Ontario Health
Cancer Care Ontario

Evidence Summary 6-25

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

First-line treatment of advanced-stage Hodgkin lymphoma

L. Mozessohn, F.G. Baldassarre, A. Prica, M.C. Cheung, C. Faught, D. Rodin, A. Singnurkar, A. Suleman, L. Hicks

Report Date: October 15, 2024

For information about this document, please contact Drs. Lee Mozessohn and Lisa Hicks, through the PEBC via:
Phone: 905-527-4322 ext. 42822 Fax: 905 526-6775 E-mail: ccopgi@mcmaster.ca

For information about the PEBC and the most current version of all reports, please visit the OH (CCO) website at https://www.cancercareontario.ca/en/guidelines-advice or contact the PEBC office at:
Phone: 905-527-4322 ext. 42822 Fax: 905 526-6775 E-mail: ccopgi@mcmaster.ca

PEBC Report Citation (Vancouver Style): Mozessohn L, Baldassarre FG, Prica A, Cheung MC, Rodin D, Singnurkar A, Suleman A, Hicks L. First-line treatment of advanced-stage Hodgkin lymphoma. Toronto (ON): Ontario Health (Cancer Care Ontario); 2024 October 15. Program in Evidence-Based Care Evidence Summary No. 6-25, available on the OH (COO) website: <insert URL when known>.

Table of Contents

First-line treatment of advanced-stage Hodgkin lymphoma

Evidence Summary

THE PROGRAM IN EVIDENCE-BASED CARE

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

The PEBC is a provincial initiative of OH (CCO) supported by the Ontario Ministry of Health (OMH). All work produced by the PEBC is editorially independent from the OMH.

INTRODUCTION

Hodgkin lymphoma (HL) is a malignancy of germinal centre B-cells with an incidence that peaks in the second and seventh decades of life [1]. An estimated 1050 Canadians have been diagnosed with HL in 2022, with an age-standardized incidence rate of 2.6 per 100,000 population [2]. Ninety-five percent of these cases are classical HL, while the remaining 5% are of nodular lymphocyte predominant type [3]. Approximately 50-60% of patients are diagnosed with advanced-stage HL (defined as Ann Arbor stage III and IV) [4,5].

Common treatment approaches include chemotherapy such as doxorubicin, bleomycin, vinblastine, and dacarbazine (ABVD), or escalated bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine and prednisolone (escBEACOPP) [6-8]. Recently, novel therapies such as brentuximab vedotin, an anti-CD30 monoclonal antibody conjugated to a microtubule disrupting agent, and immune checkpoint inhibitors, such as nivolumab and pembrolizumab, have been shown to be highly effective in patients with HL [9,10].

More aggressive regimens are associated with increased short-term and long-term toxicity, including the risk of secondary malignancies [11-13]. These toxicities are an important consideration, as elderly patients are at increased risk of infectious complications, cardiopulmonary toxicities, and debilitating neuropathy with multiagent chemotherapy [14,15]. Consequently, recent trials have aimed to use 2-(^{18}F)-fluoro-2-deoxy-D-glucose (FDG) positron emission tomography (PET) combined with computed tomography (CT) to guide the need for escalation or de-escalation of therapy based on response [16-19]. Additionally, radiotherapy has been associated with late morbidity, but with improved technology these toxicities may be mitigated, with improved survival [20]. In newer PET-adapted strategies using highly effective treatment regimens with or without novel therapies, the benefit of consolidation radiotherapy is unclear [21].

The Working Group of the Hematology Disease Site Group initially developed this systematic review to inform recommendations as part of a clinical practice guideline on the first-line treatment of advanced-stage HL. However, it was realized in October 2023 that at least two practice-changing studies, the HD21 [22] and the SWOG S1826 [23] will not be released as full-text, peer-reviewed publications soon enough to be included in this work by the intended release in October 2024. Therefore, the systematic review was summarized in this evidence summary, and recommendations were not issued.

This systematic review has been registered on the PROSPERO website (International prospective register of systematic reviews) with the following registration number CRD 4202121919 (https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021261919)

Based on the objectives of this work, the members of the Working Group derived the research questions outlined below.

OBJECTIVES

- To decide what is the preferred first-line treatment strategy for patients with advanced-stage HL among ABVD, brentuximab vedotin plus doxorubicin, vinblastine, and dacarbazine (B-AVD), A-AVD, eBEACOPP, and other treatments with or without PET response-adapted strategies?
- To determine the role of radiation as part of the first-line time treatment strategy for these patients

RESEARCH QUESTIONS

In patients with advanced-stage HL:

1. For patients younger than 60 years of age, what is the ideal treatment strategy, among A-AVD, ABVD, ABVD-PET adapted, eBEACOPP, and eBEACOPP–PET adapted, as part of the first-line therapy, to improve patient outcomes, and how does it affect adverse events (AE)?
2. For patients 60 years of age and older what is the ideal treatment strategy, among ABVD, B-AVD, **brentuximab alone or in combination,** PVAG (prednisone, vinblastine, doxorubicin, gemcitabine), CHOP (cyclophosphamide, doxorubicin, vincristine, prednisolone), CHLVPP (Chlorambucil-VinBLAStine-Procarbazine-Prednisone), **or other agents,** as part of the first-line therapy to improve patient outcomes, and how does it affect AE?
3. Does consolidation radiotherapy after first-line chemotherapy improve outcomes such as overall survival (OS), progression-free survival (PFS), recurrence, adverse events (AE), and quality of life (QoL) in adult patients with advanced-stage HL?

TARGET POPULATION

Adult patients (≥18 years of age) with advanced-stage HL that require upfront (first-line) treatment. Question 1 focuses on patients aged ≥18 and up to 60 years of age; Question 2 focuses on patients aged 60 and over, and Question 3 focuses on all adult patients.

INTENDED PURPOSE

Initially intended to be the evidentiary base for a guideline, this evidence summary reports evidence available for the first-line treatment of patients with advanced-stage HL. The Working Group decided not to proceed to issue recommendations for this population because two practice-changing studies, the HD21 [24] and the SWOG S1826 [23], are not available in full text, peer-reviewed format at the time of the latest update of this report. See Amendments to the Project Plan in Appendix 8.

INTENDED USERS

Clinicians treating patients with advanced-stage HL that require first-line treatment, including hematologists, radiation oncologists, radiologists, medical oncologists, nurse practitioners and hematology pharmacists.

METHODS

This evidence summary was developed by a Working Group consisting of hematologists, radiation oncologists, nuclear medicine radiologists, and a health research methodologist at the request of the Hematology Disease Site Group and the PET Steering Committee.

The Working Group was responsible for reviewing the identified evidence and drafting the summary. Conflict of interest declarations for all authors are summarized in Appendix 1, and were managed in accordance with the *PEBC Conflict of Interest Policy*.

This evidence review was conducted in two planned stages, including a search for guidelines/systematic reviews followed by a search for primary literature. These stages are described in subsequent sections.

Search for Guidelines

Evidence-based guidelines with systematic reviews that addressed at least one research question were included. Guidelines that were older than three years (i.e., published before 2018) and guidelines based on consensus or expert opinion were excluded.

The following sources were searched for guidelines on April 7, 2021 with the search term(s) "advanced-stage Hodgkin Lymphoma": Alberta Health Services, American Society of Clinical Oncology, European Society of Medical Oncology, National Institute for Health and Care Excellence Evidence Search, Canadian Medical Association Journal Infobase, Scottish Intercollegiate Guidelines Network, National Health and Medical Research Council – Australia Clinical Practice Guidelines Portal, and Cancer Council Australia – Cancer Guidelines Wiki, MEDLINE and EMBASE. After deduplication, 72 publications were identified. After title and abstract screening, four publications were selected for full-text review [25-28]. At full-text review three of these guidelines [25,26,28] were excluded because they did not include the population of interest, or they did not report on the outcomes of interest. The Working Group decided to use the Bröckelmann et al. [27] guideline as a source of evidence. In Appendix 2, Table 1 shows the websites of the guideline developers organizations searched and the results of the searches, and Table 2 shows the search strategies used in the MEDLINE and EMBASE databases.

Search for Systematic Reviews

We searched MEDLINE, EMBASE, the Cochrane Database of Systematic Reviews, the Agency for Health Research and Quality (AHRQ), the Canadian Agency for Drugs and Technologies in Health (CADTH), the National Institute for Health Research Health Technology Assessment (NIHR HTA [UK]) for systematic reviews published from 2017 to 2021. More details about the databases searched and the terms used are reported in Appendix 2, Tables 3 and 4.

Systematic reviews were included if they met the following criteria:

- The review addressed at least one research question with similar inclusion/exclusion criteria, and
- The review had a moderate/high overall rating as assessed with the AMSTAR 2 tool [29].
- The review was less than five years old (i.e., it was published, and search cut-off date was 2016 or later).

We included non-systematic reviews of pooled analyses with patient-level data.

Risk of bias per outcome for each included study was assessed with the AMSTAR 2 [29].

If more than one systematic review met the inclusion criteria, then one systematic review for each outcome per research question was selected by the methodologist (FGB) based on its age, quality, and the best match with our study selection criteria stated below.

Search for Primary Literature

For each outcome per comparison, within each research question, if no systematic review was included, then a search for primary literature was conducted. For any included systematic review, an updated search for primary literature was performed. If any included

systematic review was limited in scope, then a search for primary literature to address the limitation in scope was conducted.

Literature Search Strategy

We searched MEDLINE (Ovid), EMBASE (Ovid), the Cochrane library (CENTRAL), and our own files from 1998 to November 5, 2021; our own files comprised the articles that were provided by the authors and that were used to design the search strategies and to define the topic in the earliest phases of the review. We updated the search on January 12, 2024, using the same databases mentioned above and PubMed, to identify the most recent reports. We searched the Conference proceedings of the American Society of Clinical Oncology, the American Society of Hematology, the International Conference on Malignant Lymphoma, and the European Hematology Association from 2019 to April 2022. We combined keywords for HL with terms aimed at identifying randomized controlled trials (RCTs). The complete search strategies for primary studies are reported in Appendix 2, Table 5.

Since no RCTs that met the inclusion criteria were identified for Question 2, on September 26, 2022, we conducted a search for comparative trials (both RCTs and observational) that controlled for confounding and had a minimum sample of 30 patients. This number was chosen because sample sizes ≥30 are considered sufficient for the central limit theorem to hold. The central limit theorem states that the distribution of sample means approximates a normal distribution as the sample size gets larger, regardless of the population's distribution [30]. This search was updated on January 30, 2024.

Since no comparative studies met the inclusion criteria for Question 2, the Working Group decided to search for single-arm, phase 2 trials. These studies were considered important by the Working Group members, specifically for considerations related to treatments' AEs, and efficacy in the presence of significant comorbidity.

Since the search executed on September 22, 2022, was not limited by study design, and the studies that were not comparative were excluded at the title and abstract level, the methodologist, on October 19, 2023, searched the existing database for noncomparative trials of older patients with advanced-stage HL.

The methodologist retrieved the studies that met the inclusion criteria in the library and reviewed the full text.

Study Selection Criteria and Process

We included studies of adult patients with advanced-stage (i.e., stages II with adverse features, III and IV) HL treated with ABVD, eBEACOPP, A-AVD, and any other novel treatment, including PET-directed strategies; brentuximab alone or in combination, PVAG, CHOP, CHLVPP, and consolidation radiotherapy. Critical outcomes were OS, recurrence (including PFS and other measures), AEs (e.g., treatment-related death) and QoL. Pulmonary toxicity was an important outcome for Question 2. Selection criteria are described in detail in Appendix 4.

A review of the titles and abstracts was conducted by one reviewer (FGB). For studies that warranted full-text review, one reviewer (FGB) reviewed each study. If uncertainty existed, both at title and abstract and full-text screening, a second author (AP, MC) reviewed the studies, and agreement was reached through discussion.

The exclusion criteria used for the selection of single-arm trials are:

- 20% or more of the recruited patients did not meet the inclusion criteria for Question 2 (e.g., younger patients, patients with cancers other than classic HL, patients with early-stage HL).
- Sample size was <30 patients.
- Treatments included are now obsolete.

- Outcomes are not relevant (i.e., they were not important or critical outcomes for this question).
- Publications are conference abstracts reports.
- Abstract publications are non-RCTs, or interim analysis of ongoing RCTs. These are listed among the ongoing trials for this question.
- Case studies

Critical and Important Outcomes

On an initial meeting at the project plan stage, the patient representatives were asked by two PEBC methodologists which outcomes they considered critical and important for the patient populations of concern to Questions 1, 2 and 3. The clinician members of the Working Group agreed on the decisions that the patients made during this first brainstorming session.

On February 10, 2023, the Working Group met and discussed the outcomes again after considering the evidence included by the systematic review: Recurrence was viewed as a synonymous of disease control, which includes PFS, event-free survival (EFS), disease-free survival (DFS), failure-free survival (FFS), freedom from treatment failure (FFTF). The Working Group considered as most relevant the following AEs: treatment-related death, grade 3-4 infections, febrile neutropenia, pulmonary toxicity, cardiac toxicity, peripheral neuropathy; secondary malignancies, and impaired fertility as measured with female and male hormone levels. A survey was distributed to the Working Group members asking to re-assess and rate the importance of outcomes for decision making, for each research question, on a one to nine score system: one to three meaning limited or not importance for decision-making; four to six meaning important, but not critical for decision-making, and seven to nine meaning critical for decision-making [31]. See Appendix 6 for the results of this survey.

Table 1 shows the critical and important outcomes for each question.

Additionally, since this is an area of wide variability in patients' and clinicians' preferences [31], a MEDLINE search of the literature to identify reviews or surveys of patient preferences was also conducted. The key words: Hodgkin lymphoma, first-line treatment, and patient preferences were used. The studies by Khan et al. [32], and by Bröckelmann et al. [33] were identified. See Appendix 6 for the results of these studies. The Working Group discussed the results of the survey and of the data on patient preferences retrieved from the literature during the March 10, 2023, Working Group meeting. Considering that seven is the maximum number of outcomes recommended by the GRADE group [34], the outcomes were re-classified as follows:

Table 1. Critical and important outcomes

Question	Critical outcomes	Important outcomes	Outcomes that are not important
Question 1	OS Recurrence[a] Treatment-related death	Pulmonary toxicity SPM Infertility	Quality of life Other adverse events
Question 2	OS Recurrence[a] Treatment-related death Quality of life	Pulmonary toxicity	Infertility Other adverse events
Question 3	OS Recurrence[a] Treatment-related death	SPM	Other adverse events

[a]As measured by progression-free survival, and other measures such as event-free survival, disease-free survival, failure-free survival, and freedom from treatment failure.
OS=overall survival; SPM=second primary malignancies

Data Extraction and Assessment of Risk of Bias

The methodologist (FGB) extracted data from all primary studies into evidence tables. The other authors (A Suleman, MC, LM) reviewed the evidence tables for correctness. All extracted data and information were subsequently audited by an independent auditor (WJ). Ratios, including hazard ratios (HRs), were expressed with a ratio of <1.0 indicating that the group that received the experimental intervention has a lower risk of experiencing an event (e.g., death) over time compared with the control group.

During data extraction we realized that two studies, the SWOG S1826, and the HD21 [23,24] relevant for Question 1, and one study, the SWOG S1826 [23], relevant for Question 2, are now published only as interim analyses, and, therefore, cannot be included in this review at this time. Dr. Herrera, the corresponding author of the SWOG S1826 trial, was contacted and he confirmed upcoming publications for both the younger population and for a subgroup of 98 patients older than 60 years, but the date of publication of these studies is unknown at this time. As these studies are potentially practice changing, the Working Group decided to amend the project plan, and to present this work as an evidence summary and not to proceed to draft recommendations.

Assessment of Risk of Bias

Two methodologists (FGB and XY) assessed the risk of bias of included individual patient data (IPD) meta-analyses using the CheckMAP tool [35]. This tool consists of a series of 10 questions with a yes/no answer (see Table 2 in Appendix 5) as a guide in the assessment of the risk of bias. According to Tierney et al. [35], if an IPD meta-analysis is based on a systematic review, includes a high proportion of good-quality data, and uses appropriate analyses, it is most likely reliable.

For included RCTs, for each comparison, the methodologist (FGB) conducted a per-outcome, domain-based risk of bias assessment with the Cochrane Risk of Bias 2 [36] tool. Other authors (LM, MC, A Suleman, and AS) reviewed the risk of bias tables for accuracy, and disagreements were resolved by discussion. An independent auditor (WJ) audited the risk of bias assessments. The Cochrane ROB 2 tool comprises five domains: 1) Bias arising from the randomization process, 2) bias due to deviations from intended interventions, 3) bias due to missing outcome data, 4) bias in measurement of the outcome, and 5) bias in selection of the reported results. For each domain, a series of signalling questions with the answers: “yes, probably yes, no information, probably no, and no” determine the risk of bias (low risk, some concerns, and high risk). For each signalling question, text was added to show the evidence on which the judgment was based.

For observational comparative studies we planned to use the ROBINS-I tool [37].

For single-arm phase 2 trials that were included for Question 2, no risk of bias assessment was planned as these trials are at very high risk of bias because these studies did not have a comparator.

For subgroup analyses of RCTs, we followed the guidance provided by Sun et al. 2014 [38].

The certainty of the evidence, per outcome, for each comparison, considering risk of bias, inconsistency, indirectness, imprecision, and publication bias was assessed according to the GRADE system [39].

Synthesizing the Evidence

For time-to-event outcomes, when clinically and methodologically homogeneous results from two or more studies were available, a meta-analysis was conducted using Review Manager software provided by the Cochrane Collaboration, version 5.4.1. HRs, rather than the number of events at a specific time, were the preferred statistic for meta-analysis, and were used as

reported. If the HR and/or its standard error were not reported, they were derived from other information reported in the study if available, using the methods described by Parmar et al. [40]. The generic inverse variance model with random effects was used.

The chi-squared (X^2) test was used to test the null hypothesis of homogeneity, and a probability level less than or equal to 10% ($p \leq 0.10$) was considered indicative of statistical heterogeneity. If heterogeneity was detected, then the I^2 index was used to quantify the percentage of the variability in the effect estimates that was due to heterogeneity.

RESULTS

Search for Systematic Reviews

The searches resulted in 428 records: 48 articles were selected for full-text review after screening at the title and abstract level. After full text review we included nine reviews [6,11,41-47]. Six were systematic reviews [6,41-44,47]; two were IPD meta-analyses [11,46], and one was a network meta-analysis [45]. See flow chart in Appendix 3.

Two systematic reviews were excluded at quality assessment [44,47] because our confidence in the results was low or critically low as assessed with the AMSTAR 2 tool [29]. The network meta-analysis by Zhang et al. [45] and the IPD meta-analysis by Franklin et al. [46] were excluded at data extraction because the authors included studies of interventions that are currently considered of historical interest and are not used as standard therapies in HD (e.g., Stanford V [doxorubicin, vinblastine, mechlorethamine, vincristine, bleomycin, etoposide, and prednisone combined with radiation therapy], and M[C]OPP [mechlorethamine [cyclophosphamide], vincristine, procarbazine, and prednisone]). The IPD meta-analysis by Andre et al. [11] was considered unreliable after assessment with the CheckMAP tool [35], and it was excluded (Table 2, Appendix 5). The review by Aldin et al. [41] was excluded at data extraction because it included patients with both limited and advanced-stage and did not provide separate results. Finally, the review by Skoetz et al. [6] was excluded because it included patients with early-stage (stages 1 and 2) unfavourable as well as advanced-stage HL and did not provide separate results.

The systematic reviews by Amin et al. [43] and Amitai et al. [42] were included. These studies were of high quality and presented their results in a narrative manner. The review by Amin et al. [43] is recent, and it explored the fertility side effect of HL treatment in young men; its results will be used to inform this document for infertility in men for Question 1. The review by Amitai et al. 2018 [42] had a search cut-off in August 2017 and examined the effect of PET-adapted strategy, and it will be used as a source of evidence. Table 2 shows the general characteristics and the summary results of the two included reviews.

Table 2. General characteristics and summary results of included systematic reviews

Author, year, Country, Funding	Objectives Search cut-off; Design, Follow-up	Patient Population	Intervention vs. Comparison	Outcomes	Number and design of included studies	Summary results
Question 1: Patients <60 years of age						
Amin et al. 2020 [43] Country: UK Funding: UK MRC Centre for Transplantation at King's College London and the King's Medical Research Trust; The Urology Foundation, the Royal College of Surgeons of England, the Pelican Group, KMRT and the MRC Centre for Transplantation at King's College London	**Objectives:** To systematically review fertility side effect of HL treatment in young males and to identify potential strategies to preserve reproductive function. **Search cut-off**: January 2000 **Design:** systematic review with narrative synthesis Follow-up: up to 4 yrs	n=1344 pts Age: 13 to 51 yrs Stage: IIB, III and IV Recruitment: nr	ABVD vs. BEACOPP	Fertility* levels as measured by: • Sperm characteristics (sperm count, spermatogenesis, motility) • FSH, LH, testosterone • Inhibin B levels Secondary outcomes: Methods for fertility preservation	5 studies (4 prospective and one retrospective cohort trials)	Sperm characteristics ABVD: at 6 months 38% of pts had oligospermia, and 40% had azoospermia, 6-8 cycles lowered sperm count more than 2-4 cycles (p=0.05). Pts treated with ABVD+RT had significantly decreased sperm motility (p=0.001) and significant changes in sperm morphology (p=0.01) Inguinal-sparing radiotherapy maintained oligospermia at 12 months (p=0.01) ABVD: at 12-18 mos 50% of pts had recovered sperm characteristics At 24 mos: 57% of pts had recovered sperm characteristics BEACOPP: at 4 yrs: Pts who underwent 2-4 cycles of BEACOPP recovered sperm function. After 6-8 cycles of BEACOPP pts did not recover spermiogenesis. No differences in fertility rates between BEACOPP and eBEACOPP (p>0.999) Only 4% of dyspermic pts recovered spermatogenesis between 1.5-6.7 yrs post-therapy. Fewer cycles of both regimens increased the likelihood of sperm production recovery. Pts treated with BEACOPP were more likely to have oligospermia than those treated with ABVD (p values *nr*). Therefore, pts with advanced stage were less likely to have children born via natural methods post-treatment due to the more gonadotoxic treatment used for advanced-stage disease (p=0.04). Changes in sex hormones ABVD: at 6 and 12 months

Author, year, Country, Funding	Objectives Search cut-off; Design, Follow-up	Patient Population	Intervention vs. Comparison	Outcomes	Number and design of included studies	Summary results
						FSH was raised post-ABVD compared to healthy controls (p=0.008) before returning to normal LH and testosterone did not significantly change before and after treatment (p=0.203, and p=0.844 respectively) BEACOPP: 6–8 cycles of BEACOPP=lower Inhibin B/FSH ratios compared to healthy controls (impaired fertility) (p<0.001) Fertility preservation methods No data on this outcome were available
Amitai et al. 2018 [42] Country: Israel **Funding:** not reported	**Objectives:** To evaluate the effect of PET-adapted strategy on outcomes in advanced-stage HL **Search cut-off**: 1990 to Aug 2017 **Design:** systematic review with results presented in a narrative manner **Follow-up:** between 16 and 55 mos post treatment	n=6856 pts Age: *nr* Stage: IIb, III, and IV. One study included also IIa Recruitment: Studies published from 1999 to 2014	PET-adapted therapy Baseline regimens included ABVD or eBEACOPP One trial performed escalation of therapy if PET-2 was positive; one trial performed de-escalation if PET-2 was negative, and 2 trials performed both escalation and de-escalation. Escalation schemes included various regimens of BEACOPP; 2 of them included randomization with or without rituximab	PFS* OS	13 trials (10 fully published and 3 abs). Of these 4 were RCTs (used for the main analysis). In 3 trials PET-2 was interpreted by Deauville criteria, and in one by the IHP. Nine observational studies (used for secondary analyses): 7 phase 2 trials and 2 retrospective studies	*De-escalation when PET negative* The LYSA AHL 2011 study [17] that compared PET-2 adapted with non-adapted therapy, and that was still ongoing at the time of this publication was included. At a follow-up of 16.3 mos, the 2-yr PFS was similar in the adapted and non-adapted therapy (92% vs. 88%, p=0.79 – at the interim analysis). *Escalation when PET positive* The Italian GITIL/FIL HL 0607 trial [48] If PET-2 was positive, patients received four cycles of eBEACOPP plus four cycles of baseline BEACOPP (BB) and were randomly assigned to the addition of rituximab or not. PET-2 negative patients received additional four cycles of ABVD. No difference between groups with rituximab. Escalation and De-escalation according to PET results The RATHL [49] and HD18 trials [16,50,51]
Question 2: Patients ≥60 years of age						
No systematic review was included for this question						
Question 3: Consolidation radiotherapy						
No systematic review was included for this question						

*=Primary outcome
** "Radiotherapy" means in this study modern high-energy irradiation to the involved field or less, except for the comparison between involved-field and extended-field irradiation. 'Extended field' includes also the more extensive subtotal and total nodal irradiation categories (STNI and TNI).

$ “Chemotherapy” in this study means ABVD-like regimens similar to ABVD, except for the dose-intensified regimens.
a For analysis of OS, two time periods were defined because HR was not constant over time: <18 and ≥18 months

abs=abstracts; ABVD=doxorubicin, bleomycin, vinblastine, and dacarbazine; ASCT=autologous stem cell transplantation; BEACOPP=bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisolone; eBEACOPP=escalated BEACOPP; CR=complete remission; FFP=freedom from first progression; FFS=failure-free survival; FSH=follicle-stimulating hormone; HL=Hodgkin lymphoma; IHP=International Harmonization Project; IPI=International Prognostic Index; LH=luteinizing hormone; MCR=Medical Research Council; mos=months; *nr*=not reported; OS=overall survival; PET=positron emission tomography; PFS=progression-free survival; pts=patients; RCT=randomized controlled trial; RFS=relapse-free survival; RT=Radiotherapy; UK=United Kingdom; yrs=years

Search for Primary Literature

Literature Search Results

Historically, ABVD was created in 1973; it came to replace mechlorethamine, oncovin, procarbazine and prednisone (MOPP), the first successful therapy for HL, and it became the standard therapy. Looking for a better control, BEACOPP was developed in 1990 by the German Hodgkin Study Group (GHSG), and early on it was noted to be associated with increased toxicity. Subsequently, to counterbalance this toxicity while maintaining the best possible efficacy, PET-directed therapies and combinations that avoided bleomycin were tested.

The body of evidence included in this evidence summary includes all the above-mentioned chemotherapy combinations and treatment strategies, which, over time, were compared using various doses and schedules. The Working Group decided to group these comparisons in fewer broader categories for simplicity.

These broader categories include, for Question 1: 1) ABVD versus BEACOPP, 2) higher-intensity/dose versus lower-intensity/dose BEACOPP, 3) PET-adapted strategies 4) ABVD versus modified ABVD with brentuximab vedotin instead of bleomycin, 5) modified ABVD versus another modified ABVD, and 6) BEACOPP versus modified BEACOPP. For Question 2, the evidence-base includes subgroup analyses of studies already included in Question 1, and of single-arm phase 2 studies. For Question 3: 1) consolidation radiotherapy versus observation and radiotherapy versus chemotherapy are the comparisons included.

Table 3 presents a summary list of all studies included for all questions, grouped by the broader comparison categories. In this table within each broader category we present each study with the specific comparison(s) examined, the outcomes considered, and the study design.

Given the nature of the evidence base, in the following paragraphs, we will first present the results and certainty of the evidence of randomized trials included for Questions 1 and 3, as no RCT met the inclusion criteria for Question 2, then the results of observational studies included for Question 2.

Table 3. Unique primary studies selected for inclusion.

Study	Comparison(s) vs. Intervention(s)	Outcomes (critical or important)	Design
QUESTION 1: pts <60 yrs of age			
1. ABVD vs. BEACOPP			
EORTC 20012 Intergroup Trial, Carde, 2016 [52]	8×ABVD vs. $eBEACOPP_4$+$bBEACOPP_4$	OS *Recurrence*: PFS, EFS, DFS (only pts who reached CR – 83% of the sample) *AE*: TRD, SPM	RCT, superiority, single blinded (outcome assessors)
LYSA H34, Mounier, 2014 [53]	8×ABVD vs. 4×eBEACOPP+4×bBEACOPP	OS *Recurrence*: PFS, EFS *AE*: Death because of SPM, pulmonary toxicity	RCT, open label, parallel group, phase III
HD9, Diehl, 2003 [54]	8×COPP/ABVD no G-CSF vs. 8×bBEACOPP no G-CSF vs. 8×eBEACOPP G-CSF	OS *Recurrence*: PFS, FFTF AE: TRD, SPM, pulmonary toxicity	RCT open label trial
HD2000 Gruppo Italiano per lo Studio dei Linfomi Trial, Federico, 2009 [55], Merli, 2016 [56]	6×ABVD vs. 4×eBEACOPP+2bBEACOPP vs. 6×COPPEBVCAD-CEC	OS *Recurrence*: PFS, FFS, RFS (relapse-free survival) AE: TRD, SPM	RCT multicentre, open label
GITIL [57]	6×ABVD (if CR reached after 4 cycles) or 8×ABVD vs. 4×eBEACOPP + 4×bBEACOPP	OS *Recurrence*: EFS, FFFP AE: TRD, SPM	RCT multicentre, open label
2. Higher-intensity/dose vs. Lower-intensity/dose BEACOPP			
HD15, Engert, 2012 [58]	8×eBEACOPP vs. 6×eBEACOPP vs. $8\times BEACOPP_{14}$	OS *Recurrence*: PFS, FFTF, TTP AE: TRD, SPM	RCT, phase 3, open label, noninferiority trial
HD12, Borchmann, 2011 [59]	Chemotherapy comparison: 8×eBEACOPP ± RT vs. 4eBEACOPP + 4bBEACOPP± RT *NOTE: this study also appears in Question 3 as it reported a radiotherapy comparison*	OS *Recurrence*: PFS, FFTF AE: TRD, SPM	RCT 2××2 factorial, open label, noninferiority trial
HD9, Diehl, 2003 [54]	8×bBEACOPP no G-CSF vs. 8×eBEACOPP G-CSF	OS *Recurrence*: PFS, FFTF AE: see comparison 1	RCT open label, noninferiority trial*
3. PET-adapted strategies			
HD18, Borchmann, 2018 [16]	2×eBEACOPP then PET. Then: PET2– pts: 8×eBEACOPP (Standard) or 6×eBEACOPP (New Standard after amendment) vs. 4×eBEACOPP PET-2+ pts: 8×eBEACOPP or 6×eBEACOPP (Standard) vs. 8×Ritux-eBEACOPP	OS *Recurrence*: PFS AE: TRD, SPM	PET2 + pts: RCT, open label; superiority trial, stopped early for futility. PET2-negative pts: noninferiority trial.
AHL2011, Casasnovas, 2019 [17]	2×eBEACOPP, then PET2. Then: 4×eBEACOPP vs: If PET-2-negative: 2×ABVD If PET-2-positive: 2×eBEACOPP Then PET4: Consolidation If PET-4-negative: Standard treatment: 2×eBEACOPP	OS *Recurrence*: PFS AE: TRD, SPM, lung function	RCT, open label, multicentre, phase 3, noninferiority trial.

