgkin's lymphoma using 18F-FDG-PET/CT. Neoplasma 2011;58:291-7.

116. Pardal E, Coronado M, Martin A, et al. Intensification treatment based on early FDG-PET in patients with high-risk diffuse large B-cell lymphoma: a phase II GELTAMO trial. Br J Haematol 2014;167:327-36.

117. Patel K, Hadar N, Lee J, Siegel BA, Hillner BE, Lau J. The lack of evidence for PET or PET/CT surveillance of patients with treated lymphoma, colorectal cancer, and head and neck cancer: a systematic review. Journal of nuclear medicine : official publication, Society of Nuclear Medicine 2013;54:1518-27.

118. Patel V, Buckstein M, Perini R, Hill-Kayser C, Svoboda J, Plastaras JP. Computed tomography and positron emission tomography/computed tomography surveillance after combined modality

treatment of supradiaphragmatic Hodgkin lymphoma: a clinical and economic perspective. Leukemia & lymphoma 2013;54:2168-76.

119. Paulino AC, Margolin J, Dreyer Z, Teh BS, Chiang S. Impact of PET-CT on involved field radiotherapy design for pediatric Hodgkin lymphoma. Pediatric blood & cancer 2012;58:860-4.
120. Picardi M, Pugliese N, Cirillo M, et al. Advanced-stage Hodgkin lymphoma: US/chest radiography for detection of relapse in patients in first complete remission--a randomized trial of routine surveillance imaging procedures. Radiology 2014;272:262-74.
121. Pingali SR, Jewell SW, Havlat L, et al. Limited utility of routine surveillance imaging for classical Hodgkin lymphoma patients in first complete remission. Cancer 2014;120:2122-9.
122. Plastaras PJ, Advani R, Ballas LK, et al. ACR appropriateness Criteria® hodgkin lymphoma — stage III and IV. American journal of clinical oncology 2016.
123. Pommier P, Dussart S, Girinsky T, et al. Impact of 18F-fluoro-2-deoxyglucose positron emission tomography on treatment strategy and radiotherapy planning for stage I-II Hodgkin disease: a prospective multicenter study. International journal of radiation oncology, biology, physics 2011;79:823-8.
124. Press OW, Li H, Schoder H, et al. US Intergroup Trial of Response-Adapted Therapy for Stage III to IV Hodgkin Lymphoma Using Early Interim Fluorodeoxyglucose-Positron Emission Tomography Imaging: Southwest Oncology Group S0816. J Clin Oncol 2016;34:2020-7. doi: 10.1200/JCO.2015.63.1119. Epub 2016 Apr 11.
125. Pyo J, Kim KW, Jacene HA, Sakellis CG, Brown JR, Van den Abbeele AD. End-therapy positron emission tomography for treatment response assessment in follicular lymphoma: a systematic review and meta-analysis. Clinical Cancer Research 2013.
126. Radford J, Illidge T, Counsell N, et al. Results of a trial of PET-directed therapy for early-stage Hodgkin's lymphoma. N Engl J Med 2015;372:1598-607. doi: 10.056/NEJMoa1408648.
127. Radiation Dose in X-Ray and CT Exams. 2018. (Accessed August 20, 2018, at https://www.radiologyinfo.org/en/info.cfm?pg=safety-xray.)
128. Raemaekers JM, André MP, Federico M, et al. Omitting radiotherapy in early positron emission tomography-negative stage I/II Hodgkin lymphoma is associated with an increased risk of early relapse: Clinical results of the preplanned interim analysis of the randomized EORTC/LYSA/FIL H10 trial. Journal of clinical oncology 2014;32:1188-94.
129. Rashidi A, Ebadi M, Cashen AF. Allogeneic hematopoietic stem cell transplantation in Hodgkin lymphoma: a systematic review and meta-analysis. Bone marrow transplantation 2016;51:521.
130. Rathore N, Eissa HM, Margolin JF, et al. Pediatric Hodgkin lymphoma: are we over-scanning our patients? Pediatric hematology and oncology 2012;29:415-23.
131. Roberts KB, Younes A, Hodgson DC, et al. ACR Appropriateness Criteria® Hodgkin Lymphoma—Unfavorable Clinical Stage I and II. American journal of clinical oncology 2016;39:384-95.
132. Robertson VL, Anderson CS, Keller FG, et al. Role of FDG-PET in the definition of involved-field radiation therapy and management for pediatric Hodgkin's lymphoma. International journal of radiation oncology, biology, physics 2011;80:324-32.
133. Romano A, Parrinello NL, Vetro C, et al. Prognostic meaning of neutrophil to lymphocyte ratio (NLR) and lymphocyte to monocyte ration (LMR) in newly diagnosed Hodgkin lymphoma patients treated

upfront with a PET-2 based strategy. Ann Hematol 2018;97:1009-18. doi: 10.7/s00277-018-3276-y. Epub 2018 Feb 14.