Study	Comparison(s) vs. Intervention(s)	Outcomes (critical or important)	Design
	PET-driven treatment: if PET2 had been positive: 2×eBEACOPP or if PET2 had been negative: 2×ABVD If PET-4 positive: Salvage at the discretion of the investigator		
RATHL, Johnson, 2016 [49]	2 ABVD then PET2. Then: If PET negative: 4×AVD vs. 4×ABVD The PET2-positive patients were not randomized to treatment.	OS *Recurrence*: PFS AE: TRD, SPM, diffusing capacity of the lung for carbon	RCT multicentre, open label, noninferiority
GITIL/FIL HD0607, Gallamini, 2020 [60]	2×ABVD then PET: If PET-2-positive: 4×eBEACOPP+4×bBEACOPP vs. Ritux + 4×eBEACOPP+4×bBEACOPP If PET-2-negative: See Question 3	OS *Recurrence*: PFS AE: nr	RCT, open label, phase II, 2-stage design
4. ABVD vs. modified ABVD with brentuximab vedotin instead of bleomycin			
ECHELON-1, Connors, 2018 [9]	6×ABVD vs. 6×A+AVD	OS *Recurrence*: PFS AE: TRD, SPM	RCT phase 3, open-label, multicentre trial
5. Modified ABVD vs. another modified ABVD			
No fully published trial final results are available for this comparison	NA	NA	NA
6. BEACOPP vs. modified BEACOPP			
No fully published trial final results are available for this comparison	NA	NA	NA
QUESTION 2: pts ≥60 yrs of age			
Subgroup analyses of RCTs			
ECHELON-1, Evens, 2022 [61]	A-AVD vs. ABVD in pts ≥60 yrs of age	PFS, TRD	Subgroup analysis of the ECHELON-1 [9]
ECHELON-1, Hutchings, 2022 abs [62]	A-AVD vs. ABVD in pts ≥60 yrs of age	OS	Subgroup analysis of the ECHELON-1 [9]
HD9, Ballova, 2005 [63]	COPP-ABVD regimen with bBEACOPP in older pts	OS, FFTF	Subgroup analysis of the HD9 [64]
Observational comparative trials			
No comparative observational trials were identified			
Phase 2 single-arm studies			
1. Various ABVD-like treatments and RT			
Yildiz, 2021 [65]	ABVD combination (n=45, 88.2%) 1 to 8 cycles	OS, TRD	Phase 2 single-arm

Study	Comparison(s) vs. Intervention(s)	Outcomes (critical or important)	Design
	AVD (n=5, 5.9%) Bendamustine/brenduximab (n=1, 2%) Doxorubicin, bleomycin, dacarbazine (ABD) (n=1, 2%) Gemcitabine, cyclophosphamide, vincristine, dexamethasone (GCVP) (n=1, 2%) Consolidation RT (n=12, 23.6%)		
PLRG-R9 study, Wrobel, 2019 [66]	ABVD/ABVD like +/- RT 86% CHOP/PVAG: 7% BEACOPP: 3% Palliative: 4%	OS, Recurrence, AEs	Phase 2 single-arm
2. Brentuximab vedotin alone or in combination			
Evens, 2018 [67]	Sequential brentuximab vedotin (BV) given before and after AVD	OS, PFS, treatment-related morbidity	Phase 2 single-arm
BREVITY trial, Gibb, 2021 [68] (Part A)	Brentuximab vedotin monotherapy	OS, PFS, TRD	Phase 2 single-arm
BREVITY trial Friedberg, 2024 [69] (Parts B and D)	BV in combination with dacarbazine BV in combination with nivolumab	OS, PFS, treatment-related AEs	Phase 2 single-arm
ACCRU trial Cheson, 2020 [70]	BV in combination with nivolumab	OS, PFS, peripheral neuropathy	Phase 2 single-arm
3. VEPEMB			
SHIELD study, Proctor, 2012 [71]	VEPEMB (vinblastine, cyclo- phosphamide, prednisolone, procarbazine, etoposide, mitoxantrone, and bleomycin) (n=103, of which 72 with advanced stage) ABVD (n=35) CLVPP (n-19) Other (n=18)	OS, PFS, TRD, pulmonary fibrosis	Phase 2 single-arm
Levis, 2004 [72]	6×VEPEMB (vinblastine, cyclophosphamide, procarbazine, etoposide, mitoxantrone and bleomycin) +RT to bulky/not responding areas	OS, RFS, DSS, FFS	Phase 2 single-arm
4. PVAB			
Ghesquieres, 2024 [73]	6×PVAB (i.e., 6 cycles of prednisone, vinblastine, doxorubicin, and bendamustine)	OS, PFS, TRD, AEs	Phase 2 single-arm
5. ABVD			
Cokgezer, 2022 [74]	ABVD AVD and mini-CHOP (400 mg/m^2 cyclophosphamide, 25 mg/m^2 doxorubicin, 1.4 mg/m^2 vincristine, and 40 mg/m2 prednisolone)	OS, PFS	Retrospective chart review
QUESTION 3: Consolidation radiotherapy			
1. Consolidation RT vs. observation			
HD0801 FIL, Ricardi, 2021 [75]	Consolidation RT vs. Observation	EFS, PFS, OS	RCT phase 3, open label, multicentre

Study	Comparison(s) vs. Intervention(s)	Outcomes (critical or important)	Design
GITIL/FIL HD0607, Gallamini, 2020 [48,60]	Consolidation RT vs. no further treatment (only PET-2-negative pts)	OS, PFS, DFS	RCT open label, phase 2 2-stage design
EORTC 20884, Aleman, 2007 [76]	IF-RT vs. no further treatment	OS, EFS	RCT open label
HD12, Borchmann, 2011 [59], von Tresckow, 2018 [77]	Chemotherapy + RT (n=755) vs. Chemotherapy – RT (n=765)	OS, PFS, FFTF, AE (long term)	RCT open label, multicentre
2. RT vs. Chemotherapy			
ECOG E1476, Wiernik, 2009 [78]	RT vs. 3×ABVD	OS, DFI, Relapse	RCT open label
Aviles, 2000 [79]	Chemotherapy vs. RT + Chemotherapy	OS, FFS	RCT open label

NOTES:
*The HD9 trial had a noninferiority design only for its secondary objective, that is represented in our comparison 2.

A-AVD=brentuximab vedotin plus doxorubicin, vinblastine, and dacarbazine; ABD=Doxorubicin, bleomycin, dacarbazine; ABVD=doxorubicin, bleomycin, vinblastine, and dacarbazine; AE=Adverse events; AVD=doxorubicin, vinblastine, and dacarbazine; bBEACOPP=baseline BEACOPP; BEACOPP=bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone; $BEACOPP_{14}$=BEACOPP given on a shorter, 14-day cycle; BV=brentuximab vedotin; CLVPP=chlorambucil, vinblastine, procarbazine hydrochloride, and prednisolone; COPP=cyclophosphamide, vincristine, procarbazine, and prednisone; COPPEBVCAD-CEC=cyclophosphamide, lomustine, vindesine, melphalan, prednisone, epidoxirubicin, vincristine, procarbazine, vinblastine, and bleomycin; CR=complete response; DFI=disease-free interval; DFS=Disease-free Survival; DSS=Disease-specific Survival; eBEACOPP=escalated-dose BEACOPP; $eBEACOPP_4$+$bBEACOPP_4$= 4 cycles of eBEACOPP plus 4 cycles of bBEACOPP; EFS=event-free survival; FFFP=freedom from first progression; FFP=freedom from Progression; FFS=failure-free survival; FFTF=freedom from treatment failure; G-CSF=granulocyte colony-stimulating factor; GCVP=gemcitabine, cyclophosphamide, vincristine, dexamethasone; NA=not applicable; nr=not reported; OS=overall survival; PET=positron emission tomography; PFS=progression-free survival; pts=patients; PVAB=prednisone, vinblastine, doxorubicin and bendamustine; PVAG=prednisone, vinblastine, doxorubicin, and gemcitabine; RCT=randomized controlled trial; RFS=relapse-free survival; Ritux=rituximab; RT=radiotherapy; SPM=second primary malignancy; TRD=treatment-related deaths; TTP=thrombotic thrombocytopenic purpura; VEPEMB=vinblastine, cyclophosphamide, prednisolone, procarbazine, etoposide, mitoxantrone, and bleomycin

RCTs

We identified 7705 records from the searches. The methodologist (FGB) selected 514 articles for full-text review after reviewing the titles and the abstracts against the selection criteria. We retrieved the full text of these articles in the library and the methodologist screened them against the selection criteria and included 106 articles. During data extraction 39 additional publications were excluded. Among them was one abstract publication of the analysis of Part 1 (tolerability) of the ongoing HD21 trial [22] and two abstract publications of the ongoing SWOG S1826 trial [23,80].

The reasons for exclusion are reported in the study flow chart in Appendix 3B. The remaining 67 included publications represent 13 unique fully published randomized studies in 12 publications for Question 1 [9,16,17,49,52-55,57-60], and six fully published unique studies for Question 3 [59,60,75,76,78,79] (Table 3). The authors of each of the individual studies tested several comparisons of different doses and schedules that were of interest to this report.

Additionally, 15 long-term follow-up publications of the unique studies [48,50,56,64,76,77,81-89] (Table 4), 19 companion studies reporting on outcomes of interest [61-63,77,90-103] (Table 3-7A in Appendix 7), and five pooled analyses of IPD from the repository of the GHSG [104-108] (Table 4-7 A in Appendix 7) were included. The authors of these pooled analyses used and combined the raw data from some of the unique trials that we included. They reported data on some rare or long-term events in patients with advanced-stage HL that are not reported elsewhere and that can be of interest to practicing clinicians (e.g, incidence, risk factors and timing of symptomatic osteonecrosis [104]; characteristics and prognosis of late relapse [105]; the impact of dose reductions of bleomycin and vincristine [106]; the incidence, outcome, and risk factors of treatment-related acute myeloid leukemia (AML)/myelodysplastic syndromes (MDS) [107]; and the risk factors associated with treatment-related mortality in patients treated with eBEACOPP [108]). However, these pooled analyses did not compare these effects among the drug combinations that are the focus of this document; therefore, we will not discuss them any further, and the interested reader can find more information about these studies in Table 4-7A, Appendix 7.

Eleven additional ancillary studies of the 13 included studies reported outcomes that were not considered critical or important by the Working Group [109-119] and we listed them in Table 5-7A, Appendix 7.

We did not include abstracts of interim analyses of ongoing trials in our analysis, but we present them in Table 7-7A, Appendix 7.

Tables 4 and 5 present the general characteristics of the unique studies included for Questions 1 and 3 with their long-term follow-ups, if available. The characteristics and summary results of ancillary trials of the studies included for Questions 1 and 3 that addressed important or critical outcomes and the characteristics and results of the pooled analyses of the GHSG prospective IPD repository are reported in Tables 3-7A and 4-7A, in Appendix 7.

For Question 1, the studies are grouped into six comparisons: 1) ABVD versus BEACOPP; 2) higher-intensity/dose (more cycles) versus lower-intensity/dose (less cycles, lower dose) eBEACOPP; 3) PET-adapted strategies; 4) ABVD versus modified ABVD with brentuximab instead of bleomycin, 5) modified ABVD versus another modified ABVD; and 6) BEACOPP versus modified BEACOPP (Table 3).

For Question 3, the studies are grouped in two comparisons: 1) consolidation radiotherapy versus observation; and 2) radiotherapy versus chemotherapy (Table 3).

Table 4. Question 1: General characteristics of studies.

Study Name Author(s) [Reference], Years of Accrual Country Funding Notes	Study Objectives, Design, Median Follow-up	Number of Patients included / analyzed Population Age Stage Required sample size	Intervention and comparison groups, Schedule, Other treatment	Outcomes Measured
Comparison 1. ABVD vs. BEACOPP				
EORTC 20012 Intergroup Trial Carde, 2016 [52] Years of accrual: 2002 to 2010 Country: Multiple countries in Europe, Australia, and North America Funding: Industry and Fonds Cancer of Belgium *Note:* This study looks at the advanced high-risk pts, the parallel study by Mounier [53] looks at the low-risk pts in a subset of the centres.	Objectives: to clarify whether $eBEACOPP_4$+$bBEACOPP_4$, provides a better EFS and leads to a longer OS than ABVD in pts at high risk (IPS ≥3) advanced HL Design: RCT, superiority, single blinded (outcome assessors) Follow-up: 43.2 mos	N=550 included, 549 in ITT analysis; 541 in the safety analysis; 531 in the per-protocol analysis. Untreated histology-proven classic HL clinical stages III and IV, at least one bi-dimensionally measurable target lesion, performance status of 0 to 2, high-risk (IPS≥3) Age, median: 35.2 yrs (range 16.1-67.4 yrs) Gender: male 74.9% Stage: II: 0.4% III: 25.7% IV: 73.6% Sample size: 152 events and 550 pts were required over 5.5 yrs to detect a 10% increase in the 3-yr EFS rate from 70% ($ABVD_8$) to 80% ($BEACOPP_{4+4}$), for a power of 80% (2-sided log-rank test, alpha=0.05)	n=275 8×ABVD (safety population n=272) vs. n=274 $eBEACOPP_4$+$bBEACOPP_4$ (safety population n=269) Schedule: 8 cycles of ABVD given every 4 wks vs. 4 cycles of eBEACOPP + 4 cycles of bBEACOPP (BEACOPP4+4) given every 3 wks Other treatment: G-CSF mandatory with eBEACOPP	EFS* PFS* CR rate OS QoL SPM DFS Cost-effectiveness
LYSA H34 Mounier, 2014 [53] Years of accrual: February 2003 to August 2008 Country: France Funding: French Programme Hospitalier de Recherche Clinique and Chugai Company	Objectives: To assess the EFS gain after stratification in low-risk pts with an IPS ranging from 0 to 2 Design: RCT, open label, parallel group, phase III Follow-up: 66 mos	N=150 pts, 145 pts in analysis. Nodular, lymphocyte predominant type was excluded. Age (Median): 28 yrs (range 16-60 yrs) Gender: 50% male Stage: III: 52% IV: 48% IPS 0-1: 64% Sample size:	(n=80) 8×ABVD vs. (n=70) 4×eBEACOPP+4×bBEACOPP Schedule: ABVD for 8 cycles, depending on the response assessed at the end of cycles 4 (C4), 6 (C6) and 8 (C8), using clinical and computed tomography (CT) criteria. BEACOPP for 8 cycles (eBEACOPP for 4 cycles, followed by bBEACOPP for 4 cycles). Other treatment: G-CSF and trimethoprim/sulfamethoxazole was mandatory in the BEACOPP arm.	EFS* CR/Cr(u) PFS OS SPM Death Pulmonary toxicity

Study Name Author(s) [Reference], Years of Accrual Country Funding Notes	Study Objectives, Design, Median Follow-up	Number of Patients included / analyzed Population Age Stage Required sample size	Intervention and comparison groups, Schedule, Other treatment	Outcomes Measured
Note: this is a trial parallel to the larger EORTC20012 trial [52]. This study included only the LYSA centres and included low-risk pts.		26 events 150 pts were required to detect a 15% increase in the 5-year EFS rate from 75% (ABVD) to 90% (BEACOPP), one-sided log-rank test at error rates of α=0.05 and ß=0.20.		
HD9 Diehl, 2003 [54] 5-yr follow up; Diehl, 1998 [64] 1st and 2nd Interim analyses Engert, 2009 [81] 10-yr follow-up; von Tresckow, 2018 [77] 15 yrs follow-up Years of accrual: February 1993 to March 1998 Country: Germany, Switzerland, Austria, the Netherlands, and the Czech Republic Funding: Deutsche Krebshilfe and Swiss Group for Clinical Cancer Research *Notes:* 1. After the first interim analysis this study was stopped 2. This study appears also in comparison 2 Higher versus lower BEACOPP intensity/dosage.	Objectives: To test the hypotheses that: 1) BEACOPP (irrespective of the dose) results in a higher rate of FFTF than COPP-ABVD and that 2) increased-dose BEACOPP results in a higher rate of freedom from treatment failure than standard BEACOPP. Results for this second objective are reported in Comparison 2. Design: RCT 3 arm noninferiority trial Follow-up (median): 23 mos [64] at the 2nd interim analysis. At the 5-yr analysis [54]: COPP/ABVD: 72 mos bBEACOPP: 54 mos eBEACOPP: 51 mos At the 10-yr: 111 mos (range 3-167) (respectively in the 3 groups: 122, 111, 107) At 15 yrs 141 mos median	N=1282, (COPP/ABVD: 288 [260 in analysis]; bBEACOPP 498 [469 in analysis]; eBEACOPP 496 assigned [466 in analysis]) 1195 in analysis in the original publication [54], 1196 in Engert, 2009 [81] Number of events: 95 [64] Age (median yrs): 32 (COPP/ABVD), 33 (bBEACOPP), 32 (eBEACOPP) Stage (%): COPP/ABVD bBEACOPP eBEACOPP IIB 9% 14% 16% IIIA 31% 24% 21% IIIB 29% 28%. 31% IVA 9%. 9%. 10% IVB 23%. 26% 23% Male gender: 57% 63% 62% Sample size: 900 pts needed to be enrolled to obtain a statistical power of 80% to detect an absolute difference of 9% to 10% in either of the aimed comparisons. However, assignment to the COPP/ABVD was stopped after the first interim analysis for benefit of the BEACOPP arms (p=0.03), therefore the assignment to the COPP/ABVD arm was stopped, and the assignments were accrued to include 500 pts in the 2 BEACOPP arms	n=260 8×COPP/ABVD no G-CSF (enrollment for this group was stopped early for benefit at the 1st interim analysis, 2 yrs with 125 pts in the COPP/ABVD, 131 pts in the bBEACOPP arm and 65 pts in the eBEACOPP arm) n=469 8×bBEACOPP no G-CSF n=466 8×eBEACOPP G-CSF Treatment schedule: 4 double cycles of COPP+ABVD 8 cycles of bBEACOPP; 8 cycles of eBEACOPP In the COPP/ABVD and bBEACOPP arms next cycle was postponed of ≥ 2 wks if leukocyte and thrombocyte values were <2.500/µL and 80.000µ/L. In the eBEACOPP arm a stepwise reduction of cyclophosphamide and etoposide were implemented if severe toxicity. Other treatment: G-CSF only in the eBEACOPP group 300 to 480 µg depending on pt weight. Additional radiotherapy (64% COPP-ABVD and 71% b and eBEACOPP) to initial bulk (only if ≥ 5 cm, 30 Gy) and residual tumours (40 Gy) after completion of chemotherapy and restaging.	FFTF* AE OS rate CR PFS at 15 yrs

Study Name Author(s) [Reference], Years of Accrual Country Funding Notes	Study Objectives, Design, Median Follow-up	Number of Patients included / analyzed Population Age Stage Required sample size	Intervention and comparison groups, Schedule, Other treatment	Outcomes Measured
HD2000 Gruppo Italiano per lo Studio dei Linfomi Trial Federico, 2009 [55]; Merli, 2016 [56] (10-yr follow-up post-hoc analysis) Years of accrual: April 2000 to June 2007 Country: Italy Funding: Associazione Angela Serra per la Ricerca sul Cancro (Modena, Italy); Gruppo Amici dell'Ematologia (Reggio Emilia, Italy)	Objectives: To compare ABVD vs. BEACOPP vs. COPPEBVCAD-CEC in pts with advanced HL Design: RCT multicentre Follow-up: 41 mos [55] then 120 mos [56]	N=307 pts; 305 in ITT analysis; 295 in per protocol analysis Age (median): ABVD: 32 yrs BEACOPP: 29 yrs CEC: 33 yrs Stage: ABVD BEACOPP CEC IIB. 33 31 29 IIIA. 27 21 21 IIIB. 17 25 24 IVA. 13 6 10 IVB. 9 16 16 Gender: Male: 43 60 56 Sample size: 282 patients were required to detect a HR of 0.4 for FFS in the experimental arm with 80% power assuming 5-yr FFS of 65% in the ABVD arm as a reference. Since a drop-out rate of 10% was expected after randomization, a cohort of 310 pts was necessary to test the hypothesis	(n=103) 6×ABVD vs. (n=102) 4×eBEACOPP+2bBEACOPP vs. (n102) 6×COPPEBVCAD-CEC Other treatment: Partial RT program for patients with previous bulk or slowly or partially responding sites (30 Gy to 36 Gy with a boost of 6 Gy to persisting sites; bulk thoracic mass of 6 cm or 10 cm mass outside of mediastinum), administered to 46%, 44% and 43% respectively in the groups. G-CSF at the discretion of treating physician	FFS* PFS OS RFS AE (myelotoxicity, SPM) Cause of death
GITIL Viviani, 2011 [57] Years of accrual: March 2000 to March 2007 Country: Italy Funding: Fondazione Michelangelo	Objectives: To compare ABVD and BEACOPP in pts with advanced-stage HL Design: RCT multicentre Follow-up: 61 mos	N=331; 331 in ITT analysis, (322 in per-protocol analysis) pts with advanced-stage HL who were planned to receive high-dose salvage treatment Age <45 yrs: ABVD: 77% BEACOPP: 82% Stage: Stage III or IV: ABVD: 24 (53%); BEACOPP: 13 (65%) Gender: ABVD: 60% male	n=168 in the efficacy analysis, and n=166 in the safety analysis 6×ABVD (if CR reached after 4 cycles) or 8×ABVD n=163 in the efficacy analysis, and n=156 in the safety analysis 4×eBEACOPP + 4×bBEACOPP Treatment duration (Median): ABVD: 34 wks BEACOPP: 26 wks Other treatment: salvage regimen: i.e., reinduction standard-dose ifosfamide-based chemotherapy followed by a high-dose consolidation therapy with BEAM and	FFFP* rate FFSP rate EFS OS AE

Study Name Author(s) [Reference], Years of Accrual Country Funding Notes	Study Objectives, Design, Median Follow-up	Number of Patients included / analyzed Population Age Stage Required sample size	Intervention and comparison groups, Schedule, Other treatment	Outcomes Measured
		BEACOPP: 56% male Sample size: The study was designed to detect at least a 15% difference between the two groups in FFFP rate, with 80% power at a two-sided alpha of 5%. The number of required patients or events is nr	autologous hematopoietic stem-cell support; RT (66% vs. 67%)	
Comparison 2. Higher- (i.e., higher dose, more cycles) vs. Lower- intensity/dosage (i.e., lower dose, less cycles) BEACOPP				
HD15 Engert, 2012 [58], Engert, 2017 [82] Years of accrual: January 2003 to April 2008 Country: Germany, Switzerland, Austria, the Netherlands, and the Czech Republic Funding: Deutsche Krebhilfe and Swiss Federal Gorvernment	Objectives: To compare 2 reduced-intensity chemotherapy regimens followed by PET-guided RT, and to assess the reduction of toxicity while maintaining efficacy Design: RCT, phase 3, open label, noninferiority trial. Follow-up (median): 48 mos [58], and 102 mos [82]	N=2182, 2126 in ITT analysis,1741 in per-protocol analysis pts with advanced-stage HL. 4% had nodular lymphocyte predominant disease. Age (median, range): 33 yrs (18-60) Stage: IIB: 15.6% IIIA: 23% IIIB: 27% IVA: 9.9% IVB: 24.3% Gender (male): 8×eBEACOPP: 61% 6×eBEACOPP: 61% 8×$BEACOPP_{14}$: 60% Sample size: For a FFTF of 87% in the 8×eBEACOPP arm and a maximum tolerated decrease in efficacy of 6%, the non-inferiority margin for the HR was 1.51	n=728 allocated, 705 in ITT analysis 8×eBEACOPP vs n=726 allocated, 711 in ITT analysis 6×eBEACOPP vs. n=728 allocated, 710 in ITT analysis 8×$BEACOPP_{14}$ Treatment duration: Cycle frequency: $BEACOPP_{14}$: 14 ds eBEACOPP: 21 ds Number of cycles: $BEACOPP_{14}$: Up to 8 cycles eBEACOPP: up to 6 cycles Other treatment: RT: 30 Gy involved field RT if partial remission after chemotherapy (residual tumour ≥2.5 cm), and PET + Erythropoietin: All patients were randomized to receive erythropoietin or placebo	FFTF* OS PFS AE (treatment-related death, SPM)
HD12 Borchmann, 2011 [59], von Tresckow, 2018 [77] (97 mos follow-up) Years of accrual: January 1999 to January 2003	Objectives: To test a regimen that could decrease the toxicity associated with 8 cycles of eBEACOPP + RT while maintaining efficacy, specifically: 1) Chemotherapy question: to test the effect of	N=1670, 1574 in analysis for chemotherapy comparison; 1520 in analysis for RT comparison Newly diagnosed pts with histology-proven HL in stages IIB with a large mediastinal mass (≥ 1/3 maximum thoracic diameter) or extra nodal lesions, and stages III and IV.	A) 8×eBEACOPP+RT in responding pts with initial bulk or residual tumour (n=392) vs. B) 8×eBEACOPP no RT (n=395) vs. C) 4×eBEACOPP +4×bBEACOPP + RT (n=393) vs. D) 4×eBEACOPP +4×bBEACOPP no RT (n=394) Chemotherapy comparison: 8×eBEACOPP ± RT (n=787) vs.	FFTF* PFS OS AE

Study Name Author(s) [Reference], Years of Accrual Country Funding Notes	Study Objectives, Design, Median Follow-up	Number of Patients included / analyzed Population Age Stage Required sample size	Intervention and comparison groups, Schedule, Other treatment	Outcomes Measured
Country: Germany, Switzerland, Austria, the Netherlands, and the Czech Republic Funding: Deutsche Krebshilfe, the Swiss Federal Government, and the Bundesministerium fu¨r Bildung und Forschung, Germany **Note**: this study is repeated below in Question 3 for the radiotherapy comparison.	replacing the last 4×eBEACOPP with 4 cycles of bBEACOPP; 2) Radiotherapy question: to evaluate the impact of using RT for consolidation in pts responding to chemotherapy who had initial bulky disease (≥5 cm or residual disease 1.5 cm) (see Table 4-3 for details related to question 3) Design: RCT 2×2 factorial, noninferiority trial Follow-up: 78 mos	Age (Mean): Chemotherapy comparison: 35.5 yrs Gender (male): Chemotherapy comparison:61.2% Stage: IIB: 16.2% IIIA: 22.0% IIIB: 26.8% IVA: 9.5% IVB: 25.4% Sample size: The study was designed to exclude a difference in 5-yr FFTF rates of 6% in comparisons of pooled treatment arms for both treatment modalities assuming no interaction	4×eBEACOPP + 4×bBEACOPP± RT (n=787) Radiotherapy comparison: Chemotherapy + RT (n=755) vs. Chemotherapy – RT (n=765) Other treatment: none reported.	
HD9 Diehl, 2003 [54] 5-yr follow up; Diehl, 1998 [64] 1st and 2nd Interim analyses Engert, 2009 [81] 10-yr follow-up; Von Treskow, 2018 [77], 15-yr follow-up; [120] Erratum Years of accrual: February 1993 to March 1998 Country: Germany, Switzerland, Austria, the Netherlands, and the Czech Republic Funding: Deutsche Krebshilfe and Swiss Group for Clinical Cancer Research *Notes:*	Objectives: To test the hypotheses that: 2) Increased-dose BEACOPP results in a higher rate of FFTF than standard BEACOPP Design: RCT 3 arm noninferiority trial Follow-up (median): 23 mos [64] at the 2nd interim analysis. At the 5-yr analysis [54]: COPP/ABVD: 72 mos bBEACOPP: 54 mos eBEACOPP: 51 mos At the 10-yr: 111 mos (range 3-167) (respectively in the 3 groups: 122, 111, 107) At the prolongued follow-up (15 yrs) 141 mos median	N=1282, (COPP/ABVD: 288; bBEACOPP 498; eBEACOPP 496 assigned) 1195 in analysis in the original publication [54], 1196 in Engert, 2009 [81] Number of events: 95 [64] Age (median yrs): 32 (COPP/ABVD), 33 (bBEACOPP), 32 (eBEACOPP) Stage (%): COPP/ABVD bBEACOPP eBEACOPP IIB 9% 14% 16% IIIA 31% 24% 21% IIIB 29% 28%. 31% IVA 9%. 9%. 10% IVB 23%. 26% 23% Male gender: 57% 63% 62% Sample size:	n=261 8×COPP/ABVD no G-CSF (enrollment for this group was stopped early for benefit at the 1st interim analysis, 2 yrs with 125 pts in the COPP/ABVD, 131 pts in the bBEACOPP arm and 65 pts in the eBEACOPP arm) n=469 8×bBEACOPP no G-CSF n=466 8×eBEACOPP G-CSF Treatment schedule: 4 double cycles of COPP+ABVD 8 cycles of bBEACOPP; 8 cycles of eBEACOPP In the COPP/ABVD and bBEACOPP arms next cycle was postponed of ≥ 2 wks if leukocyte and thrombocyte values were <2.500/µL and 80.000µ/L. In the eBEACOPP arm a stepwise reduction of cyclophosphamide and etoposide were implemented if severe toxicity. Other treatment: G-CSF only in the eBEACOPP group 300 to 480 µg depending on pt weight.	FFTF* AE OS rate CR PFS at 15 yrs