134. Rovira J, Valera A, Colomo L, et al. Prognosis of patients with diffuse large B cell lymphoma not reaching complete response or relapsing after frontline chemotherapy or immunochemotherapy. Annals of hematology 2015;94:803-12.

135. Sawan P, Rebeiz K, Schoder H, et al. Specialized second-opinion radiology review of PET/CT examinations for patients with diffuse large B-cell lymphoma impacts patient care and management. Medicine (Baltimore) 2017;96:e9411. doi: 10.1097/MD.0000000000009411.

136. Sickinger MT, von Tresckow B, Kobe C, Borchmann P, Engert A, Skoetz N. PET-adapted omission of radiotherapy in early stage Hodgkin lymphoma-a systematic review and meta-analysis. Crit Rev Oncol Hematol 2016;101:86-92.:10.1016/j.critrevonc.2016.03.005. Epub Mar 6.

137. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA: a cancer journal for clinicians 2018;68:7-30.

138. Sohn BS, Yoon DH, Kim KP, et al. The role of (1)(8)F-fluorodeoxyglucose positron emission tomography at response assessment after autologous stem cell transplantation in T-cell non-Hodgkin's lymphoma patients. Annals of hematology 2013;92:1369-77.

139. Stewart DA, Kloiber R, Owen C, et al. Results of a prospective phase II trial evaluating interim positron emission tomography-guided high dose therapy for poor prognosis diffuse large B-cell lymphoma. Leukemia & lymphoma 2014;55:2064-70.

140. Straus DJ, Jung S-H, Pitcher B, et al. CALBG 50604: risk-adapted treatment of non-bulky early stage Hodgkin lymphoma based on interim PET. Blood 2018:blood-2018-01-827246.

141. Sun N, Zhao J, Qiao W, Wang T. Predictive value of interim PET/CT in DLBCL treated with R-CHOP: meta-analysis. BioMed research international 2015;2015.

142. Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. Blood 2016;127:2375-90.

143. Swinnen LJ, Li H, Quon A, et al. Response-adapted therapy for aggressive non-Hodgkin's lymphomas based on early [18F] FDG-PET scanning: ECOG-ACRIN Cancer Research Group study (E3404). Br J Haematol 2015;170:56-65. doi: 10.1111/bjh.13389. Epub 2015 Mar 30.

144. Taghipour M, Marcus C, Jones S, Sarangi R, Trahan TJ, Subramaniam RM. The value of fourth and subsequent post-treatment 18F-FDG PET/CT scans in the management of patients with non-Hodgkin's lymphoma. Nuclear medicine communications 2016b;37:699-704.

145. Taghipour M, Marcus C, Nunna P, Subramaniam RM. Follow-up FDG PET/CT in Patients With Non-Hodgkin Lymphoma: Value to Clinical Assessment and Patient Management. Clin Nucl Med 2016a;41:e93-7.

146. Terasawa T, Dahabreh IJ, Nihashi T. Fluorine-18-fluorodeoxyglucose positron emission tomography in response assessment before high-dose chemotherapy for lymphoma: a systematic review and meta-analysis. The oncologist 2010;15:750-9.

147. Terasawa T, Lau J, Bardet S, et al. Fluorine-18-fluorodeoxyglucose positron emission tomography for interim response assessment of advanced-stage Hodgkin's lymphoma and diffuse large B-cell lymphoma: a systematic review. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2009;27:1906-14.

148. Terasawa T, Nihashi T, Hotta T, Nagai H. 18F-FDG PET for posttherapy assessment of Hodgkin's disease and aggressive Non-Hodgkin's lymphoma: a systematic review. Journal of nuclear medicine 2008;49:13.

149. Terezakis SA, Metzger ML, Hodgson DC, et al. ACR appropriateness Criteria® pediatric Hodgkin lymphoma. Pediatric blood & cancer 2014;61:1305-12.

150. Terezakis SA, Schoder H, Kowalski A, et al. The utility and limitations of (18)F-fluorodeoxyglucose positron emission tomography with computed tomography in patients with primary mediastinal B-cell lymphoma: single institution experience and literature review. Int J Radiat Oncol Biol Phys 2014;89:376-83. doi: 10.1016/j.ijrobp.2014.02.006. Epub Apr 11.

151. Tilly H, Gomes da Silva M, Vitolo U, et al. Diffuse large B-cell lymphoma (DLBCL): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of oncology : official journal of the European Society for Medical Oncology 2015;26 Suppl 5:v116-25.

152. Toriihara A, Nakadate M, Nakamura S, Kubota K, Tateishi U. Seventy FDG-PET/CT cases in which nuclear medicine physicians suspected lymphoma: How reliable are we? Asia Oceania Journal of Nuclear Medicine and Biology 2017;5:95.