Study Name Author(s) [Reference], Years of Accrual Country Funding Notes	Study Objectives, Design, Median Follow-up	Number of Patients included / analyzed Population Age Stage Required sample size	Intervention and comparison groups, Schedule, Other treatment	Outcomes Measured
1. After the first interim analysis this study was stopped 2. This study appears also in comparison 2 Higher versus lower BEACOPP intensity/dose.		900 pts needed to be enrolled to obtain a statistical power of 80% to detect an absolute difference of 9% to 10% in either of the aimed comparisons. However, assignment to the COPP/ABVD was stopped after the first interim analysis for benefit of the BEACOPP arms (p=0.03), therefore the assignment to the COPP/ABVD arm was stopped, and the assignments were accrued to include 500 pts in the 2 BEACOPP arms	Additional radiotherapy (64% COPP-ABVD and 71% b and eBEACOPP) to initial bulk (only if ≥ 5 cm, 30 Gy) and residual tumours (40 Gy) after completion of chemotherapy and restaging.	
Comparison 3. PET-adapted strategies				
HD18 Borchmann, 2018 [16] (final analysis) Borchmann, 2017 [50] (PET-2 positive pts; results published early at the second interim analysis) Kreissl, 2021 [83] (5-yr follow-up) Years of accrual: May 2008 – July 2014 Country: Germany, Switzerland, Austria, the Netherlands, and the Czech Republic Funding: Deutche Krebshilfe, Swiss State Secretariat for Education and Research, and Roche Pharma AG	Objectives: PET-2 positive pts: To show superiority of R-eBEACOPP vs. eBEACOPP for PFS[a]. At the 3-yr interim analysis futility was concluded for this question and these pts were no longer randomized, but they all received 4 more cycles of eBEACOPP. PET-2 negative pts: To show non-inferiority of eBEACOPP with reduced number of cycles compared to standard eBEACOPP for PFS. Design: PET-2 positive pts: RCT, open label; superiority trial, stopped early for futility. PET-2 negative pts: noninferiority trial Follow-up in final analysis: PET-2 positive pts: 66 mos (range 46 to 108 mos); pts	N=1964 pts enrolled and screened with PET-2. Newly diagnosed, advanced stage pts, histology proven HL. 1945 included in the ITT[b] analysis. 920 PET-2 negative pts were included in a per-protocol analysis. 440 PET-2 positive pts were included. Age: Entire group: median 32 yrs (range18-60 yrs) PET-2 positive: median 30 yrs (range18-60) Gender: Entire group: 61% male PET-2 positive: 60% male Stage: IIB (14%), IIIA (24%), IIIB (25%), IVA (12%), IVB (25%) **Expected difference**: 15% improvement in PFS at 5 yrs in the PET-adapted group Sample size: PET positive arm:	2×eBEACOPP then PET. PET2+ pts were randomized to eBEACOPP vs. modified eBEACOPP (adding Ritux) PET2– pts were randomized to Higher dose BEACOPP vs. lower dose BEACOPP PET-2+ pts: n=220 8×eBEACOPP (Standard) (217 in ITT Analysis) + n=511 6×eBEACOPP (New Standard)[c] (not included in randomization) vs. n=220 8×R-eBEACOPP (217 in ITT Analysis) PET2– pts: (n=288) 8×eBEACOPP (Standard) + (n=216 after amendment) 6×eBEACOPP New Standard)[c] vs. (n=285) 4×eBEACOPP + (n=216 after amendment) The per-protocol population included 920 pts (92% of the total) Treatment duration: 12 weeks Other treatments: RT given 4-6 wks after completion of chemotherapy for pts with lesions ≥2.5 cm.	PFS* at 5 yrs OS SPM Treatment-related AE CR HL-specific death Treatment-related toxicities Late toxicity QoL

Study Name Author(s) [Reference], Years of Accrual Country Funding Notes	Study Objectives, Design, Median Follow-up	Number of Patients included / analyzed Population Age Stage Required sample size	Intervention and comparison groups, Schedule, Other treatment	Outcomes Measured
	who received 6×eBEACOPP: 36 mos; PET-2 negative pts: 56 mos (range, *nr*, IQR 36-72) Follow-up at 5-yrs: PET-2 positive pts (ITT cohort): Pre-amendment[c]: 73 mos This from text p.e403 #2494 Follow-up for disease status: PET-2 positive pts: pre-amendment: 8×eBEACOPP: 75 mos 8×R-eBEACOPP: 72 mos Post-amendment: 6×eBEACOPP: 58 mos PET-2 negative pts: Pre-amendment: 8×eBEACOPP: 76 mos 4×eBEACOPP: 75 mos Post-amendment: 6×eBEACOPP: 59 mos 4×eBEACOPP: 57 mos Pre-and Post-amendment combined: 8 or 6×eBEACOPP: 66 mos 4×eBEACOPP: 64 mos Follow-up for survival status: PET-2 positive pts: pre-amendment: 8×eBEACOPP: 77 mos 8×R-eBEACOPP: 79 mos PET-2 negative pts: Pre-amendment:	To detect an improvement of ≥15% in 5-yr PFS with a power of 80% and a two-sided alpha of 5%, 1050 were required. The trial was stopped at the second interim analysis, when the results of study HD15 became available, and the sample was 27 short of 1050. PET negative arm: If the lower limit of the two-sided 95% CI for the difference in the 5-yr PFS estimates was above −6% non-inferiority was established. Assuming a 5-yr PFS of 88% with standard therapy and 80% power, 870 PET negative pts were required for the per-protocol analysis. In both arms, a total of 1945 patients were to be included in the trial	Total RT dose: 30 Gy given in single fractions of 1.8-2.0 Gy 5 x wk. G-CSF given to all pts.	

Study Name Author(s) [Reference], Years of Accrual Country Funding Notes	Study Objectives, Design, Median Follow-up	Number of Patients included / analyzed Population Age Stage Required sample size	Intervention and comparison groups, Schedule, Other treatment	Outcomes Measured
	8×eBEACOPP: 83 mos 4×eBEACOPP: 78 mos Post-amendment: 6×eBEACOPP: 60 mos 4×eBEACOPP: 61 mos Pre-and Post-amendment combined: 8 or 6×eBEACOPP: 69 mos 4×eBEACOPP: 66 mos			
AHL2011 Casasnovas, 2019 [17] Casanovas, 2022 [84] (for updated late treatment - related AE such as SPM, and fertility) Years of accrual: May 2011 to April 2014 Country: France, Belgium Funding: Programme Hospitalier de Recherche Clinique	Objectives: To assess a PET-driven strategy after 2 cycles of eBEACOPP with PET-2-negative pts switching to ABVD, and PET-2-positive pts continuing with eBEACOPP compared with standard care of 6 cycles of eBEACOPP Design: RCT, open label, multicentre, phase 3, noninferiority trial. The study had a noninferiority margin of 10% on PFS at 5 yrs: PFS at 5 yrs >75% in the PET-driven group, and the upper limit of the HR was <1.77 Follow-up: 50.4 mos [17], 67.25 mos [84],	N=823; (739 in per-protocol analysis, 823 in ITT analysis, 819 in safety analysis), adult or young adult pts newly diagnosed with advanced HL and performance score <3. Age, median (IQR): 30 yrs (24-41) Stage: IIB: 10.6% III or IV: 89% IV: 60.2% Gender: 63% male **Sample size** **Expected difference**: Non-inferiority based on a clinical margin of 10% on PFS at 5 yrs; PFS >75% in the PET-driven group vs. 70% in the standard group. Non-inferiority was established if the upper limit of the HR was <1.77 with alpha=2.5%. Required sample size: 97 PFS events (810 pts) were required to obtain 80% power with a one-sided significance level of 0.025	(n=413 ITT analysis; 372 per protocol analysis) 6×eBEACOPP vs. (n=410 ITT analysis;367 per protocol analysis) PET driven strategy. All pts received induction with 2×eBEACOPP. Then PET was done: Standard treatment: 4×eBEACOPP PET-driven treatment: If PET-2-negative: 2×ABVD If PET2 positive: 2×eBEACOPP Then PET4: Consolidation If PET-4-negative: Standard treatment: 2×eBEACOPP PET-driven treatment: if PET2 had been positive: 2×eBEACOPP or if PET2 had been negative: 2×ABVD If PET-4 positive: Salvage at the discretion of the investigator. Treatment duration: eBEACOPP was given every 21 days ABVD was given every 28 days Other treatments: G-CSF was mandatory with eBEACOPP and optional with ABVD Trimethoprim/sulfamethoxazole and valacyclovir to all as opportunistic infections prevention	PFS at 5 yrs* OS AE Fertility in pts <45 yrs old Overall response

Study Name Author(s) [Reference], Years of Accrual Country Funding Notes	Study Objectives, Design, Median Follow-up	Number of Patients included / analyzed Population Age Stage Required sample size	Intervention and comparison groups, Schedule, Other treatment	Outcomes Measured
RATHL Johnson, 2016 [49,85] clinicaltrials.gov NCT00678327 Years of accrual: August 2008 to December 2012 Country: UK, Italy, Australia, New Zealand, Norway, Sweden, and Danemark Funding: Cancer Research UK; Leukaemia and Blood Cancer New Zealand, and Cancer Australia	Objectives: To test interim PET after 2 cycles of ABVD as a measure of early response to chemotherapy in order to guide treatment for patients with advanced HL In PET-2-negative pts: to test noninferiority of AVD vs. ABVD Design: RCT multicentre, noninferiority (AVD vs. ABVD) Follow-up: 41.2 mos	N=1135 pts with advanced HL PET-2-negative pts: N=935 (ITT analysis); 924 pts (per protocol analysis) PET-2-positive pts: N=172 (ITT analysis) *(Note: the authors set a 950 pts with 101 events as the required sample size to have 90% power, then revised it to 91 events and 1238 pts)* Events: 142 Age (Median): 33 yrs Stage: II: 41.6% III: 30.2% IV: 28.3% Gender: male 54.5% **Sample size** 91 events required, sample size 1230 randomized pts for the trial to have 90% power to exclude a 5% at 3-yr with alpha=0.025.	PET negative after 2×ABVD: (n=465) 4×AVD vs. (n=470) 4×ABVD PET positive after 2×ABVD: (n=94) 4×BEACOPP$_{14}$ vs. (n=78) 3×eBEACOPP Note: The BEACOPP part of the study was not randomized, therefore not further details are reported) Schedule: ABVD and AVD: on ds 1 and 15 of each 28 ds cycle for 4 cycles BEACOPP$_{14}$: in a 14 ds cycle, doxorubicin hydrochloride I and cyclophosphamide on d 1; etoposide on ds 1-3; oral procarbazine hydrochloride and oral prednisolone on ds 1-7; bleomycin and vincristine on d 8. eBEACOPP: every 21 ds for up to 3 cycles doxorubicin hydrochloride and cyclophosphamide on d 1; etoposide on ds 1-3; oral procarbazine hydrochloride on ds 1-7; oral prednisolone on ds 1-14; and bleomycin and vincristine on day 8. Drugs were given intravenously if not marked otherwise. Other treatment: RT consolidation (2.6% of pts) G-CSF subcutaneously (to pts on the BEACOPP14 arm)	PFS at 3 yrs*§ OS rate Short- and long-term AE (lung diffusion capacity for carbon monoxide, and fertility [see Anderson et al. [93] among companion studies]
GITIL/FIL HD0607 Gallamini, 2020 [60] (final analysis); Gallamini, 2018 [48] (Long-term results); Years of accrual: June 2008 to June 2014 Country: Italy, Israel Funding: Associazione Italiana per la Ricerca sul Cancro, Associazione Italiana Lotta	Objectives: 1) To test a risk-adapted strategy on PET positive pts after 2 ABVD cycles 2) To investigate the role of consolidation RT in PET2 pts treated with ABVD up-front who presented with a baseline mass ≥5 cm Design: RCT, phase II, 2-stage design	N=782 pts receiving front-line therapy for advanced stage HL. 150 PET positive, and 630 PET negative. Expected difference: 25% difference: rescue rate of 75% with R-eBEACOPP vs. 50% eBEACOPP Age (Median, range): 31 yrs (14-60 yrs) Stage: IIB: 36% III: 32% IV: 32%	2×ABVD then PET: PET-2-positive pts (n=76) 4×eBEACOPP+4×bBEACOPP vs. (n=72) Rituximab (375 mg/m^2) + 4×eBEACOPP+4×bBEACOPP PET-2-negative pts: (n=630) 6×ABVD PET-2-negative pts with a baseline mass $\geq$5 cm (n=296): RT (30 Gy) consolidation vs. no further treatment	3-yrs PFS* OS

Study Name Author(s) [Reference], Years of Accrual Country Funding Notes	Study Objectives, Design, Median Follow-up	Number of Patients included / analyzed Population Age Stage Required sample size	Intervention and comparison groups, Schedule, Other treatment	Outcomes Measured
alla Leucemia Sezione di Bergamo, Cassa di Risparmio di Cuneo **Note**: this study is reported in question 1 for Objective 1 [48] and in question 3 for Objective 2 [60] of the paper.	Follow-up: 70.8 mos in final analysis; 43.2 mos in long-term results	Gender: male 49% Sample size A minimum of 155 pts was needed to obtain a power of 90%, an alpha error of 5%, and an expectation to cure of ~85%. PET-2-positive pts: 65 pts per arm were needed to assess a benefit R-eBEACOPP with and an expected rescue rate of 75% after R+eBEACOPP and of 50% after eBEACOPP, an alpha-error of 5%, and a power of 80%. The results of the first interim analysis showed that 19% of patients had positive PET2 scans, so 684 patients had to be enrolled to reach the required sample size of 130 randomly assigned patients with PET2-positive scans	Treatment duration: nr Other treatment: nr	
Comparison 4. ABVD vs. modified ABVD				
ECHELON-1 Connors, 2018 [9] (2-yr data interim analysis); Straus, 2020 [86] (3-yr update); Radford, 2020 [87] (abs, 4-yr update); Straus, 2021 [88] (5-yr update – post-hoc analysis of PFS); Ansell, 2022 [89] Years of accrual: 2012-2017 Country: Americas, Europe, Asia Funding: Industry; Connors, Straus, and Radford received funding from sponsor	Objectives: to compare first-line therapy with A-AVD vs. ABVD in pts with stage III or IV classic HL Design: RCT phase 3 open-label, multicenter Follow-up: 24.6 mos (range 0-49); [9]; 37 mos; [86]; 48.4 mos [87]; and 60.9 mos [88] (this analysis was not pre-specified in the protocol); 73 mos Ansell, 2022 [89]	N=1334 (ITT analysis) pts with advanced-stage HL (stages III or IV) who had not been previously treated. 1321 in safety analysis Expected difference: 8% improvement in modified PFS at 2 yrs: 81% in the A-AVD vs. 73% in the ABVD groups. Age: median (range): 36 yrs (18-83 yrs) Stage: II: <1% III: 36% IV: 64% Missing: <1% Gender: 58% male Sample size: Assuming a 2-year modified PFS (mPFS) of 81% for patients in the A+AVD group	6×ABVD (n=670) vs. 6×A-AVD (n=664) Treatment duration: Given IV on ds 1 and 15 of each 28-d cycle for up to 6 cycles. Other treatment: nr Pts who failed to achieve CR after first-line therapy received salvage treatment. Pts who had grade ≥3 neutropenia received G-CSF primary prophylaxis. 25% of pts treated with A-AVD received primary prophylaxis with G-CSF – recommendation of the data and safety monitoring committee	Modified PFS*† [9] PFS after the primary analysis as a pre-specified exploratory end point per investigator in the ITT population (defined as time from randomization to first occurrence of disease progression or death from any cause or noncomplete response after completion of frontline therapy followed by subsequent anticancer therapy) CR OS Number of participants with

Study Name Author(s) [Reference], Years of Accrual Country Funding Notes	Study Objectives, Design, Median Follow-up	Number of Patients included / analyzed Population Age Stage Required sample size	Intervention and comparison groups, Schedule, Other treatment	Outcomes Measured
		and 73% for patients in the ABVD group, 260 mPFS events and approximately 1240 pts were needed to detect a HR of 0.67 for disease progression, death or modified PFS with a one-sided significance level of 0.025 and power of 90%		≥Treatment emergent AE, or serious AE Number of participants with abnormal Clinical Laboratory Values EFS DFS ORR DOR (duration of response) DOCR (duration of complete remission Percentage of pts not in CR PET negativity rate Quality of life (EORTC QLQ-C30)‡ A-AVD Cmax A-AVD AUCinf Note: secondary outcomes from https://clinicaltrials.gov/ct2/show/NCT01712490
Comparison 5. Modified ABVD vs. another modified ABVD				
No fully published trial final results are available for this comparison as of February 16, 2024. See Table 7-7A in Appendix 7 for abs of interim analyses of ongoing trials in this category.				
Comparison 6. BEACOPP vs. modified BEACOPP				
No fully published trial final results are available for this comparison as of February 16, 2024. See Table 7-7A in Appendix 7 for abs of interim analyses of ongoing trials in this category.				

Note: All outcomes measured are reported in this table; results for critical or important outcomes are reported in the results table and considered in analysis; the other outcomes are reported here only for information purposes.

* Primary outcomes.

† Connors, et al. [9] reported results on modified PFS. PFS after the primary analysis as a pre-specified exploratory end point per investigator in the ITT population (defined as time from randomization to first occurrence of disease progression or death from any cause) was used in follow-up publications. We did not report results for modified PFS because it was not one of the outcomes considered important or critical. Modified PFS (mPFS) was defined as the time from the date of randomization to the date of the first of documentation of progressive disease (PD), death due to any cause, or for participants who were confirmed non complete responders per IRF, receipt of subsequent anticancer therapy for HL after completion of frontline therapy. PD was defined as any new lesion or increase by greater than or equal to (≥) 50 percent (%) of previously involved sites from nadir. Frontline therapy is the part of standard set of treatments.

a This part of the study was closed early at a pre-planned interim analysis: no difference in PFS was shown, and futility was concluded [50]
b ITT was defined in this study as the set of all pts except those with disconfirmed diagnosis of HL, registration errors, or withdrawal of trial consent. Per protocol analysis (PET-2 negative pts) contains all randomized ITT patients without severe protocol deviation, having complete therapy documentation or progressive disease or death during therapy.
c A protocol amendment in June 2011 introduced a reduction of the Standard to 6×eBEACOPP. From then on PET-2 positive pts were no longer randomized and all received 6×eBEACOPP, while PET-2 negative pts were randomized to 6×eBEACOPP (New Standard) vs. 4×eBEACOPP.

Abbreviations: A-AVD=Brentuximab vedotin plus doxorubicin, vinblastine, and dacarbazine; ABVD=doxorubicin, bleomycin, vinblastine and dacarbazine; AE=Adverse events; AUCinf=Area Under the Plasma Concentration-time Curve From Time 0 to Infinity for Brentuximab Vedotin; AVD=doxorubicin, vinblastine, and dacarbazine; bBEACOPP=baseline BEACOPP; BEACOPP=bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone; BEAM=Carmustine, Etoposide, Cytarabine, Melphalan; CEC=carboplatin, etoposide, and cyclophosphamide; Cmax=maximum serum concentration; COPPEBVCAD-CEC=cyclo- phosphamide, lomustine, vindesine, melphalan, prednisone, epidoxirubicin, vincristine, procarbazine, vinblastine, and bleomycin; COPP/ABVD=cyclophosphamide, vincristine, procarbazine, and prednisone alternating with doxorubicin, bleomycin, vinblastine, and dacarbazine; CR=Complete remission; CR(u)=Complete remission (unconfirmed); DFS=Disease-free survival; DOCR=Duration of complete remission; DOR=Duration of response; d(s)=day(s); eBEACOPP=escalated BEACOPP; $eBEACOPP_4$+$bBEACOPP_4$ = 4 cycles of eBEACOPP + 4 cycles of bBEACCOPP; EFS=Event-free Survival; EORTC QLQ-C30=European Organization for the Research and Treatment of Cancer Quality of Life Questionnaire C-30; FFFP=Freedom from first progression; FFS=Failure-free survival; FFSP=Freedom from second progression; FFTF=Freedom from treatment failure; G-CSF=granulocyte-colony stimulating factor; Gy=Gray; HL=Hodgkin lymphoma; HR=Hazard ratio; IPS=International Prognostic Score; IQR=Interquartile range; ITT=Intention-to-treat; IV=Intravenous; mos=months; mPFS=modified PFS; nr=not reported; ORR=Overall response rate; OS=Overall Survival; PET=Positron emission tomography; PFS=Progression-free survival; pts=patients; QoL=Quality of Life; RCT=Randomized controlled trial; R-eBEACOPP=eBEACOPP with rituximab; RT=Radiotherapy; SPM=Secondary primary malignancies; wks=weeks; yrs=years.

Table 5. Question 3: General characteristics of included trials of consolidation radiotherapy.

Study Name Author [Reference], Years of Accrual Country Funding Notes	Study Objectives, Design, Median Follow-up	Number of Patients included / analyzed Population Age Stage Required sample size	Intervention and comparison groups, Schedule, Other treatment	Outcomes measured[a]
Comparison 1. Consolidation RT vs. observation				
HD0801 FIL Ricardi, 2021 [75]; Years of accrual: September 2008 to April 2013 Country: Italy Funding: Fondazione Italiana Linfomi **Note**: this study is underpowered as sample required was 60 per group to have 80% power	Objectives: To examine whether PET2-negative pts could benefit from RT consolidation for areas of bulky disease, provided they maintained negativity for the entire ABVD treatment Design: RCT phase 3 multicentre Follow-up: 71 months	N=116 pts; 116 in ITT analysis, 107 in per-protocol analysis. Pts were PET-2-negative and remained negative at the end of chemotherapy treatment (6 ABVD cycles) and with ≥1 site of bulky disease at baseline (diameter ≥5 cm) Age (median): 31 yrs Stage: IIB: 29% III: 35% IV: 35% Gender: 55% male	n=58 consolidation RT 30 Gy in 2 Gy fractions vs. n=58 observation Treatment duration: NA Other treatment: 6×ABVD	EFS* (events were relapse, second cancers, and death) PFS OS
GITIL/FIL HD0607 Gallamini, 2020 [60] (final analysis); Gallamini, 2018 [48] (Long-term results); Years of accrual: June 2008 to June 2014 Country: Italy, Israel Funding: Associazione Italiana per la Ricerca sul Cancro, Associatione Italiana Lotta alla Leucemia Sezione di Bergamo, Cassa di Risparmio di Cuneo **Note:** This study is reported in Question 3 for objective 2 and in Question 1 for objective 1 #743 belongs to question 1	Objectives: 1) To test a risk-adapted strategy on PET positive pts after 2 ABVD cycles 2) To investigate the role of consolidation RT in PET2 pts treated with ABVD up-front who presented with a baseline mass ≥5 cm Design: RCT, phase II, 2-stage design Follow-up: 70.8 mos in final analysis; 43.2 mos in long-term results	N=782 pts receiving front-line therapy for advanced stage HL Age (Median): 31 yrs Stage: IIB: 36% III: 32% IV: 32% Gender: male 49%	PET-2-positive pts (n=76) 4×eBEACOPP+4×bBEACOPP vs. (n=72) R (375 mg/m^2) - 4×eBEACOPP+4×bBEACOPP PET-2-negative pts with a baseline mass ≥5 cm (n=296): RT (30 Gy) consolidation (n=148) vs. no further treatment (n=148). Treatment duration: 1.6 to1.8 Gy in 5 fractions a week for a total dose of 30 Gy Other treatment: 6×ABVD	3-yrs PFS* OS

Study Name Author [Reference], Years of Accrual Country Funding Notes	Study Objectives, Design, Median Follow-up	Number of Patients included / analyzed Population Age Stage Required sample size	Intervention and comparison groups, Schedule, Other treatment	Outcomes measured[a]
EORTC trial No. 20884 Aleman, 2003 [121], 2007 [76] Years of accrual: September 1989 – April 2000 Country: Multiple European countries (42 centres) Funding: Ank van Vlissingen Foundation, Amsterdam, The Netherlands	Objectives: To describe the role of RT in pts with advanced HL who were in PR after mechlorethamine, vincristine, procarbazine, prednisone/doxorubicin, bleomycin, vinblastine (MOPP-ABV) hybrid chemotherapy. Design RCT **Follow-up**: 79 mos, and 93.7 mos	N=739 Previously untreated clinical Stage III or IV HL; 15–70 yrs of age. Pts had to be in CR after 4 and 6 cycles. Pts in PR after 4 cycles and in CR after 6 cycles were given 2 additional cycles of chemotherapy Pts who remained in PR after 6 cycles were not randomized and were all treated with IF-RT. Age: median 33 yrs Stage: IIIA: 26% IIIB: 32% IVA:13% IVB: 29%	Pts in complete remission (CR) after 4 cycles of chemotherapy: No further treatment vs. IF-RT Pts in partial remission (PR) after 4 cycles received 2 more cycles of chemotherapy then randomized as above. Pts who remained in PR after 6 cycles were treated with IF-RT. IF-RT: All involved area were irradiated. Pts in CR: 24 Gy to nodal areas Pts in PR: 30 Gy in 1.5-2.0 Gy fractions (with a boost of 4 to 10 Gy when indicated). IF-RT started 6-8 wks after the first d of chemotherapy and was administered in one to three courses. Other treatment: MOPP-ABV hybrid, 6 to 8 cycles of 28 ds each	RFS at 3 yrs * EFS OS SPM AE
HD12 Borchmann, 2011 [59] von Tresckow, 2018 [77] (97 mos follow-up) Years of accrual: January 1999 to January 2003 Country: Germany, Switzerland, Austria, the Netherlands, and the Czech Republic Funding: Deutsche Krebshilfe, the Swiss Federal Government, and the Bundesministerium für Bildung und Forschung, Germany	Objectives: To test a regimen that could decrease the toxicity associated with 8 cycles of eBEACOPP + RT while maintaining efficacy, specifically: 1) Chemotherapy: to test the effect of replacing the last 4×eBEACOPP with 4 cycles of baseline BEACOPP; (see question 1) 2) Radiotherapy: to evaluate the impact of using RT for consolidation in pts responding to chemotherapy who had initial bulky disease (≥5 cm) or residual disease (1.5 cm)	N=1670, 1574 in analysis for chemotherapy comparison; 1520 in analysis for RT comparison Newly diagnosed pts with histology-proven HL in stages IIB with a large mediastinal mass (≥ 1/3 maximum thoracic diameter) or extra nodal lesions, and stages III and IV. Number of events Age (Mean): 8×eBEACOPP+RT: 35.6 yrs 8×eBEACOPP no RT: 35.3 yrs 4×eBEACOPP +4×bBEACOPP + RT: 35.7 yrs 4×eBEACOPP +4×bBEACOPP no RT: 35.4 yrs	A) 8×eBEACOPP+RT in responding pts with initial bulk or residual tumour (n=392) vs. B) 8×eBEACOPP no RT (n=395) vs. C) 4×eBEACOPP +4×bBEACOPP + RT (n=393) vs. D) 4×eBEACOPP +4×bBEACOPP no RT (n=394) Chemotherapy comparison: 8×eBEACOPP ± RT (n=787) vs. 4eBEACOPP + 4bBEACOPP± RT (n=787) Radiotherapy comparison: Chemotherapy + RT (n=755) vs. Chemotherapy – RT (n=765)	FFTF* PFS OS AE

Study Name Author [Reference], Years of Accrual Country Funding Notes	Study Objectives, Design, Median Follow-up	Number of Patients included / analyzed Population Age Stage Required sample size	Intervention and comparison groups, Schedule, Other treatment	Outcomes measured[a]
Note: This study is reported in question 3 for objective 2 and in question 1 for objective 1	Design: RCT 2×2 factorial, noninferiority design Follow-up: 78 mos	Stage: Chemotherapy comparison: *8 eBEACOPP ± RT* IIB: 16.5% IIIA: 21.6% IIIB: 26.4% IVA: 8.9% IVB: 26.6% *4+4 ± RT* IIB: 15.9% IIIA: 22.5% IIIB: 27.2% IVA: 10.0% IVB: 24.3% Radiotherapy comparison: *Chemotherapy +RT* IIB: 17.7% IIIA: 21.1% IIIB: 25.6% IVA: 10.5% IVB: 25.0% *Chemotherapy −RT* IIB: 15.8% IIIA: 23.1% IIIB: 27.1% IVA: 8.8% IVB: 25.2%		
Comparison 2. RT vs. Chemotherapy				
ECOG E1476 Wiernik, 2009 [78] Years of accrual: *nr* Country: US Funding: National Cancer Institute, National Institutes of Health, and Department of Health and Human Services	Objectives: to compare the effectiveness of ABVD vs. RT consolidation Design: RCT Follow-up: 255.8 mos	N=301 pts with HL stage III or IV who were in CR or PR after 6 cycles of first-line MOPP-Bleomycin entered in study, 253 analyzed Age (median): 34 yrs Stage: III: 50.2% IV: 49.0% Gender: 71.5% male	RT vs. 3×ABVD Treatment dose/duration: ABVD was given every 35 ds for 3 cycles RT: mid-plane total dose was 15-200 Gy given at a rate of 15-20 Gy per fraction five fractions per wk. Pts > 60 yrs of age were given 50% of the full dose of drugs on the first consolidation cycle, then full dose if tolerated. Other treatment: *nr*	CR PR OS DFI PFS

<table>
<tr><th>Study Name
Author [Reference],
Years of Accrual
Country
Funding
Notes</th><th>Study Objectives, Design, Median Follow-up</th><th>Number of Patients included / analyzed
Population
Age
Stage
Required sample size</th><th>Intervention and comparison groups, Schedule, Other treatment</th><th>Outcomes measured[a]</th></tr>
<tr><td>Aviles trial

Aviles, 2000 [79]

Years of accrual:
March 1983 to December 1993

Country: Mexico

Funding: nr</td><td>Objectives:
To determine if the use of adjuvant RT to sites of initial bulky disease and adequate modern chemotherapy in pts with advanced HL could improve outcomes.

Design: RCT

Follow-up: 66 mos (range 46-108 mos)</td><td>N=110
Previously untreated Stage III or IV HL

Age (median, [range]):
Chemotherapy alone: 48.6 yrs, (23-72) yrs; Combined therapy: 45.9 yrs, (28-59 yrs)

Stage:
<table><tr><th></th><th>Chemo alone (%)</th><th>Combined therapy (%)</th></tr><tr><td>IIIB</td><td>19</td><td>18</td></tr><tr><td>IVA</td><td>7</td><td>5</td></tr><tr><td>IVB</td><td>73</td><td>75</td></tr></table>
Gender (male):
Chemo alone: 53%
Combined: 42%</td><td>N=56 Chemotherapy alone (6-mos cycles of EBVD) vs. N=54 Combined therapy (chemotherapy + RT to anatomical sites of bulky disease)

Rt dose/schedule:
tumour dose of 35 Gy delivered in 4-wks daily 1.25 fractions. Routine RT to nodal sites.

Other treatment:
EBVD</td><td>FFS
OS
AE</td></tr>
</table>

*Primary outcome

‡Change From Baseline in Patient-Reported Outcome Scores by mPFS Based on European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire-C30 (EORTC QLQ-C30) at end of treatment (approximately 1 year)
§Among pts randomly assigned to continue or stop bleomycin (ABVD vs. AVD)
a All outcomes measured are reported in this table; results for critical or important outcomes are reported in the results table and considered in analysis; the other outcomes are reported here only for information purposes.
b ITT was defined in this study as the set of all pts except those with disconfirmed diagnosis of HL, registration errors, or withdrawal of trial consent. Per protocol analysis (PET-2 negative pts) contains all randomized ITT patients without severe protocol deviation, having complete therapy documentation or progressive disease or death during therapy.
c A protocol amendment in June 2011 introduced a reduction of the Standard to 6×eBEACOPP. From then on PET-2 positive pts were no longer randomized and all received 6×eBEACOPP, while PET-2 negative pts were randomized to 6×eBEACOPP (New Standard) vs. 4×eBEACOPP.
d *nr* in this publication because data not available yet

A-AVD=Brentuximab vedotin plus doxorubicin, vinblastine, and dacarbazine; ABS=abstract; ABVD=doxorubicin, bleomycin, vinblastine, and dacarbazine; AE=adverse events; AUCinf=Area Under the Plasma Concentration-time Curve From Time 0 to Infinity for Brentuximab Vedotin; bBEACOPP=baseline BEACOPP; BEACOPP=bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, prednisone; $BEACOPP_{14}$=BEACOPP for 14 days; BrECADD=brentuximab vedotin, etoposide, cyclophosphamide, doxorubicin, dacarbazine, dexamethasone; BrECAPP=brentuximab vedotin, etoposide, cyclophosphamide, doxorubicin, procarbazine, prednisone; BV=brentuximab vedotin; cMax=Maximum Observed Serum Concentration for Brentuximab Vedotin Antibody-drug Conjugate (ADC) and Total Antibody (TAb); CMR=complete metabolic response; COPP=cyclophosphamide, vincristine, procarbazine, and prednisone; COPPEBVCAD-CEC=cyclophosphamide, lumustine, vindesine, melphalan, prednisone, epidoxirubicin, vincristine, procarbazine, vinblastine, and bleomycin; CR=complete remission; CR(u)=complete remission (unconfirmed); CT=computed tomography; DFI=Disease-free interval; DFS=disease-free survival; DOCR=duration of complete response; DOR=duration of response; d(s)=day(s); eBEACOPP=escalated BEACOPP; EBVD=epirubicin, bleomycin, vinblastine, dacarbazine; EFS=event-free survival; EORTC-C30=European Organization for Research and Treatment of Cancer Lymphoma Group quality of life questionnaire C-30; FFS=failure-free survival; FFTF=freedom from treatment

failure; G-CSF=granulocyte-colony stimulating factor; Gy=Gray; HL=Hodgkin lymphoma; IF-RT=involved field radiotherapy; ITT=intention-to-treat; IV=intravenous; MOPP-ABV=dacarbazine and mechlorethamine, vincristine, procarbazine, prednisone/doxorubicine, bleomycin, vinblastine; mos=months; NA=Not applicable; *nr*=not reported; ORR=overall response rate; OS=overall survival; PET=Positron emission tomography; PFS=progression-free survival; PR=partial remission; pts=patients; R-eBEACOPP=rituximab plus eBEACOPP; RCT=randomized controlled trial; RFS=relapse-free survival; RT=radiotherapy; SEER=Surveillance, Epidemiology and End Results; SPM=secondary primary malignancies; Stanford V=doxorubicin, vinblastine, mechlorethamine, vincristine, bleomycin, etoposide, prednisone; VEPEB=vinblastine, etoposide, bleomycin, epirubicin, and prednisone; VEPEMB=vinblastine, cyclophosphamide, procarbazine, prednisone, etoposide, mitoxantrone, bleomycin; vs=versus; wks=weeks; yrs=years

Time: 5 yr for OS, and 20 yrs for AE (second cancers, including solid tumours and AML and MDS); Clinical threshold: >10% difference in PFS would be a clinically significant difference

Outcomes for Questions 1 and 3

Critical outcomes for Questions 1 and 3 are overall OS, recurrence, and treatment-related death. Important outcomes are pulmonary toxicity, secondary primary malignancies (SPM) and infertility for Question 1, and SPM for Question 3 (see Table 1, and Appendix 6 for a more in-depth presentation of the importance of outcomes).

In the following paragraphs, we describe the results, question by question, comparison by comparison, and outcome by outcome, including the certainty of the evidence considered in context.

Additional details for the results of studies included for Questions 1 and 3, are presented in Tables 1-7A and 2-7A, Appendix 7. The results of the risk of bias assessment performed with the Cochrane ROB2 [36] is reported in Appendix 5.

Question 1

The studies that included the younger (i.e., <60 years of age) population can be grouped in six comparisons: 1) ABVD versus BEACOPP; 2) Higher- versus lower-intensity/dosage BEACOPP; 3) PET-adapted strategies, either PET-directed versus non PET-directed or standard treatment versus deescalated treatment based on PET-2 results; 4) ABVD versus modified ABVD; 5) Modified ABVD versus another modified ABVD; and 6) BEACOPP versus modified BEACOPP. Within each comparison, the authors of the included studies examined various doses and schedules of treatment. We present these details in the following paragraphs.

Comparison 1: ABVD versus BEACOPP

Table 6 presents the relative and absolute effect and the certainty of the evidence for this comparison.

Five unique RCTs of patients with advanced-stage HL, the EORTC 20012 [52], the LYSA H34 [53], the HD9 [54,64,81], the HD2000 [55], and the GITIL [57], one companion publication [103], and three pooled analyses of the GHSG [105,107,108] reported on this comparison. Included studies reported various follow-up times; for the purpose of this evidence summary, we considered follow-up times that are as consistent as possible among studies. Four different doses and schedules of both ABVD and BEACOPP were compared:

A) 6× or 8×ABVD vs. 4×eBEACOPP+4×bBEACOPP: three studies, the EORTC 20012 [52], the LYSA H34 [53], and the GITIL trial [57] at four- and seven-year follow-up.
B) 8×COPP/ABVD vs. 8×bBEACOPP: one study, the HD9 [54,64,81] at five-year follow-up;
C) 8 COPP/ABVD vs. 8eBEACOPP: one study, the HD9 at five-year follow-up [54,64,81];
D) 6ABVD vs. 4×eBEACOPP + 2×bBEACOPP: one study: the HD2000 [55] at three and half-year (41 months) follow-up.

The primary outcomes were PFS, and EFS in the EORTC 20012 trial [52], EFS in the LYSA H34 trial [53], FFTF in the HD9 [54], FFS in the HD2000 trial [55], and freedom from first progression (FFFP) in the GITIL trial [57]. See Table 4 for general characteristics, and Table 1-7A in Appendix 7 for the results of these studies.

Among the ancillary studies, Engert et al. [103] examined schedule adherence, acute hematological toxicity and the need for supportive treatment of patients included in the HD9 trial [54,64]. See Table 3-7A in Appendix 7 for more details on these studies.

Table 6 shows the relative and absolute values and the certainty of the evidence for the studies that reported on critical and important outcomes for Comparison 1.

Figures 1A to 1F show the results of meta-analyses if performed, the forest plots as well as the results of the studies that tested slightly different treatment doses and schedules in this comparison for critical and important outcomes.

Critical outcomes

OS

Five trials reported on OS: the EORTC 20012 [52], the LYSA H34 [53], the GITIL [57], the HD9 [54,64,81], and the HD2000 [55]. In the following paragraphs we present the results for this outcome grouped by similar doses and schedules.

A) 6 or 8×ABVD vs. 4×eBEACOPP+4×bBEACOPP at four-year follow-up (three studies [52,53,57]).

We pooled in meta-analysis three studies [52,53,57] as we considered them clinically homogeneous, although the patients in the GITIL [57] trial's control group received eight cycles of ABVD as opposed to six cycles as in the other two studies [52,53], and they were followed up for seven years as opposed to four years as in the other two studies [52,53].

No statistically significant difference in OS was shown between six or eight cycles of ABVD compared with four cycles of escalated (eBEACOPP) plus four cycles of baseline BEACOPP (bBEACOPP) (HR 0.69; 95% confidence interval [CI], 0.46 to 1.04), I^2=0%, (Figure 1A). We conducted a sensitivity analysis, by pooling only the EORTC 20012 [52] and the LYSA H34 [53] trials. Similarly to the main analysis, no difference in OS was found (HR, 0.54; 95% CI, 0.19 to 1.58).

We considered the certainty of the evidence for OS in this dose and schedule group as moderate, because in spite of a low risk of bias for OS in the included studies, of the large number of patients included, and consistency among most studies with different doses and schedules in Comparison 1, the 95% of the confidence interval of the point estimate crossed the threshold of non-significance and we cannot tell the direction of the effect (Figure 1A, Table 4-10).

B) 8×COPP/ABVD vs. 8×bBEACOPP at five-year follow-up (one study, HD9 [81,120]). At 48 months of follow-up, the HD9 trial [81] reported no statistically significant difference in OS rate when bBEACOPP was compared with COPP/ABVD (p=0.16). See Table 1-7A in Appendix 7 for more details on this trial.

This unblinded study was at high risk of bias because of deviations from the intended interventions that occurred since the study was stopped early for benefit. There was a risk for imprecision as the study was stopped early for benefit, and the sample in control arm was small. We considered the certainty of the evidence low (Table 6, Figure 1A).

C) 8×COPP/ABVD vs. 8×eBEACOPP at 48 months of follow-up (one study, HD9 [81,120]).

At 48 months of follow-up, the HD9 trial [81] reported a statistically significant advantage of eBEACOPP compared with COPP/ABVD for OS (p=0.002).

This study was at high risk of bias, as described above. We also noted a potential for imprecision as the study was stopped early for benefit. Inconsistency was also a factor, as the point estimate for OS was an outlier compared with other studies in this comparison; this study used a higher BEACOPP intensity compared to the other studies in this group [52,53,55,57] (Figure 1A).

D) 6×ABVD vs. 4×eBEACOPP + 2×bBEACOPP (one study, HD2000 [55]).

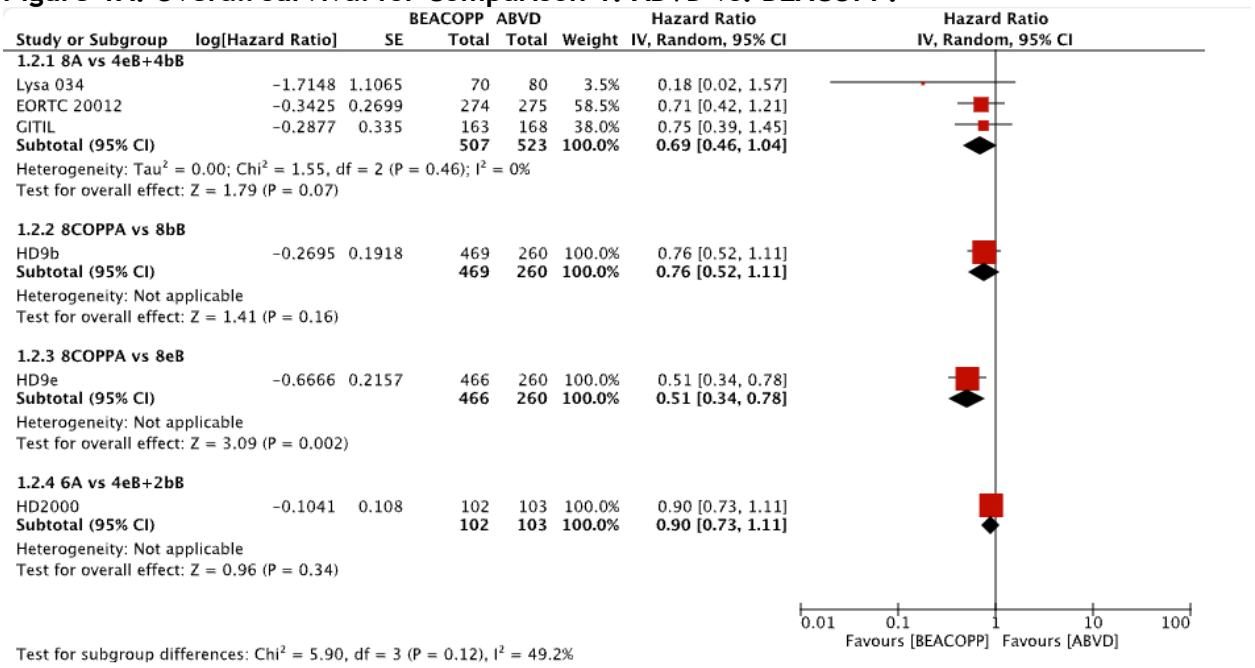
At 41 months of follow-up, the HD2000 trial reported no statistically significant difference in OS when BEACOPP was compared with ABVD (HR, 1.070; 95% CI, 0.401 to 2.850, p=0.893). See Figure 1A, and Tables 6, in this chapter, and 1-7A, in Appendix 7 for more details.

This open label study was initially designed for another purpose, and then, after it was started, the objective was changed to evaluate ABVD versus BEACOPP. It was at high risk of bias because allocation was not concealed. We downgraded the certainty of the evidence for imprecision because the sample of this study was <400 and this was just one study for this

comparison. Therefore, we considered the certainty of the evidence low for this dose and schedule group.

All studies in this Comparison 1, except the HD9 [54], were consistent in showing no statistically significant difference in OS between patients treated with ABVD regimens versus BEACOPP-based regimens. The HD9 [54], unlike the other trials, randomized patients to the highest number of cycles of the increased-dose BEACOPP (8 cycles of eBEACOPP) versus eight cycles of ABVD, and it did not use any standard-dose BEACOPP (bBEACOPP), and it was at high risk of bias. In summary, the certainty of the evidence for OS across the various schedules and doses combinations was moderate to low.

Figure 1A. Overall survival for Comparison 1: ABVD vs. BEACOPP.



Note: HD9e and HD9b represent two randomizations in the same HD9 study [54,64]. In text we considered each randomization as a separate study.
4eB=4 cycles of escalated BEACOPP; 4bB=4 cycles of baseline BEACOPP; A=ABVD; CI=confidence interval; COPPA=COPP/ABVD

Recurrence

PFS

Five studies in four publications reported results for PFS: the EORTC 20012 [52], the LYSA H34 [53], the HD9 [54,64], and the HD2000 [55,56]. We pooled in meta-analysis the EORTC 20012 [52] and the LYSA H34 [53] studies, while the other three studies (the HD9 of bBEACOPP [HD9b], the HD9 of eBEACOPP [HD9e] [54,64], and the HD2000 [55,56]) included for this outcome were clinically heterogeneous (i.e., they used slightly different treatment regimens), and we considered statistical pooling of the results inappropriate. Table 6 shows the relative and absolute values as well as the certainty of the evidence for the studies that reported on PFS. Figure 1B shows the meta-analysis and the forest plot for all the studies against the threshold of no effect.

In the following paragraphs we present the results for this outcome more details can be found in Table 1-7A, Appendix 7.

A) 6 or 8×ABVD vs. 4×eBEACOPP+4×bBEACOPP at four-year follow-up (two studies [52,53]).

The results of our meta-analysis, Figure 1B, show a benefit for the BEACOPP combination over the ABVD regimen (HR, 0.48; 95% CI, 0.27 to 0.86), I^2=39%. the certainty of the evidence for PFS was moderate to high for the EORTC 20012 [52], and LYSA H34 [53] trials because neither of these studies were blinded. The difference in point estimates fell in both studies above the minimal clinical important difference of 10% that we set up at the project plan stage.

B) 8×COPP/ABVD vs. 8×bBEACOPP at five-year follow-up (one study, HD9b [54]).

The HD9 trial [54,64] at 60-month follow-up reported a progression rate of 10% for COPP/ABVD vs. in the bBEACOPP group. The difference was not statistically significant (p values not reported). The certainty of the evidence was low because of risk of bias and imprecision, Table 6, and Figure 1B.

C) 8×COPP/ABVD vs. 8×eBEACOPP at five-year follow-up (one study, HD9 [54]).

The HD9 trial [54,64] at 60-month follow-up showed a statistically significant difference in progression rate between COPP-ABVD (10%) and eBEACOPP (2%) in favour of eBEACOPP (p=0.001). This statistically significant advantage of eBEACOPP over ABVD was maintained at 180 months of follow-up: the difference in PFS between eBEACOPP and COPP/ABVD was 17%, (HR, 0.53; 95% CI, 0.41 to 0.69), p<0.0001.

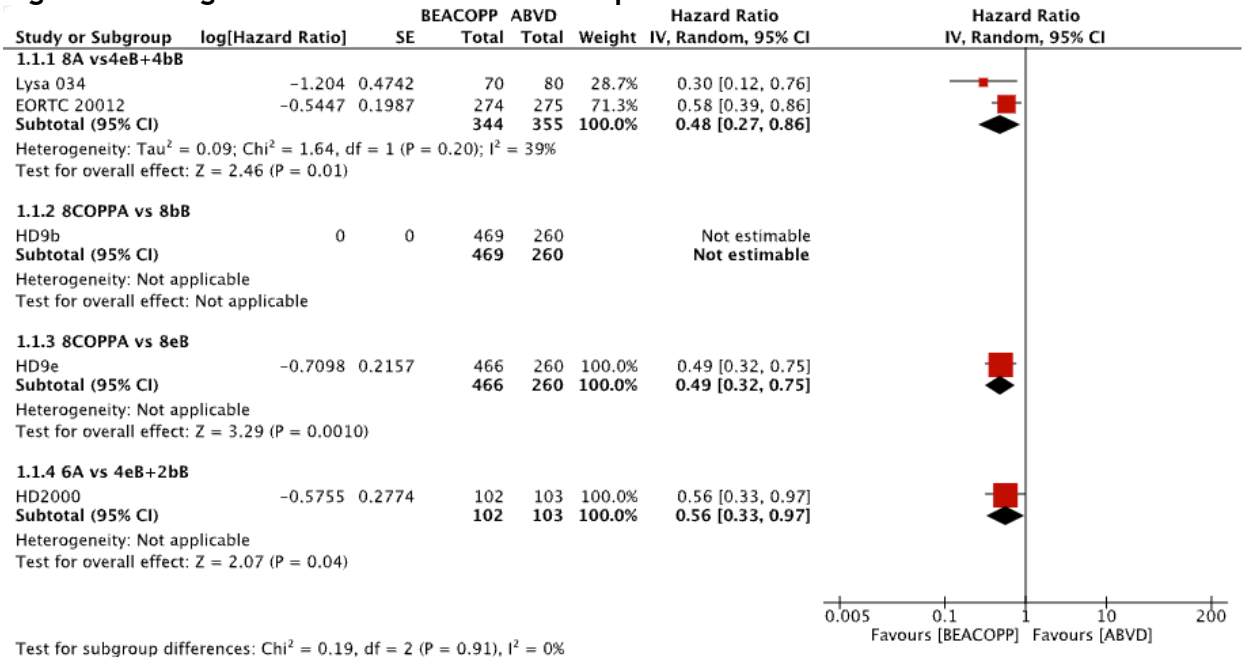
In this study patients, clinicians and outcome assessors were not blinded to patients' assignment, leading to potential risk of bias. The study was stopped early for benefit, leading to a potential overestimation of treatment effect and imprecision; therefore, we considered the certainty of the evidence to be low.

D) 6×ABVD vs. 4×eBEACOPP + 2×bBEACOPP (one study, HD2000 [55]).

The HD2000 trial [55] at 41-month follow-up showed a statistically significant difference in PFS rate in favour of the BEACOPP combination (81% for BEACOPP vs. 68% for ABVD, p=0.038). We considered the certainty of the evidence for this trial low because of potential risk of bias (i.e., allocation was not concealed, patients, clinicians and outcome assessors were not blinded to patients' assignment, and it was one small study for this schedule).

Table 6 shows the relative and absolute values as well as the certainty of the evidence for this outcome. Table 1-7A in Appendix 7 presents more details.

Figure 1B. Progression-free Survival for Comparison 1: ABVD vs. BEACOPP.



Note: HD9e and HD9b represent two randomizations in the same HD9 study [54,64]. In text we considered each randomization as a separate study.

A=ABVD; eB=escalated BEACOPP; bB=baseline BEACOPP; COPPA=COPP/ABVD; CI=Confidence Interval

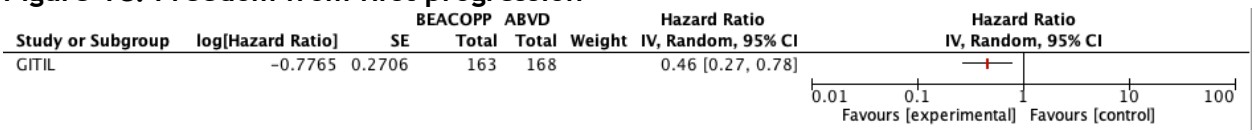
Other measures of recurrence

FFFP

6 or 8×ABVD vs. 4×eBEACOPP+4×bBEACOPP. The GITIL RCT [57] reported results for FFFP at 61 months of follow-up. FFFP rate was statistically significantly better for the BEACOPP combination than the standard ABVD (respectively 85%; 95% CI, 78 to 91 vs. 73%; 95% CI, 66 to 80 [HR, 0.46; 95% CI, 0.27to 0.78, p=0.004]). See Figure 1C, Tables 6, and 1-7A in Appendix 7 for more details.

We had some concerns for the potential risk of bias of this study for this outcome because participants, clinicians and outcome assessors were not blinded to patients' assignment. The results for this outcome were reported solely in this trial, leading to potential imprecision; therefore, we rated the certainty of the evidence as low.

Figure 1C. Freedom from first progression



CI=Confidence Interval

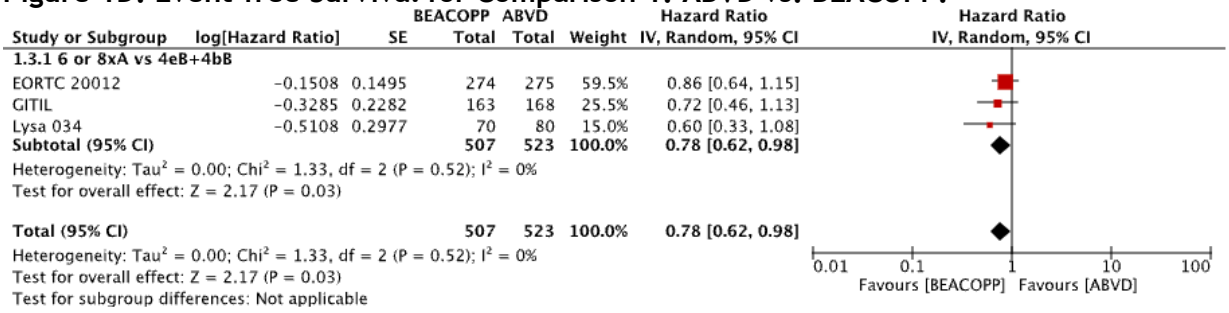
EFS

6× or 8×ABVD vs. 4×eBEACOPP+4×bBEACOPP. Three studies reported on EFS: The EORTC 20012 [52], the LYSA H34 [53], and the GITIL [57] at 48, 60, and 61 months of follow-up, respectively. The meta-analysis of these three studies showed a statistically significant difference in favour

of the BEACOPP-based regimens (HR, 0.78; 95% CI 0.62 to 0.98), I^2=0% when compared with ABVD. See Figure 1D, and Tables 6 in this section and 1-7A in Appendix 7 for more details.

We considered the certainty of the evidence for this outcome low because the included trials were at risk of bias as patients, clinicians and outcome assessors were not blinded to patients' assignment; these studies were at risk for indirectness because this outcome, which is a composite outcome, was defined slightly differently in two of the three studies [52,53], and not defined in the third one [57].

Figure 1D. Event-free Survival for Comparison 1: ABVD vs. BEACOPP.



A=ABVD; eB=escalated BEACOPP; bB=baseline BEACOPP; CI=Confidence Interval

DFS

8×ABVD vs. 4×eBEACOPP+4×bBEACOPP. The EORTC 20012 [52] reported data on DFS only for the subgroup of patients who reached complete remission (83% of the sample), and we do not report these results here because they are not generalizable to all patients. See Table 1-7A in Appendix 7 for more details.

FFTF

8×COPP/ABVD vs. 8×bBEACOPP and 8×COPP/ABVD vs. 8×eBEACOPP. The HD9 [54] reported on this measure at a median 60-month follow-up and showed that both baseline and escalated BEACOPP regimens were statistically significantly better than COPP/ABVD: COPP/ABVD 69% (95% CI, 63% to 75%) versus bBEACOPP 76% (95% CI, 72% to 80%), p=0.04, and COPP/ABVD 69% (95% CI, 63% to 7%5) versus eBEACOPP 87% (95% CI, 83% to 91%), p<0.001). See Tables 6 in this section, and 1-7A in Appendix 7 for more details.

The certainty of the evidence for this outcome was low: the HD9 trial was at high risk of bias: clinicians, patients, and outcome assessors were not blinded to patients' assignment, the study was stopped early for benefit, and there were deviations from the intended interventions because of the trial context. We considered also the study at risk for imprecision because the sample was smaller than planned.

FFS

6×ABVD vs. 4eBEACOPP + 2bBEACOPP. The HD2000 trial [55] reported on this outcome, and four cycles of escalated BEACOPP plus two cycles of baseline BEACOPP were shown to be better than ABVD: 79.6% versus 65.7%, p=0.036 (one sided log-rank test).

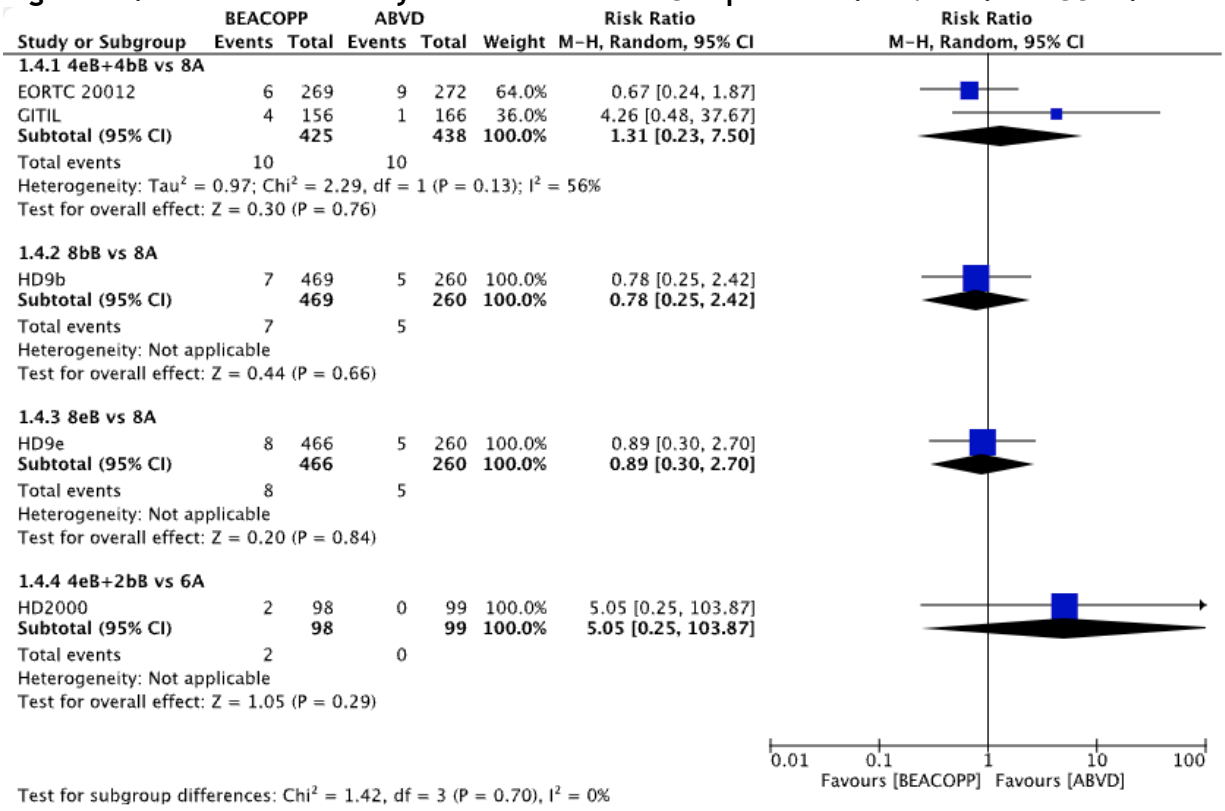
The certainty of the evidence was low for this outcome: allocation was not concealed; patients, clinicians and outcome assessors were aware of patients' assignment, and the plan of the study was different from what is reported in the results. Additionally, this study was at risk for imprecision because the sample was <400 patients, and there were no other studies reporting on this outcome.

Adverse Events
Critical Outcomes
TRD

Five studies in four publications, the EORTC20012 [52], the HD9 [54], the HD2000 [56], and the GITIL trial [57] reported on TRD. We combined in meta-analysis the results of the EORTC 20012 [52] and of the GITIL [57] trials. Event rates were generally low, and all studies reported no statistically significant difference between groups. See Figure 1E, Tables 6 in this section and 1-7A for more details.

The certainty of the evidence for these studies was very low. We had some concerns for risk of bias because this outcome was ascertained and reported by local investigators, and it was subjective; the direction of the results were inconsistent with some of the studies showing that the intervention with BEACOPP-based regimens would reduce the risk of TRD (EORTC20012 [52]: number needed to treat [NNT]=90.91; HD9 [54], NNT=250), while others showing that the BEACOPP-based regimens may harm the patients (HD2000 [56] number needed to harm [NNH]=50; GITIL [57], NNH=67). Additionally, we had concerns about imprecision; because of the very few events confidence intervals were very wide (Figure 1E, Table 6).

Figure 1E. Treatment toxicity-related death for Comparison 1: ABVD vs. BEACOPP.



Note: HD9e and HD9b represent two randomizations in the same HD9 study [54,64]. In text we considered each randomization as a separate study.