153. 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. 2015.

154. van Hamersvelt HP, Kwee TC, Fijnheer R, Beek FJ, de Klerk JM, Nievelstein RA. Can full-dose contrast-enhanced CT be omitted from an FDG-PET/CT staging examination in newly diagnosed FDG-avid lymphoma? Journal of computer assisted tomography 2014;38:620-5.

155. Vermoolen MA, Kersten MJ, Fijnheer R, van Leeuwen MS, Kwee TC, Nievelstein RA. Magnetic resonance imaging of malignant lymphoma. Expert review of hematology 2011;4:161-71.

156. Wang C, Li P, Wu S, et al. The role of fluorine-18 fluorodeoxyglucose PET in prognosis evaluation for stem cell transplantation of lymphoma: a systematic review and meta-analysis. Nuclear medicine communications 2016;37:338-47.

157. Ward E, DeSantis C, Robbins A, Kohler B, Jemal A. Childhood and adolescent cancer statistics, 2014. CA: a cancer journal for clinicians 2014;64:83-103.

158. West S, King V, Carey TS, et.al. Systems to Rate the Strength of Scientific Evidence. Evidence Report/Technology Assessment No. 47 (Prepared by the Research Triangle Institute-University of North Carolina Evidence-based Practice Center, Contract No. 290-97-0011). Agency for Healthcare Research and Quality, Rockville, MD; 2002.

159. Winkfield KM, Advani RH, Ballas LK, et al. ACR appropriateness criteria® recurrent Hodgkin lymphoma. ONCOLOGY (United States) 2016;30.

160. Wu LM, Chen FY, Jiang XX, Gu HY, Yin Y, Xu JR. 18F-FDG PET, combined FDG-PET/CT and MRI for evaluation of bone marrow infiltration in staging of lymphoma: a systematic review and meta-analysis. European journal of radiology 2012;81:303-11.

161. Yang M, Sun J, Bai HX, et al. Diagnostic accuracy of SPECT, PET, and MRS for primary central nervous system lymphoma in HIV patients: A systematic review and meta-analysis. Medicine (Baltimore) 2017;96:e6676. doi: 10.1097/MD.0000000000006676.

162. Younes A, Hilden P, Coiffier B, et al. International Working Group consensus response evaluation criteria in lymphoma (RECIL 2017). Annals of oncology : official journal of the European Society for Medical Oncology 2017;28:1436-47.

163. Zelenetz AD, Gordon LI, Abramson JS, et al. NCCN Clinical Practice Guidelines in Oncology B-Cell Lymphomas Version 4.2018, . Clinical Practice Guideline. Ft. Washington, PA2018 05/15/2018.

164. Zhou X, Lu K, Geng L, Li X, Jiang Y, Wang X. Utility of PET/CT in the diagnosis and staging of extranodal natural killer/T-cell lymphoma: a systematic review and meta-analysis. Medicine (Baltimore) 2014;93:e258. doi: 10.1097/MD.0000000000000258.

165. Zhu Y, Lu J, Wei X, Song S, Huang G. The predictive value of interim and final [18F] fluorodeoxyglucose positron emission tomography after rituximab-chemotherapy in the treatment of non-Hodgkin's lymphoma: a meta-analysis. BioMed research international 2013;2013.

166. Ziakas PD, Poulou LS, Voulgarelis M, Thanos L. The Gordian knot of interim 18-fluorodeoxyglucose positron emission tomography for Hodgkin lymphoma: a meta-analysis and commentary on published studies. Leukemia & lymphoma 2012;53:2166-74.

167. Zijlstra JM, Lindauer-van der Werf G, Hoekstra OS, Hooft L, Riphagen, II, Huijgens PC. 18F-fluoro-deoxyglucose positron emission tomography for post-treatment evaluation of malignant lymphoma: a systematic review. Haematologica 2006;91:522-9.

168. Zinzani PL, Broccoli A, Casadei B, et al. The role of rituximab and positron emission tomography in the treatment of primary mediastinal large B-cell lymphoma: experience on 74 patients. Hematological oncology 2015;33:145-50.

169. Zinzani PL, Broccoli A, Gioia DM, et al. Interim Positron Emission Tomography Response-Adapted Therapy in Advanced-Stage Hodgkin Lymphoma: Final Results of the Phase II Part of the HD0801 Study. J Clin Oncol 2016;34:1376-85. doi: 10.200/JCO.2015.63.0699. Epub 2016 Feb 16.

170. Zou Y, Tong J, Leng H, Jiang J, Pan M, Chen Z. Diagnostic value of using 18F-FDG PET and PET/CT in immunocompetent patients with primary central nervous system lymphoma: A systematic review and meta-analysis. Oncotarget 2017;8:41518-28. doi: 10.18632/oncotarget.7456.