A=ABVD; eB=escalated BEACOPP; bB=baseline BEACOPP; CI=Confidence Interval

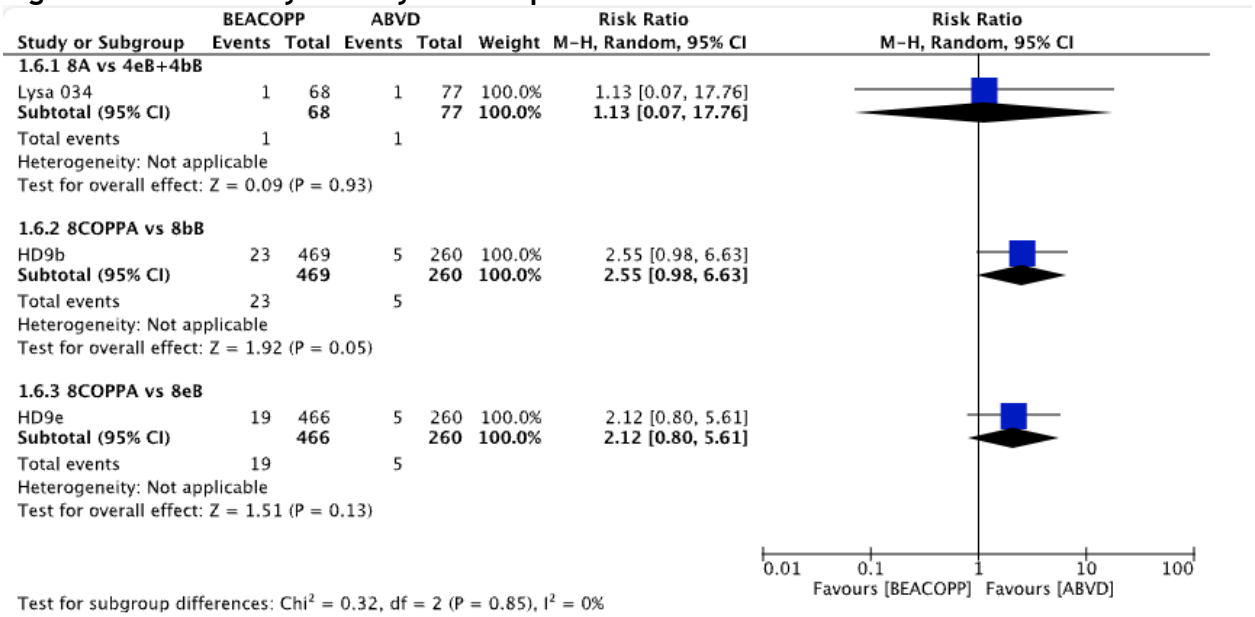
Important Outcomes

Pulmonary toxicity

The LYSA H34 reported pulmonary toxicity rate of 1.3% in the ABVD and 1.5% in the BEACOPP groups (p values were not reported). Similarly, the HD9 trial [54] reported very low rates of pulmonary toxicity (i.e., COPP/ABVD 2% vs. bBEACOPP 5% vs. eBEACOPP 4%) (p values were not reported). Other studies did not report on this outcome.

The certainty of the evidence for this outcome was very low because of risk of bias as described above, inconsistency, because the point estimates lay on different sides of the threshold in different studies, and imprecision as the number of events is very low. See Figure 1F, Tables 6 in this section, and 1-7A in Appendix 7 for more details.

Figure 1F. Pulmonary toxicity for Comparison 1: ABVD vs. BEACOPP.



Note: HD9e and HD9b represent two randomizations in the same HD9 study [54,64]. In text we considered each randomization as a separate study.

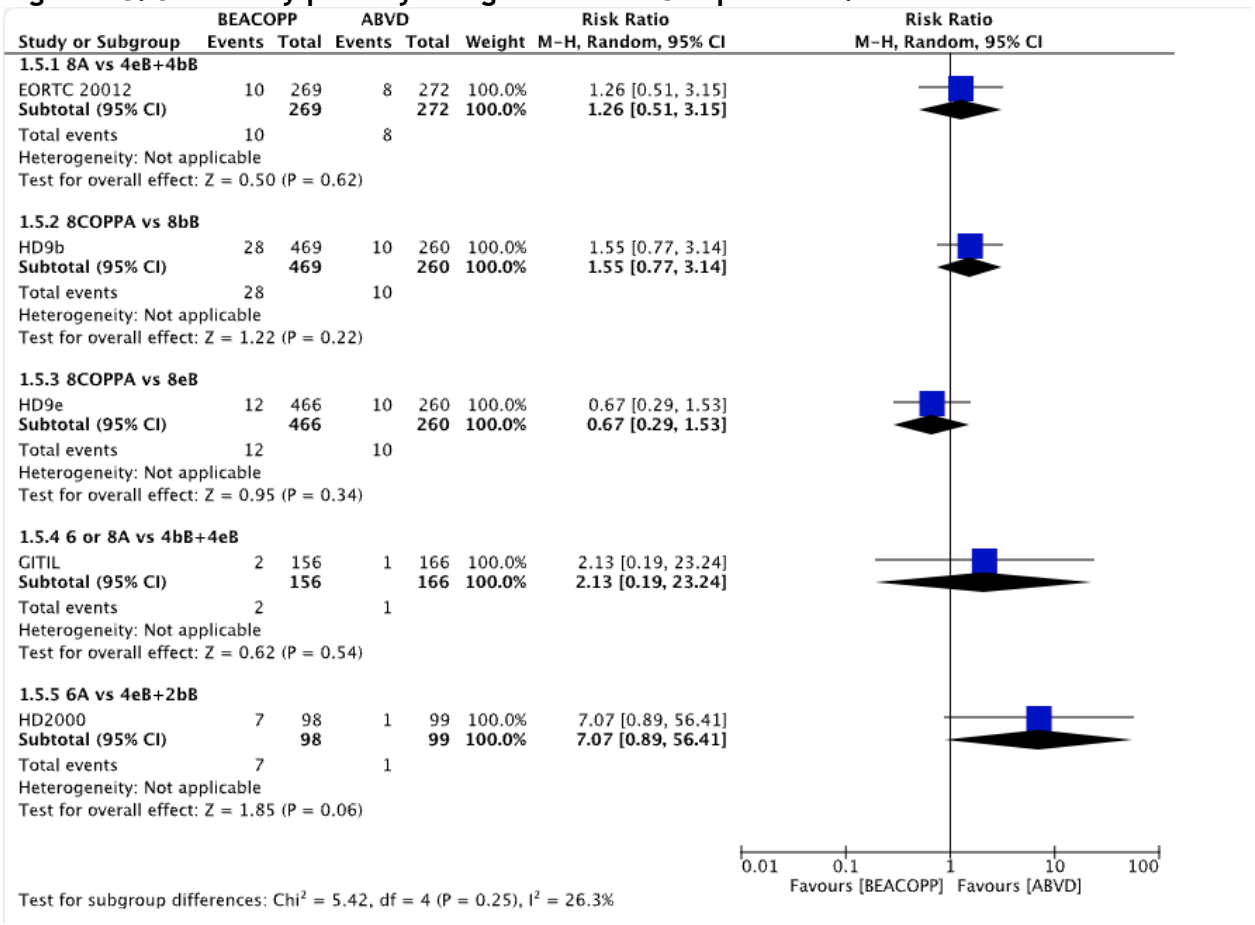
A=ABVD; eB=escalated BEACOPP; bB=baseline BEACOPP; COPPA=COPP/ABVD

SPM

Five studies in four publications reported on SPM: the EORTC20012 [52], the HD9 [54], the HD2000 [55], and the GITIL [57] trials. We did not pool the results in meta-analysis because the studies measured this outcome in different ways or were otherwise heterogeneous. The EORTC20012 [52] reported no significant difference in the cumulative incidence of SPM at four-year follow-up (Figure 1G, 1.5.1). In the HD9, [81], when the authors compared the 10-year cumulative incidence of all secondary malignancies, including solid tumours, no statistically significant difference was detected: 6.5% for eBEACOPP versus 7.9% for bBEACOPP versus 5.3% for COPP/ABVD, p=0.82 (Figure 1G). However, the cumulative incidence for AML/MDS alone was lower for patients in the COPP/ABVD group (rate: 0.4%) compared with the bBEACOPP (rate: 2.2%) and eBEACOPP group (rate: 3.2%), p=0.030. See Table 1-7A in Appendix 7 for more details. The HD2000 [55,56] reported a statistically significant higher cumulative incidence of SPM rate for the BEACOPP regimen combination compared with the ABVD and the COPPEBVCAD-CEC regimens (6.6% for BEACOPP vs. 0.9% for ABVD vs. 6% for COPPEBVCAD, p=0.02).

The certainty of the evidence for this outcome is low because of risk of bias and imprecision in the included studies.

Figure 1G. Secondary primary malignancies for Comparison 1.



Note: HD9e and HD9b represent two randomizations in the same HD9 study [54,64]. In text we considered each randomization as a separate study.

A=ABVD; eB=escalated BEACOPP; bB=baseline BEACOPP; COPPA=COPP/ABVD; CI=Confidence Interval

Infertility

None of the primary studies in this comparison reported data on infertility.

The systematic review by Amin et al. 2020 [43] reported that 50 to 57% of young male patients who received ABVD recovered sperm characteristics at 12-18 months of follow-up. Among patients who received BEACOPP, those who had two to four cycles of BEACOPP recovered sperm function at four years, while whose who received six to eight cycles did not recover spermiogenesis. No difference in infertility rates between patients who received baseline versus escalated BEACOPP was reported. Patients treated with BEACOPP were more likely to have oligospermia than those treated with ABVD (p values *nr*). In patients treated with ABVD, follicle-stimulation hormone (FSH) was raised post-ABVD compared to healthy controls (p=0.008) before returning to normal; luteinizing hormone and testosterone did not significantly change before and after treatment (p=0.203 and p=0.844, respectively). In patients treated with BEACOPP, six to eight cycles of BEACOPP were associated with a lower inhibin B/FSH ratios compared to healthy controls (impaired fertility) (p<0.001).

In summary, the certainty of the evidence for the studies included in this comparison was moderate-to-low for OS and recurrence, and very low for AEs. For OS, the direction of the results was consistent among all trials that compared lower intensity BEACOPP with ABVD.

These trials showed no between-groups statistically significant difference for OS. The HD9 [81,120] in its ABVD versus higher-intensity BEACOPP comparison showed a benefit for the BEACOPP regimen; however, the certainty of the evidence for this trial was low. Most measures of recurrence showed a benefit of BEACOPP compared with ABVD regimens; the certainty of the evidence for these comparisons was generally low. Most measures of adverse events showed no statistically significant difference between groups; the certainty of the evidence for these measures was very low.

Table 6. Comparison 1. ABVD compared with BEACOPP

Certainty assessment							№ of patients		Effect		Certainty of the evidence	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	BEACOPP	ABVD	Relative (95% CI)	Absolute (95% CI)		
Overall Survival												
A) 6 or 8×ABVD vs. 4×eBEACOPP+4×bBEACOPP (follow-up: median 4 yrs; assessed with: HR)												
3	Meta-analysis of randomised trials EORTC20012, LYSAH34, and GITIL [52,53,57]	Not serious	Not serious	Not serious	Serious[a]	None	507	523 7.0%	HR 0.69 (0.46 to 1.04) [death]	21 fewer per 1000 (from 37 fewer to 3 more)	Moderate	CRITICAL
B) 8×COPP/ABVD vs. 8×bBEACOPP (follow-up: median 5* yrs; assessed with HR)												
1	RCT HD9 [54]	Serious[b]	Not serious	Not serious	Serious[c]	None	469	260 /18.8%	HR 0.76 (0.52 to 1.11)	47 fewer per 1000 (from 85 fewer to 18 more)	Low	CRITICAL
C) 8×COPP/ABVD vs. 8×eBEACOPP (follow-up: median 5* yrs; assessed with HR)												
1	RCT HD9 [54]	Serious[b]	Serious[d]	Not serious	Serious[c]	None	466	260/ 18.8%	HR 0.51 (0.34 to 0.79)	87 fewer per 1000 (from 120 fewer to 36 fewer)	Very Low	CRITICAL
D) 6×ABVD vs. 4×eBEACOPP + 2×bBEACOPP (follow-up: median 5 yrs assessed with HR)												
1	RCT HD2000 [55]	Serious[be]	Not serious	Not serious	Serious[f]	None	102	103/ 32%	HR 0.90 (0.73 to 1.11)	8 fewer per 1000 (from 21 fewer to 8 more)	Low	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty of the evidence	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	BEACOPP	ABVD	Relative (95% CI)	Absolute (95% CI)		
PFS												
A) 6 or 8×ABVD vs. 4×eBEACOPP+4×bBEACOPP) (follow-up: median 4 yrs; assessed with: HR)												
2	Meta-analysis of RCTs EORTC20012 [52], and LYSAH34 [53]	Serious[g]	Not serious	Not serious	Not serious	None	344	355/ 80%	HR 0.48 (0.27 to 0.86)	98 more per 1000 (from 25 more to 142 more)	Moderate	CRITICAL
B) 8×COPP/ABVD vs. 8×bBEACOPP (follow-up median 5 yrs; assessed with HR)												
1	RCT HD9b Engert, 2009 [81] Diehl, 2003 [54]	Serious[h]	Not serious	Not serious	Serious[i]	None	469	260 /18.8%	HR not estimable **	Not estimable	Low	CRITICAL
C) 8×COPP/ABVD vs. 8×eBEACOPP (follow-up median 5 yrs; assessed with HR)												
1	RCT HD9e Engert, 2009 [81] Diehl, 2003 [54]	Serious[h]	Not serious	Not serious	Serious[k]	None	466	260/ 18.8%	HR 0.49 (0.32 to 0.75)	83 fewer per 1000 (from 112 fewer to 40 fewer)	Low	CRITICAL
D) 6×ABVD vs. 4×eBEACOPP+2×bBEACOPP (follow-up median 41 months assessed with HR)												
1	RCT HD2000 [55]	Serious[j]	Not serious	Not serious	Serious[l]	None	102	103/32%	HR 0.56 (0.33 to 0.97)	208 more per 1000 (from 11 more to 367 more)	Low	CRITICAL

FFFP (follow-up 7 yrs assessed with HR): 6 or 8×ABVD vs. 4×eBEACOPP+4×bBEACOPP												
1	RCT GITIL [57]	Serious[b]	Not serious	Not serious	Serious[m]	None	163	168/15%	HR 0.46 (0.27 to 0.78)	78 fewer per 1000 (from 107 fewer to 31 fewer)	Low	CRITICAL
EFS (follow-up 5 yrs assessed with HR): 6 or 8×ABVD vs. 4×eBEACOPP+4×bBEACOPP												
3	Meta-analysis of RCTs EORTC20012 [52], and LYSAH34 [53] GITIL [57]	Serious[b]	Not serious	Serious[n]	Not serious	None	507	523/30%	HR 0.78 (0.62 to 0.98)	57 fewer per 1000 (from 102 fewer to 5 fewer)	Low	CRITICAL
FFTF (follow-up 5 yrs assessed with HR: 6 or 8×COPP/ABVD vs. 8×eBEACOPP or 8×bBEACOPP												
A) **8×COPP/ABVD vs. 8×bBEACOPP**												
1	RCT HD9b [54]	Serious[b]	Not serious	Not serious	serious[c]	none	469	260/30%	HR 1.349 (1.014 to 1.794)	82 more per 1000 (from 3 more to 173 more)	Low	CRITICAL
B) **8×COPP/ABVD vs. 8×eBEACOPP**												
1	RCT HD9e [54]	Serious[b]	Not serious	Not serious	Serious[c]	None	466	260 /30%	HR 1.507 (1.180 to 1.923)	116 more per 1000 (from 44 more to 196 more)	Low	CRITICAL
FFS (follow-up 3.4 yrs assessed with HR) 6×ABVD vs. 4×eBEACOPP + 2×bBEACOPP												
1	RCT HD2000 [55]	Serious[e]	Not serious	Not serious	Serious[j]	None	102	103	HR 1.466 (1.025 to 2.096)	79 more per 1000 (from 4 more to 174 more)	Low	CRITICAL

Adverse events												
Treatment toxicity-related death												
A) 6 or 8×ABVD vs. 4×eBEACOPP+4×bBEACOPP (follow-up: median 4 yrs; assessed with: RR)												
2	RCT EORTC20012 [52], GITIL [57]	Serious[o]	Serious[q]	Not serious	Serious[p]	None	10/425 (2.4%)	10/438 (2.3%)	RR 1.31 (0.23 to 7.50)	7 more per 1000 (from 17 fewer to 147 more)	Very low	CRITICAL
B) 8×ABVD vs. 8×bBEACOPP (follow-up: median 5 yrs; assessed with: RR)												
1	RCT HD9b [54]	Serious[o]	Serious[q]	Not serious	Serious[p]	None	7/469 (1.5%)	5/260 (1.9%)	RR 0.78 (0.25 to 2.42)	4 fewer per 1000 (from 14 fewer to 27 more)	Very low	CRITICAL
C) 8×ABVD vs. 8×eBEACOPP (follow-up: median 5 yrs; assessed with: RR)												
1	RCT HD9e [54]	Serious[o]	Serious[q]	Not serious	Serious[p]	None	8/466 (1.7%)	5/260 (1.9%)	RR 0.89 (0.30 to 2.70)	2 fewer per 1000 (from 13 fewer to 33 more)	Very low	CRITICAL
D) 6×ABVD vs. 4eBEACOPP + 2bBEACOPP (follow-up: median 3.4 years; assessed with: RR)												
1	RCT HD2000 [56]	Serious[o]	Serious[q]	Not serious	Serious[p]	None	2/98 (2%)	0/99 (0.0%)	RR Undefined	Not estimable	Very low	CRITICAL
Pulmonary toxicity												
A) 6 or 8×ABVD vs. 4×eBEACOPP+4×bBEACOPP) (follow-up: median 4 yrs; assessed with: RR)												
1	RCT LYSA H34 [53]	Serious[b]	Serious[q]	Not serious	Serious[p]	None	1/68 (1.5%)	1/77 (1.3%)	RR 1.13 (0.07 to 17.76)	2 more per 1000 (from 12 fewer to 213 more)	Very low	IMPORTANT
B) 8×ABVD vs. 4×bBEACOPP (follow-up: median 5 years; assessed with: RR)												

1	RCT HD9b [54]	Serious[b]	Serious[q]	Not serious	Serious[p]	None	23/469 (5%)	5/260	RR 2.55 (0.98 to 6.63)	30 more per 1000 (from 0 fewer to 108 more)	Very low	IMPORTANT
C) 8×ABVD vs. 4×eBEACOPP (follow-up: median 5 yrs; assessed with: RR)												
1	RCT HD9e [54]	Serious[b]	Serious[q]	Not serious	Serious[p]	None	19/466 (4.1%)	5/260 (1.9%)	RR 2.12 (0.80 to 5.61)	22 more per 1000 (from 4 fewer to 89 more)	Very low	IMPORTANT
Secondary primary malignancies												
A) 6 or 8×ABVD vs. 4×eBEACOPP+4×bBEACOPP (follow-up: median 4 yrs; assessed with: RR)												
1	RCT EORTC20012 [52],	Serious[b]	Serious[q]	Not serious	Serious[a]	None	10/269	8/272	RR 1.26 (0.51 to 3.15)	8 more per 1000 (from 14 fewer to 63 more)	Very Low	IMPORTANT
B) 8×COPP/ABVD vs. 4×bBEACOPP (follow-up: median 5 yrs; assessed with: RR)												
1	RCT HD9b [54]	Serious[b]	Serious[q]	Not serious	Serious[a,c]	None	28/469 (6.0%)	10/260 (3.8%)	RR 1.55 (0.77 to 3.14)	21 more per 1000 (from 9 fewer to 82 more)	Very Low	IMPORTANT
C) 8×COPP/ABVD vs. 4×eBEACOPP (follow-up: median 5 yrs; assessed with: RR)												
1	RCT HD9e [54]	Serious[b]	Serious[q]	Not serious	Serious[a,c]	None	12/466 (2.6%)	10/260 (3.8%)	RR 0.67 (0.29 to 1.53)	13 fewer per 1000 (from 27 fewer to 20 more)	Very Low	IMPORTANT
D) 6 or 8×ABVD vs. 4×eBEACOPP+4×bBEACOPP (follow-up: median 4 yrs; assessed with: RR)												
1	RCT GITIL [57]	Serious[b]	Serious[q]	Not serious	Serious[a]	None	2/156 (1.3%)	1/166 (0.6%)	RR 2.13 (0.19 to 23.24)	7 more per 1000 (from 5 fewer to 134 more)	Very Low	IMPORTANT

E) 6×ABVD vs. 4×eBEACOPP + 2×bBEACOPP (follow-up: median 3.4 yrs; assessed with: RR)												
1	RCT HD2000 [56]	Serious[b]	Serious[q]	Not serious	Serious[a]	none	7/102 (6.9%)	1/103 (1.0%)	RR 7.07 (0.89 to 56.43)	59 more per 1000 (from 1 fewer to 538 more)	Very Low	IMPORTANT

NS=not significant; na=not applicable.
* The follow-up time of five years was chosen because it is the most similar to the follow-up of the studies that we pooled in meta-analysis.
**The comparison COPP/ABVD vs. bBEACOPP for PFS was not reported in the study at any time point.

a We downgraded one level for imprecision because the CI crosses 1.
b Clinicians, patients and outcome assessors were unblinded.
c The study was stopped early for benefit and the sample in the ABVD group was very small. The confidence interval around the point estimate crossed 1, therefore we do not know the direction of the effect.
d The point estimate is in a different direction from other studies of similar comparisons, and the confidence interval show minimal overlap
e Allocation was not concealed; patients, clinicians and outcome assessors were not blinded. The study was planned to detect myelotoxicity, then the objectives were changed to test the efficacy of the drugs combinations.
f We downgraded for imprecision because the sample was <400 and this was only one study.
g Both studies were considered at some concerns of risk of bias for PFS because patients, clinicians, and outcome assessors were not blinded.
h This was an unblinded study, and there were deviations from the intended interventions because of the trial context (the trial was stopped early for benefit)
i This study was stopped early for benefit.
k The number of patients was <300 because this study was stopped early.
j The plan of the study was different from what was reported in the results, and the CI around the point estimate was wide.
l We downgraded for imprecision because the sample was <400 and this was only one study
m The results for this outcome were reported only in this trial leading to potential imprecision
n This composite outcome was defined slightly differently in two of the studies [52,53], and it was not defined in the third trial [57].
o This outcome was local investigator reported and subjective, and we rated it at some concerns for risk of bias.
p Very few events led to wide confidence intervals, therefore we considered this result imprecise.
q The point estimates lay on different sides of the threshold in studies included for this outcome

ABVD=doxorubicin, bleomycin, vinblastine, and dacarbazine; bBEACOPP=baseline BEACOPP; BEACOPP=bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, prednisone; CI=confidence interval; COPP/ABVD=cyclophosphamide, vincristine, procarbazine, and prednisone alternating with doxorubicin, bleomycin, vinblastine, and dacarbazine; eBEACOPP=escalated BEACOPP; EFS=Event-free survival; FFFP=Freedom from first progression; FFS=Failure-free survival; FFTF – Freedom from treatment failure; HR=Hazard ratio; PFS=Progression-free survival; RCT=randomized controlled trial; RR=Relative risk; yrs=years.

Comparison 2: Higher- vs. lower-intensity/dosage BEACOPP

Table 7 presents the relative and absolute effects and the certainty of the evidence for this comparison.

The publications of three unique noninferiority RCTs were included for this comparison: The HD15 [58,82], the HD12 [59,77] for its chemotherapy comparison, and the HD9 [54,64,81] for its secondary objective (i.e., comparing bBEACOPP with eBEACOPP). These studies reported various follow-up periods; for the purpose of this evidence summary, we considered follow-up times that were as consistent as possible among studies. Five different doses and schedules of both higher and lower intensities/dosages were compared in the studies included for this comparison:

A) 8×eBEACOPP vs. 6×eBEACOPP, in the HD15 trial [58,82]
B) 8×eBEACOPP vs. 8×BEACOPP$_{14}$, in the HD15 trial [58,82]
C) 6×eBEACOPP vs. 8×BEACOPP$_{14}$, in the HD15 trial [58,82]
D) 8×eBEACOPP vs. 4×eBEACOPP + 4×bBEACOPP, in the HD12 trial [59,77]
E) 8×eBEACOPP vs. 8×bBEACOPP, for objective 2 of the HD9 trial [54,64,81].

The primary outcome of the three studies was FFTF. We considered these studies clinically heterogeneous; therefore, we did not pool the results in meta-analysis.

Among the companion publications, Engel et al. [103] reported the TRD of patients treated in the bBEACOPP versus eBEACOPP groups for Objective 2 in the HD9 [54,64,81].

Behringer et al. [96], an ancillary study of the HD15 [58], analyzed the gonadal function of 649 included both male and female survivors. However, this study does not report comparative data among the drug combinations of interest to this report, and we will not discuss it any further.

Kreissl et al. [97], in a companion of the HD15 trial [58], described the health-related quality of life (as measured with the EORTC QLQ-C30) at up to five-year follow-up, and analyzed the influence of patient, lymphoma, and treatment characteristics. However, quality of life was not one of the critical or important outcomes for the population of younger patients focus of our Question 1, and we will not discuss this study any further. For more details on these studies, see Table 3-7A in Appendix 7.

Among the pooled analyses Borchmann et al. [104] analyzed the incidence, risk factors and timing of osteonecrosis in patients included in studies HD12 [59], HD15 [58], and HD18 [16] who were treated with various doses and schedules of BEACOPP. However, osteonecrosis was not one of the critical or important outcomes for this question, and, therefore, this study will not be discussed any further. Haverkamp et al. [106] analyzed the impact of bleomycin and vincristine dose reductions in case of AEs in patients included in studies HD12 [59] and HD15 [58]. This study did not compare the drug combinations of interest; therefore, it will not be discussed any further. For more details on these studies, see Table 4-7A in Appendix 7.

Critical outcomes

OS

The HD15 trial [58], at 48 months follow-up, reported that OS was statistically significantly higher in the lower intensity/dosage group (6×eBEACOPP) than in the higher intensity/dosage (8×eBEACOPP) group (survival rates: 95.3% [97.5% CI, 93.4 to 97.2] vs. 91.9% [97.5% CI, 89.4 to 94.4], HR, 0.60 [97.5% CI, 0.36 to 0.98; calculated 95% CI, 0.39 to 0.92], p=0.019), while these authors did not detect any statistically significant difference between the 8×eBEACOPP and the 8×BEACOPP$_{14}$ groups (HR, 0.68 [97.5% CI, 0.42 to 1.10; calculated 95% CI, 0.42 to 1.10], p=0.07), and between the 6×eBEACOPP and the 8×BEACOPP$_{14}$ (HR, 0.88 [97.5% CI, 0.51 to 1.51; calculated 95% CI, 0.63 to 1.13], p=0.6) groups (Figure 2A). At 102 months of

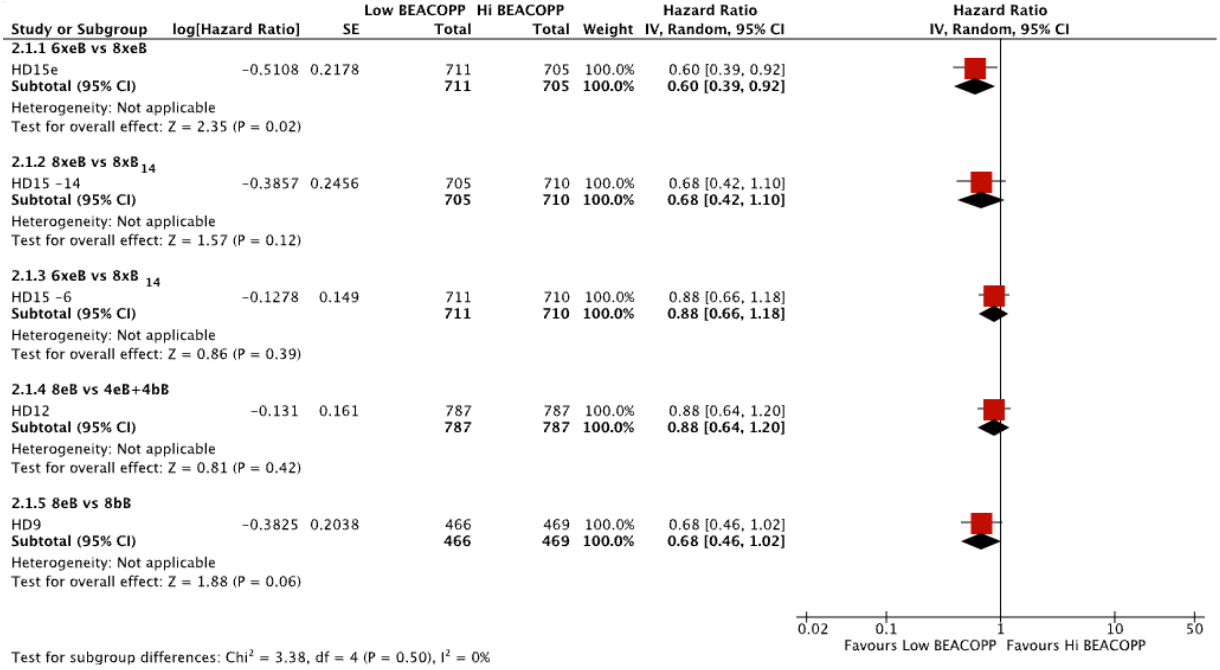
follow-up [82] these results remained unchanged: The 6×eBEACOPP versus 8×eBEACOPP comparison remained statistically significant: 6×eBEACOPP (90.4% [97.5% CI, 87.5 to 93.2]) versus 8×eBEACOPP, (87.6% [97.5% CI, 84.6 to 90.6]), HR, 0.7 (97.5% CI, 0.5 to 0.999), p=0.0245, and the other two comparisons remained not statistically significant; the data for longer follow-up are not represented in Figure 2A and in Table 7. See Table 1-7A in Appendix 7 for more detailed results.

The HD12 [59] showed no statistically significant difference between 8×eBEACOPP and 4×eBEACOPP plus 4×bBEACOPP both at 78- [59] (HR, 1.14; 95% CI, 0.83 to 1.56), as well as at 97-month follow-up [77] (HR, 1.02; 95% CI, 0.77 to 1.36).

The HD9 trial showed no statistically significant difference between bBEACOPP and eBEACOPP at 72 months of follow-up [54] (88% vs. 91%, p=0.06).

The overall certainty of the evidence across the five different doses and schedules for OS was moderate because of inconsistency in the direction of the results: the 6×eBEACOPP versus 8×eBEACOPP comparison in the HD15 trial [58] showed a statistically significant difference in OS in favour of the higher intensity regimen when comparing more with less cycles of eBEACOPP while the other studies. Table 7 presents the relative and absolute effects, as well as for certainty of the evidence.

Figure 2A. Overall survival for Comparison 2: Higher- versus Lower BEACOPP intensity/dosage.



Note: HD15e and HD15-14 and HD15-6 represent three randomizations in the same HD15 study [58]. In text we considered each randomization as a separate study.

B_{14}=BEACOPP$_{14}$; bB=baseline BEACOPP; CI = Confidence Interval; eB=escalated BEACOPP

Recurrence

PFS

A), B), and C) 8×eBEACOPP vs. 6×eBEACOPP, 8×eBEACOPP vs. 8×BEACOPP$_{14}$, and 6×eBEACOPP vs. 8×BEACOPP$_{14}$

The HD15 reported a PFS rate at 48 months of follow-up [58] of 85.6% (97.5% CI, 82.3 to 88.9) for 8×eBEACOPP, 90.3% (97.5% CI, 87.6 to 93.0) for 6×eBEACOPP, and 85.8% (97.5% CI,

82.4 to 89.2) for 8×BEACOPP$_{14}$. The authors of the HD15 changed the primary endpoint from FFTF to PFS for the extended follow-up analysis. The PFS rates at 102-month follow-up [82] were: 8×eBEACOPP: 80.9% (97.5% CI, 77.3 to 84.6), 6×eBEACOPP: 83.7% (97.5% CI, 79.9 to 87.4), and 8×BEACOPP$_{14}$: 83.6% (97.5% CI,80.2 to 87.0). The authors reported that none of the comparisons were statistically significant. Noninferiority of both the lower intensity/dosage regimens compared with the higher intensity/dosage (i.e., 8×eBEACOPP) was confirmed as the margin of 1.51 for the HR could be excluded for both (6×eBEACOPP, HR, 0.7; 97.5% CI, 0.5 to 1.0 [calculated 95% CI, 0.547 to 0.897]; 8×BEACOPP$_{14}$, HR, 0.9; 97.5% CI, 0.7 to 1.2 [calculated 95% CI, 0.703 to 1.153]) (see Table 7).

D) 8×eBEACOPP vs. 4×eBEACOPP +4×bBEACOPP

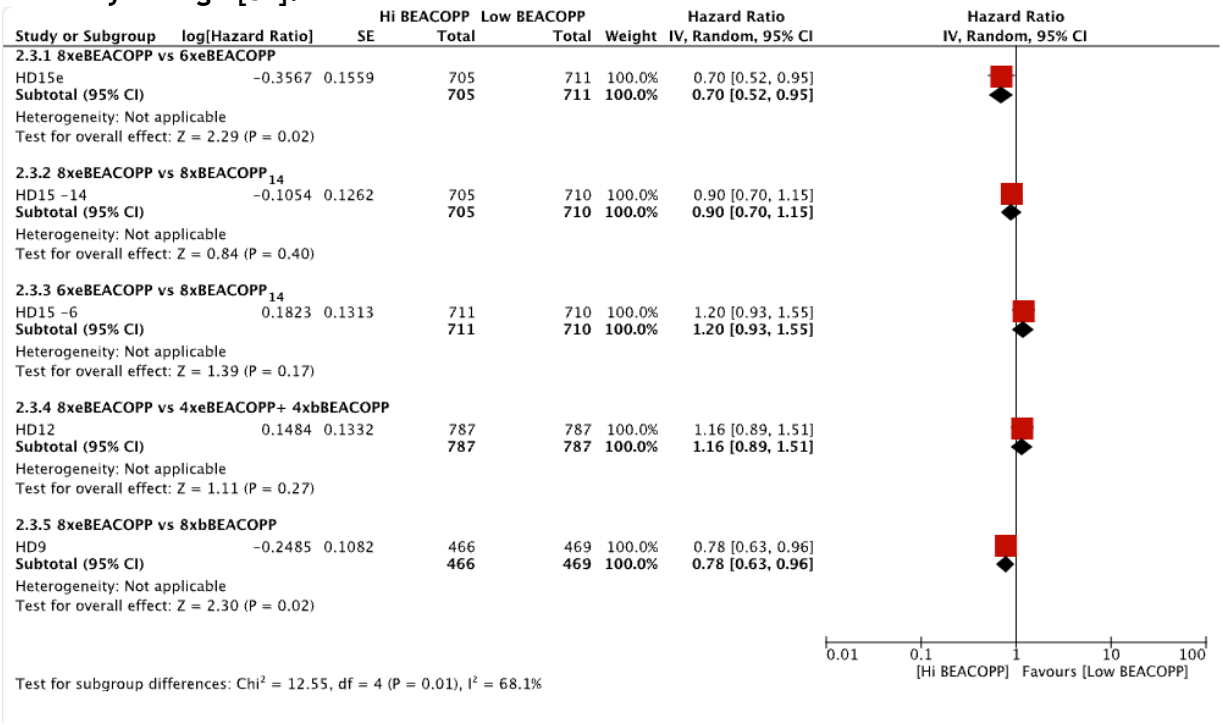
The HD12 [59] reported at both 78- and 97-month follow-up (HR, 1.16 [95% CI, 0.89 to 1.50], and HR, 1.13 [95% CI, 0.89 to 1.43], respectively); reduced intensity/dosage chemotherapy (4+4 combination) was non inferior to the standard, high-intensity/dosage regimen (8×eBEACOPP) within the noninferiority margin of 1.50. See Tables 7 in this section and 1-7A in Appendix 7, for more details.

E) 8×eBEACOPP vs. 8×bBEACOPP

The HD9 [54] five-year follow-up, on the other hand, reported a statistically significantly better progression rate for the higher-intensity/dosage (eBEACOPP) than for the lower intensity/dosage regimen (eBEACOPP): 2% and 8% respectively, p=0.001. See Tables 7 in this section and 1-7A in Appendix 7.

The overall certainty of the evidence for PFS was moderate because patients, clinicians and outcome assessors were not blinded to patients' assignment and because there was an inconsistency in the direction of the results among the HD9 [54], the HD15 [58] and the other studies (Table 7).

Figure 2B. Progression-free survival. Comparison 2: Higher- versus lower BEACOPP intensity/dosage [82].



Note: HD15e and HD15-14 and HD15-6 represent three randomizations in the same HD15 study [58]. In text we considered each randomization as a separate study.

$BEACOPP_{14}$=BEACOPP given on a shorter, 14-day cycle; bBEACOPP=baseline BEACOPP; CI=Confidence Interval; eBEACOPP=escalated BEACOPP.

Other measures of recurrence

FFTF

A) 8×eBEACOPP vs. 6×eBEACOPP, B) 8×eBEACOPP vs. 8×$BEACOPP_{14}$, and C) 6×eBEACOPP vs. 8×$BEACOPP_{14}$, in the HD15 trial [58,82]

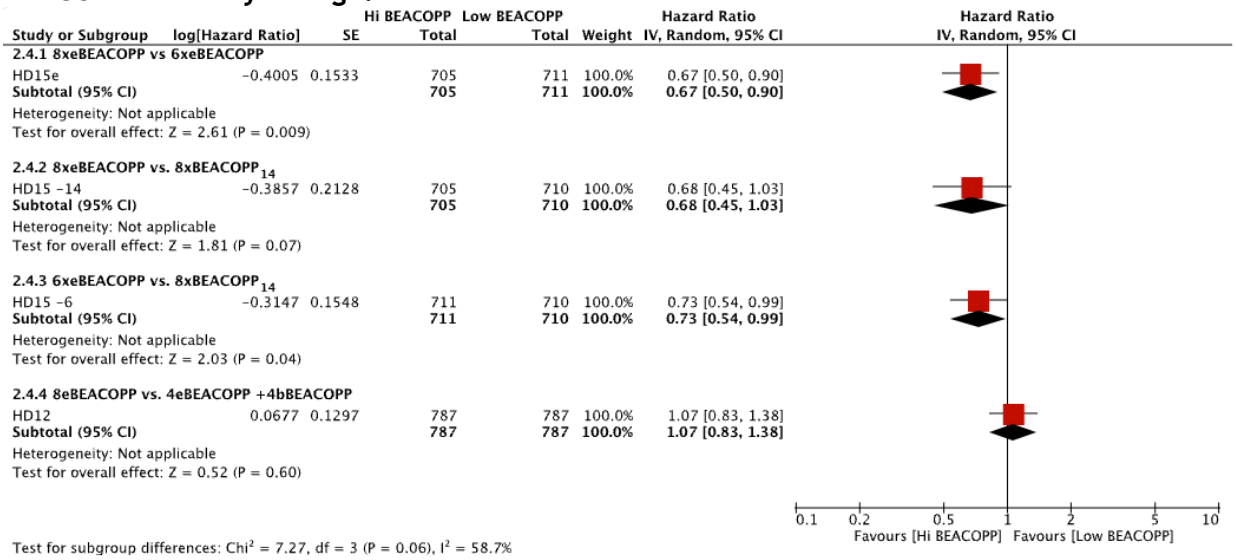
The HD15 trial at 48-month follow-up [58] reported that 6×eBEACOPP and 8×$BEACOPP_{14}$ were non-inferior to 8×eBEACOPP as their 97.5% CI for the HR excluded the noninferiority margin of 1.51 (6×eBEACOPP vs. 8×eBEACOPP: 97.5% CI, 0.44 to 1.02; 8×$BEACOPP_{14}$ vs. 8×eBEACOPP: 97.5% CI, 0.62 to 1.36).

In the superiority analysis, the HD15 trial [58] reported that treatment with 6×eBEACOPP was statistically significantly superior to 8×eBEACOPP (FFTF rates 89.3% [97.5% CI, 86.5 to 92.1] and 84.4% [97.5% CI, 81.0 to 87.7], respectively; HR, 0.67 [97.5% CI, 0.47 to 0.95; calculated 95% CI, 0.496 to 0.905], p=0.009). Treatment with 6×eBEACOPP was also statistically significantly superior to 8×$BEACOPP_{14}$ (FFTF rates 89.3% [97.5% CI, 86.5 to 92.1] and 85.4% [97.5% CI, 82.1 to 88.7], respectively; HR, 0.73 [97.5% CI, 0.51 to 1.03, calculated 95% CI, 0.539 to 0.989], p=0.042). This study did not show any statistically significant difference between 8×$BEACOPP_{14}$ versus 8×eBEACOPP (HR, 0.68 [97.5% CI, 0.42 to 1.10, calculated 95% CI, 0.448 to 1.032], p=0.07).

D) 8×eBEACOPP vs. 4×eBEACOPP +4×bBEACOPP [59,77]

At 78-month follow-up, the HD12 [59] reported a FFTF rate of 86.4% for 8×eBEACOPP compared with 84.8% for the 4×eBEACOPP + 4×bBEACOPP combination (HR, 1.07; 95% CI, 0.83 to 1.38).

Figure 2C. Freedom from treatment failure for Comparison 2: Higher- versus lower BEACOPP intensity/dosage.

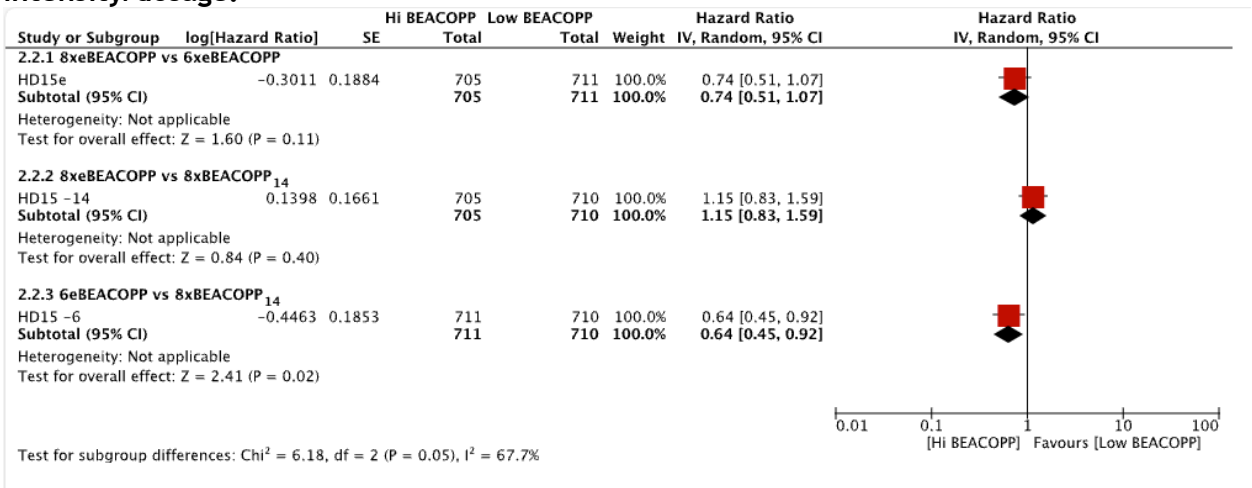


bBEACOPP=baseline BEACOPP; $BEACOPP_{14}$=BEACOPP given on a shorter, 14-day cycle; CI=Confidence Interval; eBEACOPP=escalated BEACOPP

TTP

The HD15 trial [58] reported that treatment with 6×eBEACOPP had a statistically significantly better TTP than treatment with 8×$BEACOPP_{14}$ (HR, 0.64 [97.5% CI, 0.43 to 0.97; calculated 95% CI, 0.45 to 0.92], p=0.016), whereas the comparisons 6×eBEACOPP versus 8×eBEACOPP (HR, 0.74 [97.5% CI, 0.48 to 1.13; calculated 95% CI, 0.512 to 1.071], p=0.11), and 8×$BEACOPP_{14}$ versus 8×eBEACOPP (HR, 0.15 [97.5% CI, 0.78 to 1.68; calculated 95% CI, 0.831 to 1.592], p=0.4) did not show any statistically significant difference (Table 1-7A in Appendix 7).

Figure 2D. Time to progression for Comparison 2: Higher- versus lower BEACOPP intensity/dosage.



$BEACOPP_{14}$=BEACOPP given on a shorter, 14-day cycle; CI=Confidence Interval; eBEACOPP=escalated BEACOPP

The overall certainty of the evidence for the recurrence outcome (including all the measures that we reported) was moderate-to-low because the studies were at risk of bias, and because there was an inconsistency in the direction of the results among trials.

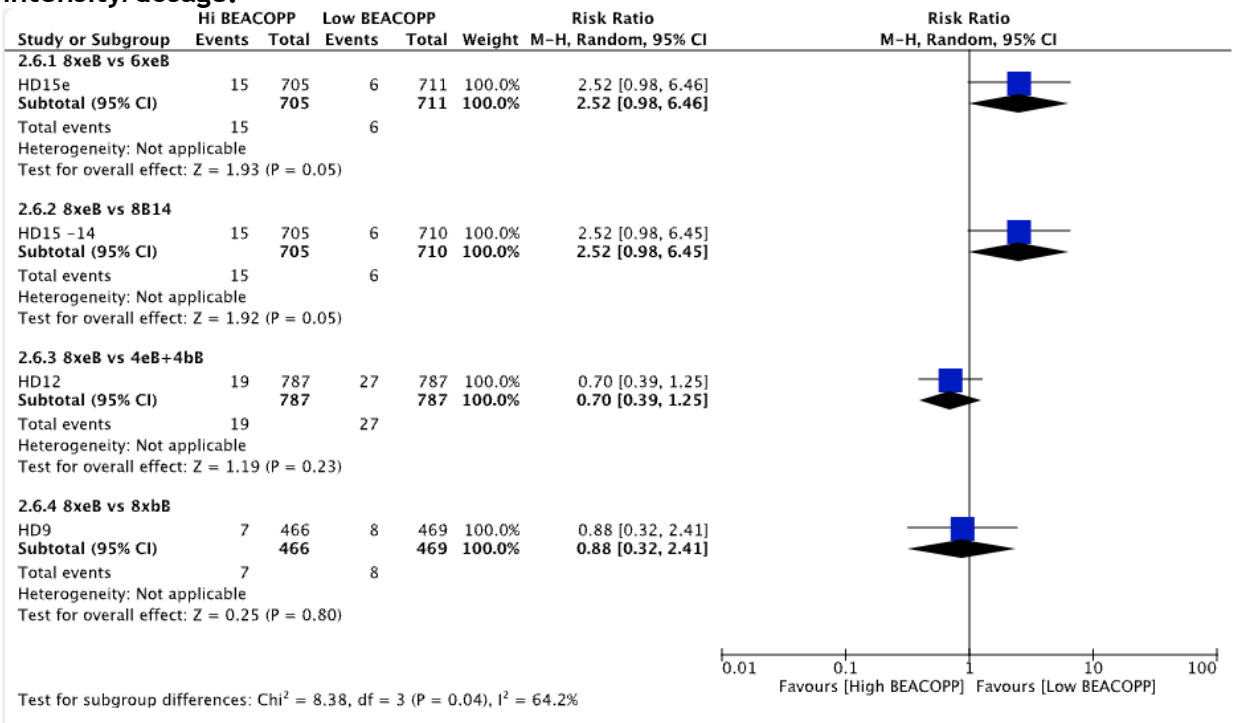
Adverse effects

TRD

The HD15 trial [58] reported low event rates at 98-month follow-up [82]: 15 patients (2.1%) for 8×eBEACOPP, six patients (0.8%) for 6×eBEACOPP, and seven patients (1%) for 8×BEACOPP$_{14}$; p values were not reported. The HD12 trial, at 97-month follow-up [77] also reported a very small number of events: 11 (3%) patients for 8×eBEACOPP, and 10 (2.5%) patients for 4×eBEACOPP+4×bBEACOPP. Von Tresckow et al. [77] reported the TRD of patients treated in the bBEACOPP versus eBEACOPP groups for Objective 2 in the HD9: numbers were small (seven patients [1.5%] in the 8×bBEACOPP versus seven patients [1.5%] in the 8×eBEACOPP); p values were not reported.

The certainty of the evidence for this outcome was very low because of risk of bias and imprecision. Event rates were very low, although clinically important given their serious nature. Therefore, we expect the confidence interval around the point estimates to be very large, and we are unable to determine if an apparent difference is due to a difference in effect or to chance alone.

Figure 2E. Treatment-related death for Comparison 2: Higher- versus lower BEACOPP intensity/dosage.



B_{14}=BEACOPP given on a shorter, 14-day cycle; B=BEACOPP; bB=baseline BEACOPP; CI=Confidence Interval; eB=escalated BEACOPP

Important outcomes

Pulmonary toxicity

None of the trials included for this comparison (i.e., the HD15 [58,82], the HD12 [59,77], and the HD9 [54,64,81]) reported on this outcome.

We considered the certainty of the evidence for pulmonary toxicity for the comparison higher-intensity versus lower-intensity BEACOPP very low.

SPM

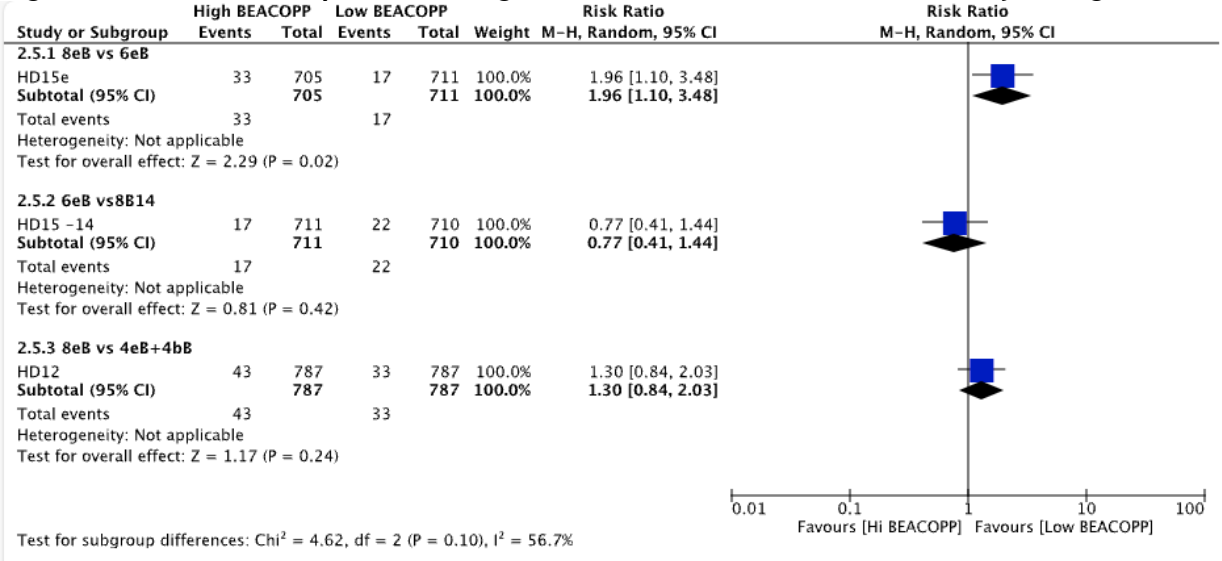
The HD15 trial [58] at 48-month follow-up reported 33 SPM (4.7%) in the 8×eBEACOPP group, 17 (2.4%) in the 6×eBEACOPP group, and 22 (3.1%) in the 8×BEACOPP$_{14}$ group. Patients in the 8×eBEACOPP group experienced more events than patients in the 6×eBEACOPP group (p=0.02). The cumulative incidence of SPM was 10% (97.5% CI, 7% to 13%) in the 8×eBEACOPP group, 7% (97.5% CI, 4% to 9%) in the 6×eBEACOPP group, and 7% (97.5% CI, 4% to 10%) in the 8×BEACOPP$_{14}$ group. When comparing 6×eBEACOPP versus 8×eBEACOPP and 8×BEACOPP$_{14}$ versus 8×eBEACOPP (HR, 0.7 [97.5% CI the difference in SPM was not statistically significant]). See Table 1-7A in Appendix 7 for more information.

The HD12 trial at 78-month follow-up [77] reported low SPM rates, (i.e., 43 events [5.5%] in the 8×BEACOPP group versus 33 [4.2%] in the 4+4 combination; p not significant).

The HD9 [54] reported no difference for all SPM (including solid tumours), which was the outcome considered important for this report. At 10-year follow-up, the cumulative incidence of all SPM was 5.3% in the COPP/ABVD group, 7.9% in the 8×bBEACOPP, and 6.5% in the 8×eBEACOPP (p=0.82). However, at the five-year follow-up, a statistically significant higher rate of AML/MDS was reported in the eBEACOPP arm compared with bBEACOPP and with COPP/ABVD (3.2% vs. 2.2% vs. 0.4%, p=0.03), and this effect remained at 15-year follow-up [77].

The certainty of the evidence for SPM was very low because of risk of bias, inconsistency and imprecision.

Figure 2F. SPM for Comparison 2: Higher- versus lower BEACOPP intensity/dosage.



B$_{14}$=BEACOPP given on a shorter, 14-day cycle; B=BEACOPP; bB=baseline BEACOPP; CI = Confidence Interval; eB=escalated BEACOPP

Infertility

None of the RCTs included for this comparison reported on infertility. Behringer et al. [96], a noncomparative, corollary study of the HD15 [58], reported on the gonadal function of 232 women and 417 men enrolled in the HD15 trial [58]. Results of this study are reported in Appendix 7, Table 3-7A.

Table 7 reports the relative and absolute results, outcome by outcome for the different drug dosages and schedules, when available, as well as the certainty and the importance of the evidence. See Table 1-7A in Appendix 7 for more details on the results of the studies described above.

Table 7. Comparison 2: Higher- versus lower intensity/dosage BEACOPP

Certainty assessment							№ of patients		Effect		Certainty of the evidence§	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Higher intensity BEACOPP	Lower intensity BEACOPP	Relative (95% CI)	Absolute (95% CI)		
Overall Survival												
Overall survival 8×eBEACOPP vs. 6×eBEACOPP (follow-up: median 48 months; assessed with: HR)												
1	RCT HD15 [58]	Not serious	Not serious	Not serious	Not serious	none	705	711 / 0.7%*	HR 0.60 (0.39 to 0.92)†	3 fewer per 1000 (from 4 fewer to 1 fewer)	High	CRITICAL
Overall survival 8×eBEACOPP vs. 8×BEACOPP$_{14}$ (follow-up: median 48 months; assessed with: HR)												
1	RCT HD15 [58]	Not serious	Not serious	Not serious	Not serious	none	705	710 / 0.9%*	HR 0.68 (0.42 to 1.10)†	3 fewer per 1000 (from 5 fewer to 1 more)	High	CRITICAL
Overall survival 6×eBEACOPP vs. 8×BEACOPP$_{14}$ (follow-up: median 48 months; assessed with: HR)												
1	RCT HD15 [58,82]	Not serious	Not serious	Not serious	Not serious	none	711	710 / 0.9%	HR 0.88 (0.63 to 1.13)†	1 fewer per 1000 (from 2 fewer to 1 more)	High	CRITICAL
Overall survival 8×eBEACOPP vs. 4×eBEACOPP + 4×bBEACOPP (follow-up: median 78 months; assessed with: HR)												
1	RCT HD12 (chemotherapy comparison) [59]	Not serious	Not serious	Not serious	Not serious	none	787	787 / 10%	HR 1.14 (0.83 to 1.56)	13 more per 1000 (from 16 fewer to 52 more)	High	CRITICAL

Overall survival 8×eBEACOPP vs. 8×bBEACOPP (follow-up: median 72 months; assessed with: HR)												
1	RCT HD9 [54]	Serious[a]	Not serious	Not serious	Serious[b]	none	466	469 / 12.0%	HR 1.466 (0.984 to 2.185)	51 more per 1000 (from 2 fewer to 124 more)	Low	CRITICAL
Progression-free survival												
PFS 8×eBEACOPP vs. 6×eBEACOPP (follow-up: median 48 months; assessed with HR)												
1	RCT HD15 [58]	Serious[c]	Not serious	Not serious	Not serious	none	705	711 / 16%	HR 0.7 (0.547-0.897)	53 fewer per 1000 (from 81 fewer to 18 fewer)	Moderate	CRITICAL
PFS 8×eBEACOPP vs. 8×BEACOPP$_{14}$ (follow-up: median 102 months; assessed with HR)												
1	RCT HD15 [58]	Serious[c]	Not serious	Not serious	Not serious	none	705	710 / 16%	HR 0.9 (0.703 to 1.153)	15 fewer per 1000 (from 45 fewer to 22 more)	Moderate	CRITICAL
PFS 6×eBEACOPP vs. 8×BEACOPP$_{14}$ (follow-up: median 102 months; assessed with HR)												
1	RCT HD15 [58]	Serious[c]	Not serious	Not serious	Not serious	none	711	710 / 16%	HR 1.2 (0.928 to 1.552)	29 more per 1000 (from 11 fewer to 77 more)	Moderate	CRITICAL
PFS 8×eBEACOPP vs. 4×eBEACOPP+ 4×bBEACOPP (follow-up: median 78 months; assessed with: HR)												
1	RCT HD12 (chemotherapy comparison) [59]	Serious[c]	Not serious	Not serious	Not serious	none	787	787	HR 1.16 (0.89 to 1.50)	22 more per 1000 (from 15 fewer to 66 more)	Moderate	CRITICAL
PFS 8×eBEACOPP vs. 8×bBEACOPP (follow-up: median 72 months; assessed with: HR)												
1	RCT HD9 [54] (Objective 2)	Serious[a]	Serious[b]	Not serious	Serious[c]	none	466	469 / 8%	HR 2.667 (1.487 to 4.785)	119 more per 1000 (from 37 more to 249 more)	Low	CRITICAL

Time to Progression

Time to progression 8×eBEACOPP vs. 6×eBEACOPP (follow-up: median 48months; assessed with: HR)											
1	RCT HD15 [58]	Serious[c]	Not serious	Not serious	Not serious	none	705	711 / 10%	HR 0.74 (0.512 to 1.071) †	25 fewer per 1000 (from 47 fewer to 7 more)	moderate	CRITICAL
Time to progression 8×eBEACOPP vs. 8×BEACOPP$_{14}$ (follow-up: median 48 months; assessed with: HR)											
1	RCT HD15 [58]	Serious[c]	Not serious	Not serious	Not serious	none	705	710 / 10%	HR 1.15 (0.831 to 1.592)†	14 more per 1000 (from 16 fewer to 54 more)	moderate	CRITICAL
Time to progression 6eBEACOPP vs. 8×BEACOPP$_{14}$ (follow-up: median 48 months; assessed with: HR)											
1	RCT HD15 [58]	Serious[c]	Not serious	Not serious	Not serious	none	711	710 / 14%	HR 0.64 (0.445 to 0.920)†	48 fewer per 1000 (from 75 fewer to 10 fewer)	moderate	CRITICAL

Freedom from treatment failure

Freedom from treatment failure (FFTF) 6×eBEACOPP vs. 8×eBEACOPP (follow-up: median 48 months; assessed with: HR)											
1	RCT HD15 [58]	Serious[d]	Not serious	Not serious	Not serious	none	705	711 / 10%	HR 0.67 (0.496 to 0.905)†	32 fewer per 1000 (from 49 fewer to 9 fewer)	Moderate	CRITICAL
FFTF 6×eBEACOPP vs. 8×BEACOPP$_{14}$ (follow-up: median 48 months; assessed with: HR)											
1	RCT HD15 [58]	Serious[d]	Not serious	Not serious	Not serious	none	711	710 / 14%	HR 0.73 (0.539 to 0.989)†	36 fewer per 1000 (from 62 fewer to 1 fewer)	Moderate	CRITICAL
FFTF 8×eBEACOPP vs. 8×BEACOPP$_{14}$ (follow-up: median 48 months; assessed with: HR)											

1	RCT HD15 [58]	Serious[d]	Not serious	Not serious	Not serious	none	705	710 / 14%	HR 0.68 (0.448 to 1.032)†	42 fewer per 1000 (from 75 fewer to 4 more)	Moderate	CRITICAL
FFTF 8×eBEACOPP vs. 4×eBEACOPP + 4×bBEACOPP (follow-up: median 78 months; assessed with: HR)												
1	RCT HD12 (chemotherapy comparison) [59]	Serious[d]	Not serious	Not serious	Not serious	none	787	787	HR 1.07 (0.83 to 1.38)	9 more per 1000 (from 22 fewer to 48 more)	moderate	CRITICAL
Adverse events												
Treatment-related death												
TRD 8×eBEACOPP vs. 6×eBEACOPP (follow-up: median 48 months; assessed with: RR)												
1	RCT HD15 [58]	Serious[d]	Not serious	Not serious	Very serious[e]	none	15 / 705 (2.1%)	6 / 711 (0.8%)	RR 2.52 (0.98 to 6.46)	13 more per 1000 (from 0 fewer to 46 more)	Very low	CRITICAL
TRD 8×eBEACOPP vs. 8×BEACOPP14 (follow-up: median 48 months; assessed with: RR)												
1	RCT HD15 [58]	Serious[d]	Not serious	Not serious	Very serious[e]	none	15 / 705 (2.1%)	6 / 710 (0.8%)	RR 2.52 (from 0.98 to 6.46)	13 more per 1000 (from 0 fewer to 46 more)	Very low	CRITICAL
TRD 8×eBEACOPP ± RT vs. 4×eBEACOPP + 4×bBEACOPP ± RT (follow-up: median 78 months; assessed with: RR)												
1	RCT HD12 (chemotherapy comparison) [77]	Serious[d]	Not serious	Not serious	Very serious[e]	none	19 / 787 (2.4%)	27 / 787 (3.4%)	RR 0.70 (from 0.39 to 1.25)	10 fewer per 1000 (from 21 fewer to 9 more)	Very low	CRITICAL
TRD 8×eBEACOPP vs. 8×bBEACOPP (follow-up: median 28 months; assessed with: assessed with RR)												
1	RCT HD9 [77]	Serious[a]	Not serious	Not serious	Very serious[ce]	none	7 / 466 (1.5%)	8 / 469 (1.7%)	RR 0.88 (from 0.32 to 2.41)	2 fewer per 1000 (from 12 fewer to 24 more)	Very low	CRITICAL

Secondary primary malignancies												
Secondary primary malignancies 8×eBEACOPP vs. 6×eBEACOPP (Follow-up 48 months, assessed with RR)												
1	RCT HD15 [58]	Serious[d]	Not serious	Not serious	Very serious[e]	none	33 / 705 (4.7%)	17 / 711 (2.4%)	RR 1.96 (1.10 to 3.48)	23 more per 1000 (from 2 more to 59 more)	Very low	IMPORTANT
Secondary primary malignancies 6×eBEACOPP vs. 8×BEACOPP$_{14}$ (Follow-up 48 months, assessed with RR)												
1	RCT HD15 [58]	Serious[d]	Not serious	Not serious	Very serious[e]	none	17/711 (2.4%)	22/710 (3.1%)	RR 0.77 (0.41 to 1.44)	7 fewer for 1000 (from 18 fewer to 14 more)	Very low	IMPORTANT
Secondary primary malignancies 8×eBEACOPP vs. 4×eBEACOPP +4×bBEACOPP (Follow-up 78 months, assessed with RR)												
1	RCT HD12 (chemotherapy comparison) [77]	Serious[d]	Not serious	Not serious	Very serious[e]	none	43 / 787 (5.5%)	33 / 787 (4.2%)	RR 1.30 (0.84 to 2.03)	13 more per 1000 (from 7 fewer to 43 more)	Very low	IMPORTANT
Infertility												
No data were reported for this outcome												
Pulmonary toxicity												
No data were reported for this outcome												

* For overall survival an event was considered death.
†The 97.5% CI is reported in the HD15 study. We estimated the 95% CI that is used in the GRADE pro application by using the provided survival rates.
‡ The HR was not provided and it was calculated based on the p value an on the number of patients randomized to each treatment group.
§ Definitions of degrees of certainty:
High: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low: Any estimate of effect is very uncertain.
a This was an unblinded study that was stopped early for benefit, and there were deviations from the intended interventions because of the trial context
b The direction of the results of the HD9 trials were in the opposite direction compared to the other studies comparing higher versus lower intensity BEACOPP
c This study was stopped early for benefit.
d Patients, clinicians, and outcome assessors were unblinded to patients' assignment.
e The event rate was very low in both groups. With such a low event rate, the confidence interval is very wide. It is difficult to determine whether the observed effect is due to the treatment or simply to chance.
bBEACOPP=baseline BEACOPP; BEACOPP=bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, prednisone; CI=Confidence interval; eBEACOPP=escalated BEACOPP; FFTF=Freedom from treatment failure; HR=Hazard ratio; PFS=Progression-free survival; RCT=Randomized controlled trial; RR=Relative risk; RT=Radiotherapy; TRD=Treatment-related death

Comparison 3: PET-adapted strategies: PET-directed versus non-PET-directed strategies or standard treatment versus deescalated treatment based on PET-2 results

Four RCTs constitute the body of evidence for this comparison: the AHL2011 [17], the RATHL [49], the HD18 [16], and the GITIL/FIL HD0607 [48] trials. Of these trials, the AHL2011 [17] randomized patients to a PET-driven strategy versus a standard strategy without PET. The other studies used PET for identifying subgroups of PET-positive and PET-negative patients, which were in turn randomized to different interventions. Because of the importance of PET in the selection of patients we placed all these studies in this section.

The trials that reported results for PET-positive patients (i.e., the RATHL trial [49], the GITIL/FIL0607 [48], and the HD18 [16]), either did not randomize PET-positive patients to different treatments (i.e., RATHL trial [49]), or randomized PET-positive patients to a BEACOPP-based regimen with or without rituximab (i.e., the GITIL/FIL0607 [48] and the HD18 [16] trials). As rituximab is not one of the interventions of interest for this report, we will not describe the studies for this subgroup of patients in any further detail here. The interested reader can find more information in Table 1-7A, Appendix 7. In the following sections we describe first the results of the AHL2001 trial [17], then the results for the subgroup of PET-negative patients included in the HD18 and in the RATHL trials [16,49].

The AHL2001 trial [17,84] was a noninferiority RCT that tested a PET-directed approach to treatment. Patients were assigned to a standard group vs. a PET-driven treatment group. Patients in both groups received two cycles of upfront eBEACOPP, and then PET (i.e., PET-2).

In the standard treatment group, all patients, regardless of the results of PET-2, received 2×eBEACOPP; then PET-4.

In the PET-driven arm, if PET-2 was positive, patients received 2×eBEACOPP, then PET-4; if PET-2 was negative, therapy was de-escalated and patients received two cycles of ABVD, then PET-4.

In both, the standard treatment arm, and the PET-2-positive group of the PET-driven arm, if PET-4 was positive, patients received salvage treatment. If PET-4 was negative, patients received 2×eBEACOPP. In the PET-2-negative arm, if PET-4 was positive, patients were given salvage treatment, and if it was negative they would be given two more cycles of ABVD. The primary outcome for this trial was PFS at five-year follow-up.

In the noninferiority RATHL trial [49,85] all patients were given two preliminary cycles of ABVD, then PET-2. Those who were PET-2 negative were randomized to 4×ABVD versus 4×AVD. The primary outcome was PFS at three years.

In the HD18 trial [16], patients were given two preliminary cycles of eBEACOPP, then PET was performed. PET-2 negative patients were randomized to six or eight cycles (standard) or four cycles of eBEACOPP with the aim to show noninferiority for PFS of fewer compared with more eBEACOPP cycles.

The GITIL/FIL0607 [48] randomized PET-2-negative patients to 6×ABVD with or without radiotherapy; this study is relevant to Question 3, and we will not describe it any further here.

Among the companion studies, the cohort study by Demeestere et al. [90] studied gonadal dysfunction in patients included in the AHL2011 [17,84] who switched from a higher-intensity regimen to a lower one. Kobe et al. [91] in another cohort study analyzed PFS and OS according to Deauville scores to identify the best cut-off to separate risk groups in patients included in the HD18 [16]. Anderson et al. [93] studied the ovarian function of women exposed to PET-adapted chemotherapy as part of the RATHL trial [49].

The results of these ancillary observational studies are presented in detail in Table 3-7A, Appendix 7.

A. PET-directed vs. standard (no PET) strategy

Table 8A presents the relative and absolute effect and the certainty of the evidence for this comparison.

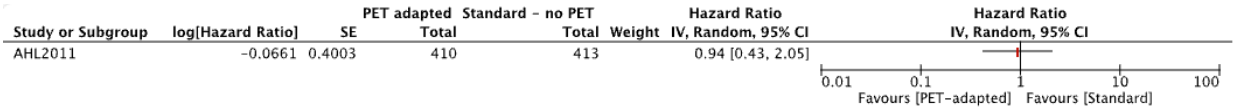
Critical Outcomes

OS

The AHL2011 trial [17], after two cycles of eBEACOPP, reported no statistically significant difference between standard treatment and PET-driven strategy both at 50.4 months (95.2% vs. 96.4%; HR, 0.936 [95% CI, 0.427 to 2.051], p=0.43), and at 67.25 months (97.7% vs. 97.7%; HR, 1.012 [95% CI, 0.50 to 2.12], p=0.53).

The certainty of the evidence for OS is moderate because of serious risk of bias due to of deviations from the intended interventions and missing data.

Figure 3A.1. PET-directed versus Standard treatment at 50.4 months: OS



OS=overall survival; PET=Positron emission tomography; SE=standard error

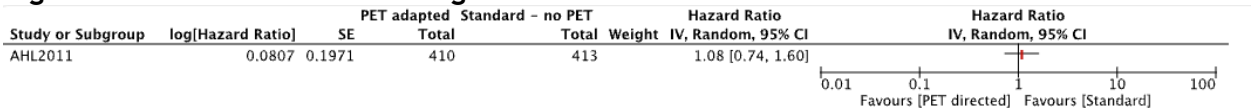
Recurrence

PFS

The AHL2011 trial [17] was planned as a noninferiory trial with a predefined margin of 10 per cent difference in PFS between the PET-driven and standard strategy groups. At 50.4-month follow-up, the between-group difference in PFS was lower (–0.5%; 95% CI, –6.1 to 5.0) than the pre-specified difference and the condition for noninferiority was met. In the intention-to-treat analysis PFS rate was 86.2% (95% CI, 81.6 to 89.0) in the standard group, and 85.7% (95% CI, 81.4 to 89.1) in the PET-directed group (HR, 1.084; 95% CI 0.737 to 1.596). See Tables 8A and 8B in this section and 1-7A in Appendix 7 for more details.

The certainty of the evidence for PFS was very low because of very serious risk of bias, and imprecision.

Figure 3A.2. PET-directed strategies versus Standard treatment at 50.4 months: PFS



PET=positron emission tomography; PFS=progression-free survival; SE=standard error

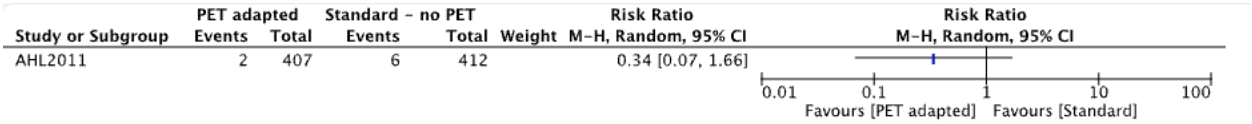
Adverse effects

TRD

The AHL2011 [17] reported a death rate from treatment-related adverse effects of 1.4% (6 of 412 patients) in the standard group vs. 0.48% (2 of 407 patients) in the PET-adapted group; p values were not reported.

The certainty of the evidence for TRD was very low because of very serious risk of bias and imprecision.

Figure 3A.3. PET-directed strategies versus standard treatment at 50.4 months: TRD



PET=positron emission tomography; TRD=treatment-related death

Important outcomes

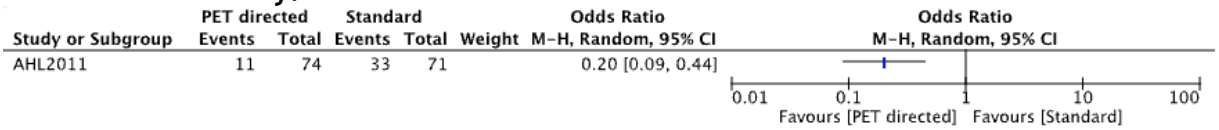
SPM

The AHL2011 [84] reported the rate of SPM at 67.25 months as 3.2% in the 6×eBEACOPP (standard arm) compared with 2.2% in the PET-driven arm; p values were not reported.
The certainty of the evidence for TRD was very low because of risk of bias and imprecision.

Infertility

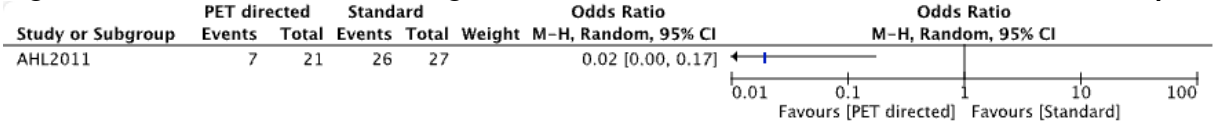
Data on infertility are reported by a cohort study (Demeestere et al. 2021 [90]) companion of the AHL2011 trial [17]. In women, hormone status favoured the PET-driven strategy as measured with premature ovarian insufficiency (FSH >24 IU/L, odds ratio [OR], 0.20; 95% CI, 0.08 to 0.50, p=0.001; age-adjusted OR [aOR], 0.09; 95% CI, 0.03 to 0.32; p<0.001); low ovarian reserve (anti-müllerian hormone, <0.5 ng/mL; OR, 0.15; 95% CI, 0.04 to 0.56, p=0.005), and ovarian function recovery (i.e., FSH level <25 IU/L or pregnancy; HR, 2.52; 95% CI, 1.73 to 3.67; p<0.0001). Men had a lower risk of severe testicular damage (OR, 0.26; 95% CI, 0.13 to 0.5, p<0.0001) and higher sperm recovery parameters in the PET-driven group compared with the standard eBEACOPP group.

Figure 3A.4. PET-directed strategies versus standard treatment at 60 months: Premature ovarian insufficiency.



PET=Positron emission tomography

Figure 3A.5. PET-directed strategies versus Standard treatment at 29 months: Azoospermia



PET=Positron emission tomography

Table 8A. Comparison 3A. PET-adapted strategies: PET-directed vs. standard (no PET) strategy

Certainty assessment							№ of patients		Effect		Certainty of the evidence§	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	PET-adapted	Standard no PET	Relative (95% CI)	Absolute (95% CI)		
Overall survival PET-driven vs. standard strategy (no PET) (follow-up: median 50.4 months; assessed with: HR)												
1	RCT AHL2001 [17,84]	Serious[a]	Not serious	Not serious	Not serious	None	410	413/5%	HR 0.936 (0.427 to 2.051)	2 fewer per 1000 (from 28 fewer to 50 more)	Moderate	CRITICAL
Progression-free survival												
Progression-free survival PET-driven strategy vs. standard strategy (no PET) (follow-up: median 50.4 months; assessed with: HR)												
1	RCT AHL2001 [17,84]	Very serious[b]	Not serious	Not serious	Serious[c]	None	410	413/15%	HR 1.084 (0.737 to 1.596)	12 more per 1000 (from 37 fewer to 78 more)	Very low	CRITICAL
Adverse events												
Treatment-related death: PET-driven strategy vs. standard strategy (no PET) (follow-up: median 50.4 months)												
1	RCT AHL2001 [17,84]	Very serious[b]	Not serious	Not serious	Serious[c]	None	407	412/0.5%	RR 0.34 (from 0.07 to 1.66)	10 fewer per 1000 (from 14 fewer to 9 more)	Very low	CRITICAL
Secondary primary malignancies cumulative incidence PET-driven strategy vs. standard strategy (no PET) (follow-up: median 67.25 months; assessed with HR)												
1	RCT AHL2001 [17,84]	Very serious[b]	Not serious	Not serious	Serious[c]	None	407/5 (1.2%)	412/10 (2.4%)	0.51 (0.17 to 1.47)	12 fewer per 1000 (from 20 fewer to 11 more)	Very low	IMPORTANT

Premature ovarian insufficiency (FSH ≥25 UI/L (follow-up: median 60 months; assessed with OR)												
1	RCT AHL2001 [17,84,90]	Very serious[b]	Not serious	Not serious	Serious[cd]	None	11/74 (14%)	33/71 (46.5%)	OR 0.09 (0.03 to 0.32)	392 fewer per 1000 (from 439 fewer to 247 fewer)	Very low	IMPORTANT
Azoospermia (follow-up median 29 months; assessed with OR)												
1	RCT AHL2001[90]	Very serious[b]	Not serious	Not serious	Serious[cd]	None	7/21 (33.3%)	26/27 (96.3%)	OR 0.01923 (0.00214 to 0.17247)	630 fewer per 1000 (from 910 fewer to 145 fewer)	Very low	IMPORTANT

a There were some deviations from intended interventions, and missing data (10% and 13% respectively in the standard and in the intervention group).
b Patients, clinicians and outcome assessor were aware of patient assignment, and there were some deviations from the intended interventions because of this.
c Very few events, we expect wide confidence intervals.
d Data were available for this outcome on a very small subgroup of patients

CI=Confidence interval; FSH=Follicle stimulating hormone; HR=Hazard ratio; OR=odds ratio; PET=Positron emission tomography; RCT=Randomized controlled trial; RR=Relative risk

B. PET-adapted strategies. PET-2 negative patients randomized, after two cycles of ABVD and PET, to AVD versus ABVD: the RATHL trial [49,85].

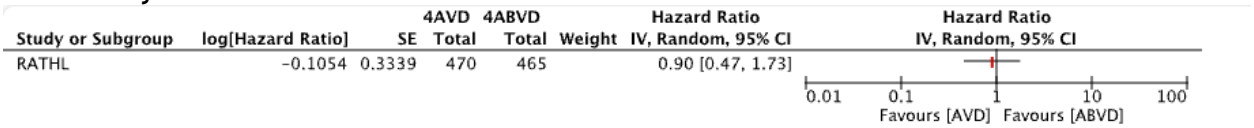
Table 8B presents the relative and absolute effect and the certainty of the evidence for this comparison.

Critical outcomes

OS

At 36-month follow-up, the RATHL trial [49] reported no statistically significant between-groups difference in PET-2-negative patients who, after the initial two cycles of ABVD were treated with either 4×ABVD or 4×AVD. OS rates were 97.2% (95% CI, 95.1 to 98.4) versus 97.6% (95% CI, 95.6 to 98.7); HR, 0.90 (95% CI, 0.47 to 1.74), p=0.76.

Figure 3B.1. PET-adapted strategies: PET-2 negative patients randomized to 4×ABVD vs. 4×AVD after two cycles of ABVD and PET: OS

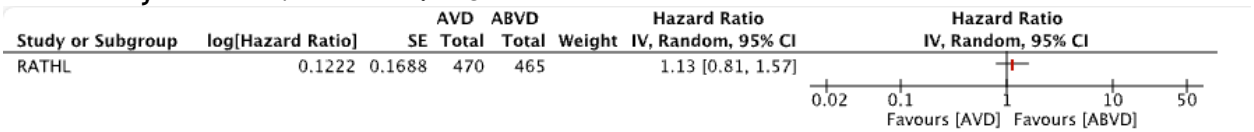


4ABVD=4 cycles of doxorubicin hydrochloride, bleomycin sulfate, dacarbazine; 4AVD= 4 cycles of doxorubicin hydrochloride, vinblastine sulfate, and dacarbazine; SE=standard error.

PFS

The RATHL trial [49] was a noninferiority trial and the prespecified upper boundary margin of the 95% confidence interval of the difference between ABVD and AVD for PFS was five percentage points at three-year follow-up. At three-year (36-month) follow-up [49], in the per-protocol analysis the difference between 4×ABVD and 4×AVD was 1.3% (95% CI, −3.7 to 5.1), HR, 1.10 (95% CI, 0.79 to 1.53), p=0.58, and this result is slightly above the prespecified margin, so the noninferiority condition was not met at this time point. At 87.2 months of follow-up [85] the difference in three-year PFS was 1.3% (95% CI, -3.0 to 4.7), and it fell within the prespecified margin. In the intention-to-treat analysis, PFS rates were 85.7% in the ABVD group, compared with 84.4% in the AVD group, HR 1.13 (95% CI, 0.81 to 1.57).

Figure 3B.2. PET-adapted strategies: PET-2 negative patients randomized to 4×ABVD vs. 4×AVD after two cycles of ABVD and PET: PFS



ABVD=doxorubicin hydrochloride, bleomycin sulfate, dacarbazine; AVD=doxorubicin hydrochloride, vinblastine sulfate, and dacarbazine; SE=standard error.

Adverse effects

TRD

The RATHL trial [49,85] reported a TRD rate of 0.85% in the group receiving 4×ABVD versus 0% in the group receiving 4×AVD; p values were not reported.

Important outcomes

Pulmonary toxicity

The RATHL [49] reported that the absolute difference between the ABVD and the AVD groups in the diffusion capacity of the lung for carbon monoxide (DL_{CO}) was statistically

significant in favour of AVD: −7.4% (95% CI, 5.1 to 9.7; p<0.001) from baseline to the completion of therapy, and it remain significant at one year −4.6% (95% CI, 1.6 to 7.5; p=0.003).

SPM

The RATHL trial [85] reported the cumulative incidence of SPM at seven-year follow-up to be 5.1% (95% CI, 3.2 to 8.1) for patients allocated to ABVD, and 5.8% (95% CI, 3.8 to 9.0) for patients allocated to AVD; p values were not reported.

Infertility

Data on women's infertility are reported by a cohort study [93] ancillary of the RATHL trial [49]. Patients' PET status was not considered in the analyses. Therefore, we will not discuss this study in further details here. The interested reader can find more information about these studies in Table 3-7A, Appendix 7.

Table 8B. Comparison 3B. PET-adapted strategies: PET-2 negative patients randomized to 4×ABVD vs. 4×AVD after two cycles of ABVD and PET.

Certainty assessment							№ of patients		Effect		Certainty of the evidence§	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	4×AVD	4×ABVD	Relative (95% CI)	Absolute (95% CI)		
Overall survival 4×AVD vs. 4×ABVD (follow-up: 36 months; assessed with HR)												
1	RCT RATHL [49]	Serious[a]	Not serious	Not serious	Not serious	None	465	470 / 2.8%	HR 0.90 (0.47 to 1.74)	2 fewer per 1000 (from 15 fewer to 20 more)	Moderate	CRITICAL
Recurrence												
Progression-free survival 4×AVD vs. 4×ABVD (follow-up: 36 moths; assessed with HR)												
1	RCT RATHL [49]	Very serious[b]	Not serious	Not serious	Not serious	None	465	470 / 14.3	HR 1.13 (95% CI, 0.81 to 1.57) [c]	17 more per 1000 (from 25 fewer to 72 more)	Low	CRITICAL
Adverse events												
Treatment-related death 4×AVD vs. 4×ABVD (follow-up: 36 moths; assessed with HR)												
1	RCT RATHL [49]	Very serious[b]	Not serious	Not serious	Serious[d]	None	0/465	4 / 470	RR 0	Not estimable	Very low	CRITICAL
Second primary malignancies 4×ABVD vs. 4×AVD (follow-up median 87.2 months; assessed with RR)												
1	RCT RATHL [85]	Very serious[b]	Not serious	Not serious	Serious[d]	None	27/465	24 / 470	RR 1.14 (0.11 to 11.32)	7 more per 1000 (from 45 fewer to 527 more)	Very low	IMPORTANT

Pulmonary toxicity at during cycles 3 to 6 4×AVD vs. 4×ABVD (follow-up: 36 moths; assessed with RR)												
1	RCT RATHL [49]	Very serious[b]	Not serious	Not serious	Serious[d]	None	3/457	15 / 468	RR 0.205 (0.06 to 0.703)	25 fewer per 1000 (from 30 fewer to 10 fewer)	Very low	IMPORTANT

a Outcome assessors were aware of patients' assignment, and there were deviations from intended interventions because of the trial context.
b Patients, clinicians and outcome were aware of patients' assignment. There were deviations from the intended interventions for 7.4% of patients which most likely affected the outcome. There were missing data.
c Data from ITT analysis.
d Very few events lead to very wide confidence intervals.

ABVD=doxorubicin, bleomycin, vinblastine, and dacarbazine; AVD=doxorubicin, vinblastine, and dacarbazine; HR=Hazard ratio; PET=Positron emission tomography; RCT=Randomized controlled trial; RR=Relative risk

C. PET-adapted strategies: PET-2 negative patients randomized to 4×eBEACOPP vs. 8× or 6×eBEACOPP after two cycles of eBEACOPP and PET

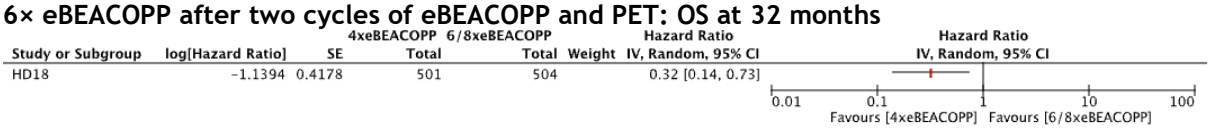
Table 8C presents the relative and absolute effect and the certainty of the evidence for this comparison.

Critical outcomes

OS

The HD18 trial [16] reported in their per-protocol analysis that lower-intensity chemotherapy (i.e., 4×eBEACOPP) was statistically significantly better for PET-2-negative patients than higher-intensity chemotherapy (i.e., 8× or 6×eBEACOPP) at 55 months: 97.7% vs. 95.4%, HR, 0.32 (95% CI, 0.14 to 0.72), p=0.0037 [16]; and at 66 months: 98.1% vs. 95.3%, HR, 0.36 (95% CI, 0.17 to 0.74), p=0.0038 [83].

Figure 3C.1. PET-adapted strategies: PET-2 negative patients randomized to 4×eBEACOPP vs. 8× or 6× eBEACOPP after two cycles of eBEACOPP and PET: OS at 32 months

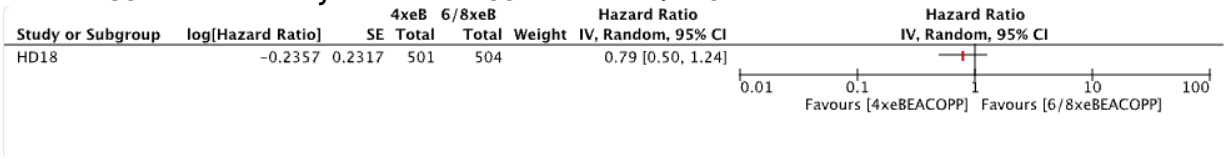


4×eBEACOPP=4 cycles of escalated bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, prednisone; 6/8×eBEACOPP; 6 or 8 cycles of BEACOPP; SE= standard error.

PFS

The HD18 [16] found that 4×eBEACOPP was noninferior to 8×eBEACOPP (or 6×eBEACOPP). Both at 55 and at 66 months follow-up the difference (respectively −1.4% [95% CI, −2.7% to 5.4%] and 1.9% [95% CI, −1.8% to 5.5%]) excluded the prespecified noninferiority margin of -6%; HR, 0.79 (95% CI, 0.50 to 1.24).

Figure 3C.2. PET-adapted strategies: PET-2 negative patients randomized to 4×eBEACOPP vs. 8× or 6× eBEACOPP after two cycles of eBEACOPP and PET: PFS at 55 months



4×eB=4 cycles of escalated bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, prednisone; 6/8×eB=6 or 8 cycles of BEACOPP; SE= standard error.

Adverse effects

At 78 months the HD18 [83] reported a TRD rate of 1% versus 0% for patients treated with 8×eBEACOPP (or 6×eBEACOPP) versus 4×eBEACOPP; p values not reported.

Important outcomes

Pulmonary toxicity

In the HD18 [83], data are reported on a very small number of participants for this outcome at 78-month follow-up. Among the patients for which this outcome was reported, 33 women had a mean DL_{CO} at five years of 75.2% (standard deviation [SD] 20.3) and 37 men had a 84.5% (SD 18.4) DL_{CO}; no comparison is reported, therefore, these data will not be discussed any further here.

SPM

The HD18 trial [83] reported the cumulative incidence of SPM: 3.7% (95% CI, 2.0 to 5.4), for patients treated with the standard regimen (i.e., 8×eBEACOPP or 6×eBEACOPP) versus 3.3% (95% CI, 1.6 to 5.0), HR, 0.87 (95% CI, 0.46 to 1.63), p=0.66.

Infertility

No comparative data on infertility are reported in the HD18 trial [16,83].

Table 8C. Comparison 3: PET-adapted strategies, PET-2 negative patients randomized to 4×eBEACOPP vs. 8×eBEACOPP after two cycles of eBEACOPP and PET*

Certainty assessment							№ of patients		Effect		Certainty of the evidence	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	4× eBEACOPP	8× eBEACOPP	Relative (95% CI)	Absolute (95% CI)		
Overall Survival 4×eBEACOPP vs. 8×eBEACOPP (follow-up median 55 months [16]; assessed with HR)												
1	RCT HD18 [16]	Very High[a]	Not serious	Not serious	Not serious	None	501 (285 8eB + 216 4eB)	504 (288 8×eB) + 216 (4×eB) / 4.6%	HR 0.32 (95% CI, 0.14 to 0.72)	31 fewer per 1000 (from 39 fewer to 13 fewer)	Low	CRITICAL
Progression-free survival												
Progression-free survival 4×eBEACOPP vs. 8×eBEACOPP (follow-up median 32 months [16]; assessed with HR)												
1	RCT HD18 [16]	Very High[a]	Not serious	Not serious	Not serious	None	501 (285 8eB + 216 4eB)	504 (288 8×eB) + 216 (4×eB) / 9.2%	HR 0.79 (95% CI 0.50 to 1.24)	19 fewer per 1000 (from 45 fewer to 21 more)	Low	CRITICAL
Adverse events												
Treatment-related death 4×eBEACOPP vs. 8×eBEACOPP (follow-up median 78 months [122]; assessed with RR)												
1	RCT HD18 [16]	Very High[a]	Not serious	Not serious	Serious[b]	None	0/501	6/504	RR 0	Not estimable	Very Low	CRITICAL
Secondary primary malignancies (cumulative incidence) 4×eBEACOPP vs. 8×eBEACOPP (follow-up median 78 months [122]; assessed with HR)												
1	RCT HD18 [16]	Very High[a]	Not serious	Not serious	Serious[b]	None	501	504/3.7%	HR 0.87 (0.46 to 1.63)	5 fewer per 1000 (from 20 fewer to 23 more)	Very Low	IMPORTANT

*Data reported are from the per-protocol analysis of this study
a Patients, clinicians, and outcome assessors were aware of patients' assignment; there were deviations to the intended interventions because of the trial context.
b Very small number of events.

bBEACOPP=baseline BEACOPP; BEACOPP=bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone; CI=Confidence interval; eBEACOPP=escalated-dose BEACOPP; HR=Hazard ratio; RCT=Randomized controlled trial; RR=Relative risk

Comparison 4: ABVD vs. modified ABVD

The ECHELON-1 trial [9,86-89] represents the body of evidence for this comparison. This trial compared six courses of ABVD with six courses of a modified regimen where brentuximab vedotin was added and bleomycin was withdrawn from the drug combination (A-AVD). ECHELON-1 was a multicentre, industry-funded, phase 3, open label randomized trial. The primary outcome of the study for the primary analysis [9] was modified PFS (i.e., time to disease progression, death, or noncomplete response after completion of frontline therapy followed by subsequent anticancer therapy). In subsequent analyses [89] PFS per investigator was reported (i.e., time from randomization to first documentation of progressive disease or death due to any cause). For this report we considered this latter measure of recurrence.

We identified four corollary publications of the ECHELON-1 trial [61,62,94,95]. Ramchandren et al. [94] reported on efficacy and safety for the subgroup of patients who were treated in North America; Evens et al. [61] publication was the pre-specified analysis of the subgroup of patients 60 years old and older included in the original study; Grigg et al. [95] abstract publication was a 62.9-month follow-up analysis of the adolescents and young adults included in the original trial; and Hutchings et al. conference abstract [62] reported the pre-specified OS analysis at 73 months follow-up, and included younger, as well as patients 60 years old and older, but presented the results separately. We will discuss Evens et al. [61] and Hutchings et al. [62] in the section dedicated to Question 2.

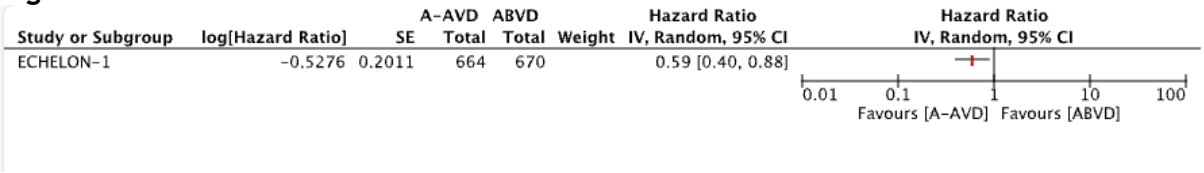
Critical Outcomes

OS

The final analysis for OS after the occurrence of 112 deaths has not been published at the time of our latest update on January 12, 2024.

At 78 months of follow-up, Ansell et al. [89] reported a significantly better OS for patients allocated to the A-AVD regimen compared to ABVD (39 vs. 64 patients died, HR, 0.59 [95% CI, 0.40 to 0.88], p=0.009).

Figure 4.1. ABVD versus modified ABVD. OS at 78 months.



A-AVD=brentuximab vedotin in combination with doxorubicin, vinblastine, and dacarbazine; ABVD=doxorubicin, bleomycin, vinblastine, and dacarbazine; SE=standard error.

Recurrence

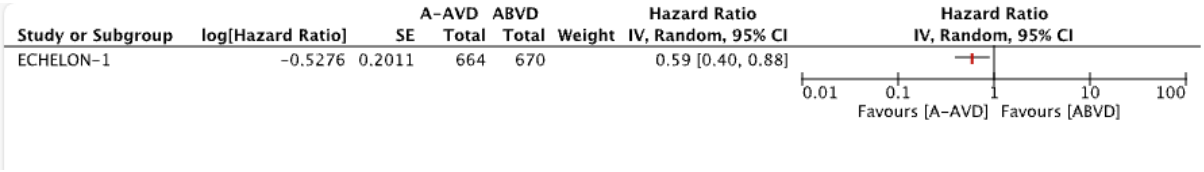
PFS

This outcome was defined as the time from randomization to first occurrence of disease progression or death from any cause per investigator in the intention-to-treat population.

The 78-month estimated [89] PFS rates were statistically significantly better for A-AVD than with ABVD (82.3% vs. 74.5%) (HR for disease progression, 0.68 [95% CI, 0.53 to 0.86]), p values not reported). This result was similar for earlier time points (see Table 1-7A in Appendix 7 for more details). Straus et al. [88] reported results for subgroups at 60.9 months (5-year) follow-up: PFS rates for PET-2 negative patients were 78.9% (95% CI, 75.2 to 82.1) for ABVD compared with 84.9% (95% CI, 81.7 to 87.6) for A-AVD (HR, 0.66; 95% CI, 0.50 to 0.88, p=0.0035). The five-year PFS rates for PET-2-positive patients were 45.9% (95% CI, 32.7 to 58.2) for the ABVD group versus 60.6% (95% CI, 45 to 73.1) for the A-AVD group (HR, 0.70; 95% CI, 0.39 to 1.26, p=0.23). A similar pattern was seen when examining patients of age younger than 60 years (p<0.0034) and those 60 year or older (p=0.44). See Table 1-7A in Appendix 7 for more details.

Among the corollary studies (Table 3-7A in Appendix 7), Ramchandren et al. [94] reported similar results for the subgroup of patients treated in North America as the first publication [9] for modified PFS (progression, death, confirmed non-response): at 24.7 months, 57 for ABVD versus 38 for A-AVD events HR was 0.596 (95% CI, 0.365 to 0.899, p=0.012). Evens et al. [61] presents in detail a subgroup analysis at 60.9 months of people of 60 years of age or older that was briefly mentioned in the Ansell et al. paper [89], and the five-year PFS per investigator assessment was 61.6% (95% CI, 50.9 to 70.7) in the ABVD group versus 67.1% (95% CI, 55.1 to 76.5) in the A-AVD group (HR, 0.820; 95% CI, 0.494 to 1.362, p=0.443). Hutchings et al. in a recent abstract [62] reported similar results for the population of older adults. See Table 3-7A in Appendix 7 and Question 2 section for more details.

Figure 4.2. ABVD versus modified ABVD. PFS at 78 months.



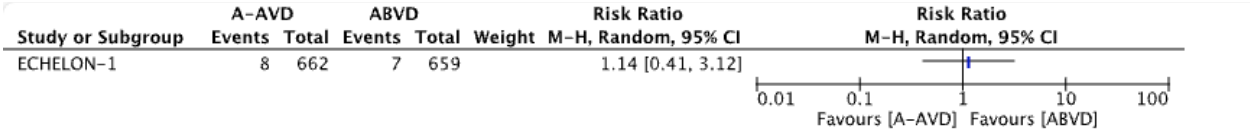
A-AVD=brentuximab vedotin in combination with doxorubicin, vinblastine, and dacarbazine; ABVD= doxorubicin, bleomycin, vinblastine, and dacarbazine; SE=standard error.

Adverse effects

TRD

Connors et al. (2-year interim analysis) [9] reported a TRD of 1% in each group (p values not reported). Similar results were reported by Ansell et al. [89]: eight patients and seven patients in the A-AVD and ABVD groups, respectively, died of drug-related adverse events.

Figure 4.3. ABVD versus modified ABVD. TRD at 78 months.



Important outcomes

Adverse effects

SPM

Ansell et al. [89] at 72.6 months of follow-up reported SPM rates to be 4.9% (32 patients) in the ABVD versus 3.5% (23 patients) in the A-AVD groups (p values not reported).

Pulmonary toxicity

Connors et al. [9] reported that in the A-AVD group (n=662), 12 patients experienced pulmonary toxicity, compared with 44 patients in the ABVD group (n=659).

Infertility

No data for this outcome was reported in this study.

Table 9. Comparison 4: ABVD versus modified ABVD with brentuximab instead of bleomycin

Certainty assessment							№ of patients		Effect		Certainty of the evidence	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	6×A-AVD	6×ABVD	Relative (95% CI)	Absolute (95% CI)		
Overall survival 6×ABVD vs. 6×A-AVD (follow-up median 72.6 months; assessed with HR)												
1	RCT ECHELON-1 [89]	Serious[a]	Not serious	Not serious	Not serious	None	664	670/10.0%	HR 0.59 (0.40 to 0.88)	40 fewer per 1000 (from 50 fewer to 11 fewer)	Moderate	CRITICAL
Progression-free survival												
Progression-free survival 6×ABVD vs. 6×A-AVD (follow-up median 72.6 months; assessed with HR)												
1	RCT ECHELON-1 [89]	Very serious[b]	Not serious	Not serious	Not serious	None	664	670/25.5%	HR 0.68 (0.53 to 0.86)	74 fewer per 1000 (from 111 fewer to 31 fewer)	Low	CRITICAL
Adverse events												
Treatment-related death 6×ABVD vs. 6×A-AVD (follow-up median 72.6 months; assessed with RR)												
1	RCT ECHELON-1 [89]	Very serious[b]	Not serious	Not serious	Serious[c]	None	8/662	7/659	RR 1.14 (0.41 to 3.12)	1 more per 1000 (from 6 fewer to 23 more)	Very low	CRITICAL
Secondary primary malignancies 6×ABVD vs. A-6×AVD (follow-up 7 median 6.2 months; assessed with HR)												
1	RCT ECHELON-1 [89]	Very serious[b]	Not serious	Not serious	Serious[c]	None	23/662	32/659	RR 0.72 (0.05 to 10.07)	14 fewer per 1000 (from 46 fewer to440 more)	Very low	IMPORTANT
Pulmonary toxicity 6×ABVD vs. A-6×AVD (follow-up median 24.6 months; assessed with HR)												

Certainty assessment							№ of patients		Effect		Certainty of the evidence	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	6×A-AVD	6×ABVD	Relative (95% CI)	Absolute (95% CI)		
1	RCT ECHELON-1 [9]	Very serious[b]	Not serious	Not serious	Serious[c]	None	12/662	44/659	RR 0.27 (0.00 to 20.17)	49 fewer per 1000 (from 0 to 1,000)	Very low	IMPORTANT

a No information about how randomization was conducted, about whether deviations from the intended interventions were present.
b In addition to what in a), patients, clinicians, and outcome assessors were aware of patients' assignment.
c Very few events.

A-AVD=Brentuximab vedotin plus doxorubicin, vinblastine, and dacarbazine; ABVD=doxorubicin, bleomycin, vinblastine and dacarbazine; CI=Confidence interval; HR=Hazard ratio; RR=Relative risk.

Comparison 5: Modified ABVD vs. another modified ABVD

No fully published, final, results of any studies for this comparison are available as of January 12, 2024. Our searches captured two abstracts of interim analyses of the randomized, phase 3, SWOG S1826 trial [23,123], which was still active at the time of our latest search, and is estimated to be completed in April 2025. This study included children, adults, and older adults; therefore, it will be relevant for both our Question 1, and Question 2. The comparison was nivolumab combined with AVD (doxorubicin, vinblastine, and dacarbazine) (N-AVD) versus brentuximab vedotin (A) combined with AVD (A-AVD). See Table 11, Ongoing trials for more information on these abstracts.

Comparison 6: BEACOPP vs. modified BEACOPP

No fully published, final, results of any studies for this comparison are available as of January 12, 2024. We are aware of HD21 trial [22] that is at least partly still ongoing, and it is expected to be completed in September 2025. This study compares PET2-guided four to six cycles of either standard eBEACOPP (escalated doses of bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, prednisone) or experimental BrECADD (brentuximab vedotin, etoposide, cyclophosphamide, doxorubicin, dacarbazine, dexamethasone) treatment, and it is composed by two parts: tolerability, and efficacy. The tolerability part has been completed, and we included a published abstract of the final analysis [22]. See Table 1-7A in Appendix 7 for the results of this abstract.

Question 3: Radiotherapy consolidation

Studies relevant to Question 3 can be grouped into 2 comparisons: 1) Consolidation radiotherapy versus observation, and 2) Consolidation radiotherapy versus chemotherapy.

Comparison 1: Consolidation radiotherapy vs. observation

Four unique RCTs of patients with advanced-stage HL met the inclusion criteria for this comparison: the FIL HD0801 [75], the GITIL/FIL HD0607 [60], the EORTC 20884 [76], and the HD12 [59,77].

The HD0801 FIL [75], a phase 3 randomized trial, compared consolidation radiotherapy in PET-2-negative patients treated with six cycles of ABVD with observation; the primary outcome of this study was EFS where relapse, second cancers and death were considered events. The GITIL/FIL HD0607 [60], a phase 2, two-stage design trial, randomized PET-2-negative patients with a mass ≥five cm to radiotherapy consolidation or no further treatment; the primary outcome was three-year PFS. The EORTC 20884 [76] randomized patients who were in partial or complete remission after six or four cycles of chemotherapy (MOPP-ABV) to no further treatment or involved field radiotherapy (IF-RT). The primary outcome was three-year relapse-free survival (RFS).The HD12 [59] was a noninferiority, two-by-two factorial design RCT. The chemotherapy arm is discussed in the Question 1 section. In the radiotherapy arm patients who responded to chemotherapy and had an initial bulky disease (≥5 cm) or residual disease (1.5 cm) were randomized to consolidation radiotherapy or no radiotherapy. The primary outcome was FFTF.

We pooled in meta-analysis two studies, the HD0801 FIL [75] and the GITIL/FIL HD0607 [48,60] for PFS. These studies were clinically homogeneous, and, although executed in the same sites and at the same time, the patients were not allowed to participate in both studies (Levis, Gallamini, personal communication, October 16, 2023).

Table 10A reports the relative, the absolute effect, and the certainty of the evidence for this comparison. Table 2-7A in Appendix 7 reports additional details for the results of studies included for Question 3.

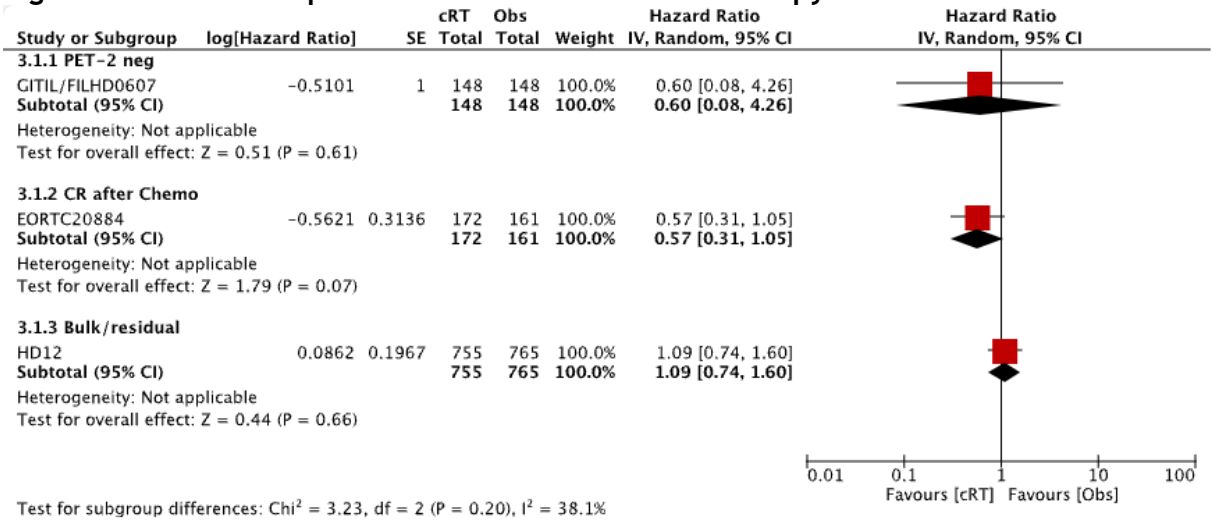
Critical outcomes

OS

The HD0801 FIL [75] did not report results for OS, although this outcome was listed in the protocol. We contacted the authors on October 25, 2023. No response was received as of February 20, 2024. The GITIL/FIL HD0607 [60], at 72-month follow-up, the EORTC 20884 [76] at 96-month follow-up, and the HD12 [59] at both 78- and 97-month reported no statistically significant difference between radiotherapy and no further treatment.

The certainty of the evidence for this outcome was low because of risk of bias and imprecision.

Figure 5A. OS for Comparison 1: Consolidation radiotherapy vs. observation.



CI=Confidence interval; CR=Complete remission; cRT=Consolidation radiotherapy; Obs=Observation; PET=Positron emission tomography

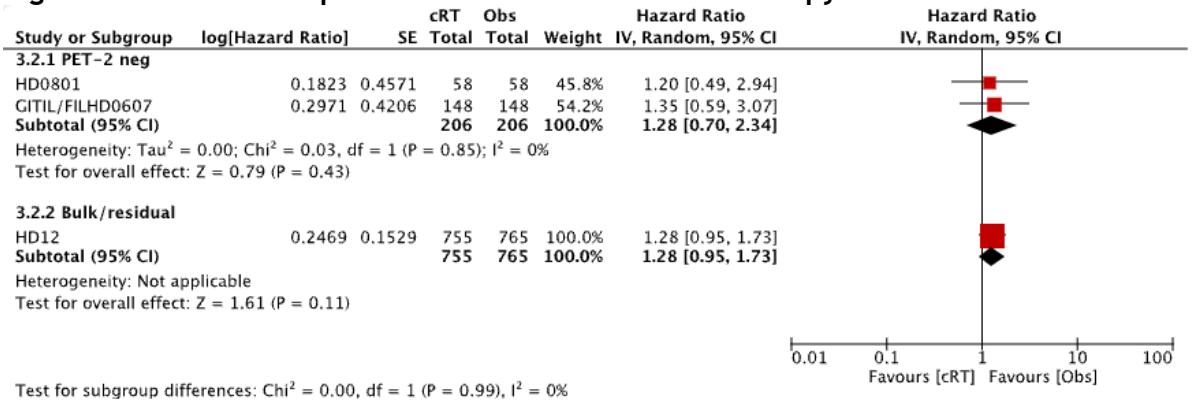
Recurrence

PFS

Our meta-analysis of the HD0801 FIL [75] and the GITIL/FIL HD0607 [48,60] trials showed no statistically significant difference between consolidation RT compared with observation for PFS (HR, 1.28; 95% CI, 0.70 to 2.34).

The HD12 [59], at 97-month follow-up reported a better PFS rate for radiotherapy versus no radiotherapy: 86.8% (95% CI, 84.1 to 89.6) versus 82.2% (95% CI, 79.0 to 85.4) (HR, 1.34; 95% CI, 1.02 to 1.75) with the upper limit of the 95% CI being just above the non-inferiority margin of 1.74.

Figure 5B. PFS for Comparison 1: Consolidation radiotherapy vs. observation.



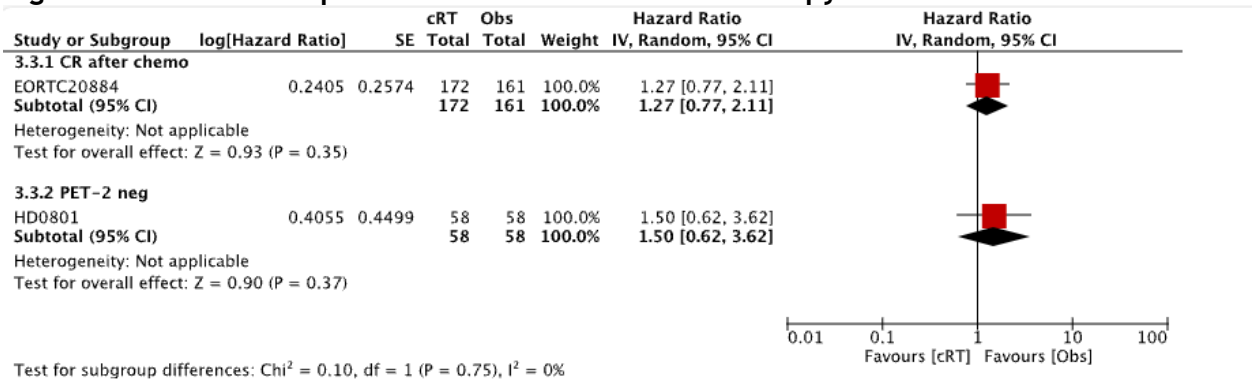
CI=Confidence interval; cRT=Consolidation radiotherapy; Obs=Observation; PET=Positron emission tomography; PFS=Progression-free survival

EFS

The HD0801FIL trial [75] reported no statistically significant difference at two-year follow-up in patients who were PET-2 negative after chemotherapy; the EFS rates were 87.8% for radiotherapy compared with 85.8% for observation (HR, 1.5; 95% CI, 0.6 to 3.5, p=0.34). The

EORTC 20884 at five-year follow-up [121] reported no statistically significant difference between no treatment and consolidation radiotherapy (84% vs. 79%, p=0.35) in patients who were in complete remission after chemotherapy.

Figure 5C. EFS for Comparison 1: Consolidation radiotherapy vs. observation.

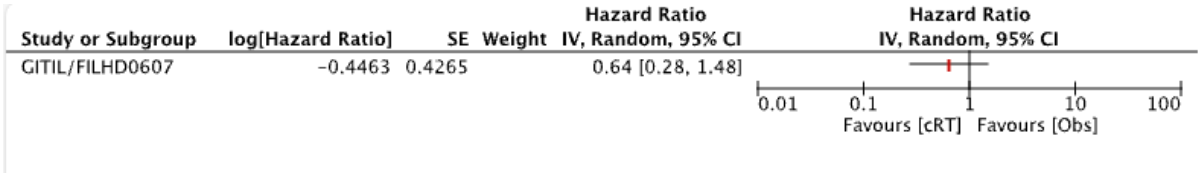


CI=Confidence interval; CR=Complete remission; cRT=Consolidation radiotherapy; Obs=Observation; PET=Positron emission tomography

DFS

The GITIL/FIL HD0607 [60] at five-year follow-up reported no difference in DFS from the time of complete remission (CR) rate between no further treatment and consolidation radiotherapy (91% [95% CI, 86% to 96%] vs. 94% [95% CI, 89% to 98%], p values not reported).

Figure 5D. DFS for Comparison 1: Consolidation radiotherapy vs. observation.

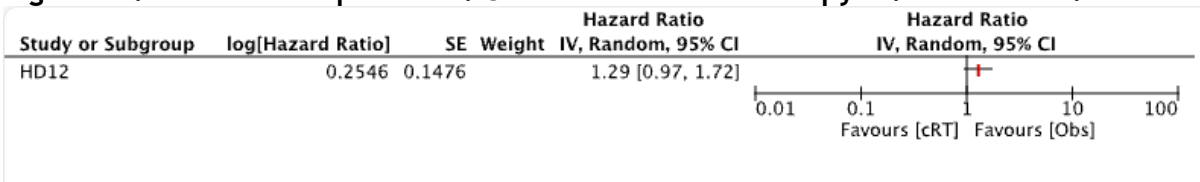


CI=Confidence interval; cRT=Consolidation radiotherapy; DFS=Disease-free survival; Obs=Observation

FFTF

The HD12 [59] reported no statistically significant difference in FFTF at 78 months of follow-up (90.4% vs. 87%, p=0.08).

Figure 5E. FFTF for comparison 1: Consolidation radiotherapy vs. observation.



CI=Confidence interval; cRT=Consolidation radiotherapy; FFTF=Freedom from treatment failure; Obs=Observation

We considered certainty of the evidence for recurrence as moderate to low; the included studies for the outcomes in this group were of moderate to very low certainty of the evidence; however, they consistently showed no difference between consolidation radiotherapy

compared with observation across different patient populations and different levels of certainty of evidence for the individual measures.

Adverse events

Critical outcomes

TRD

None of the trials reported this outcome in relation to radiotherapy treatment.

Important outcomes

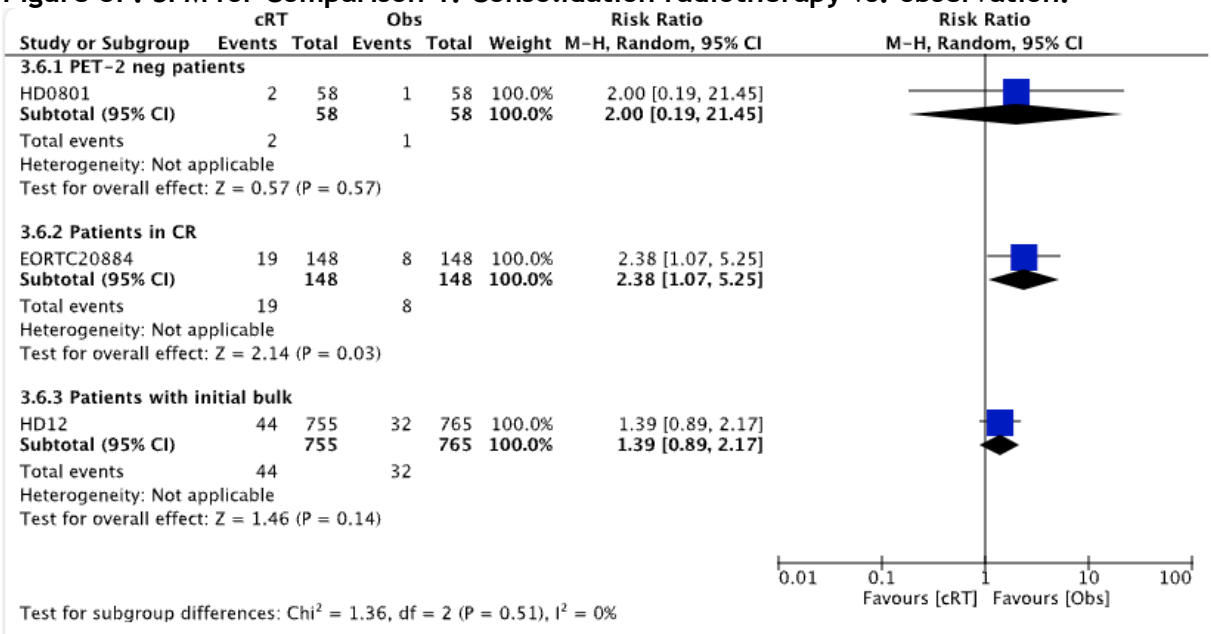
SPM

The HD0801 FIL [75] reported a SPM rate of 3.4% in the consolidation radiotherapy and of 1.7% in the observation group in PET-2-negative patients (p values were not reported).

The EORTC 20884 [76] reported the cumulative SPM rates and compared patients who were in partial remission and had radiotherapy (5.7%) versus patients in complete remission who did not have radiotherapy (5.6%), and patients in complete remission who had radiotherapy (12.9%); the comparison among the three groups was statistically significant (p=0.0177). We reported only the comparison of patients in complete remission in Figure 5F and in Table 10A.

The HD12 [59] reported a SPM rate of 5.8% in the group assigned to radiotherapy and 4.2% in the group assigned to no radiotherapy.

Figure 5F. SPM for Comparison 1: Consolidation radiotherapy vs. observation.



CI=Confidence interval; CR=Complete remission; cRT=Consolidation radiotherapy; Obs=Observation; SPM=Second primary malignancy

We considered the certainty of the evidence for adverse events, both for the critical and for the important outcomes to be low or very low because of risk of bias, inconsistency among the different subpopulations, and imprecision.

Table 10A. Comparison 1. Consolidation radiotherapy versus observation

Certainty assessment							№ of patients		Effect		Certainty of the evidence	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RT	Obs	Relative (95% CI)	Absolute (95% CI)		
Overall survival												
Overall survival HD0801FIL - PET-2-negative patients (follow-up median 71 months; assessed with HR)												
1	RCT HD0801 FIL [75]	Serious[a]	Serious[a]	Very serious[a]	Serious[a]	none	58	58/0%	Not estimable	Not estimable	Very low	CRITICAL
Overall survival GITIL/FIL PET-2-negative patients (follow-up median 70.8 months; assessed with HR)												
1	RCT GITIL/FIL HD0607 [60]	Serious[b]	Not serious	Not serious	Serious[c]	none	148	148/2%	HR 1.665 (0.235 to 11.823)	13 more per 1000 (from 15 fewer to 192 more)	Low	CRITICAL
Overall survival EORTC 20884 - Patients in partial or complete remission after chemotherapy (follow-up median 79 months; assessed with HR)												
1	RCT EORTC 20884 [76]	Very serious[d]	Not serious	Not serious	Serious[e]	none	172	161/10%	HR 0.57 (0.31 to 1.06)	42 fewer per 1000 (from 68 fewer to 6 more)	Very low	CRITICAL
Overall survival HD12 - patients with bulky disease who responded to chemotherapy (follow-up median 78 months; assessed with HR)												
1	RCT HD12 [59]	Not serious	Not serious	Not serious	Serious[e]	none	755	765/7.1%	HR 1.09 (0.74 to 1.60)	6 more per 1000 (from 18 fewer to 40 more)	Moderate	CRITICAL

Progression-free survival												
Progression-free survival (follow-up: median 71 months; assessed with HR)												
2	RCTs HD0801 FIL [75] and GITIL/FIL HD0607 [48,60]	Serious[f]	Not serious	Not serious	Serious[e]	None	206	206/10%	HR 1.28 (0.70 to 2.34)	26 more per 1000 (from 29 fewer to 119 more)	Moderate	CRITICAL
Progression-free survival (follow-up 78 months; assessed with HR)												
1	RCT HD12 [59]	Serious[g]	Not serious	Not serious	Serious[e]	None	755	765/10%	HR 1.28 (0.95 to 1.73)	26 more per 1000 (from 5 fewer to 67 more	Moderate	CRITICAL
Event-free survival												
Event-free survival: Patients in complete remission after chemotherapy (follow-up: median 79 months; assessed with: HR)												
1	RCT EORTC20884 [121]	Very serious[d]	Not serious	Not serious	Serious[e]	None	172	161/16%	HR 1.272 (0.768 to 2.106)	39 more per 1000 (from 35 fewer to 147 more)	Very low	CRITICAL
Event-free survival: PET-2-negative patients (follow-up: median 24 months; assessed with: HR)												
1	RCT HD0801FIL trial [75]	Serious[g]	Not serious	Not serious	Serious[e]	None	58	58/14%	HR 1.50 (0.6 to 3.5)	62 more per 1000 (from 53 fewer to 27 more)	Low	CRITICAL
Disease-free survival												
Disease-free survival (follow-up: median 71 months; assessed with HR)												
1	RCT GITIL/FIL HD0607 [60]	Serious[d]	Not serious	Not serious	Serious[e]	None	148	148/6%	HR 0.64 (0.28 to 1.49)	21 fewer per 1000 (from 43 fewer to 28 more)	Low	CRITICAL

Freedom from treatment failure												
Freedom from treatment failure (follow-up: median 78 months; assessed with HR)												
1	RCT HD12 [59]	Serious[g]	Not serious	Not serious	Serious[e]	None	787	787/13%	HR 1.29 (0.97 to 1.73)	34 more per 1000 (from 4 fewer to 84 more)	Low	CRITICAL
Adverse events												
Treatment-related death:												
0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	Very low	CRITICAL
Second primary malignancies: PET-2 negative patients (follow-up 71 months; assessed with RR)												
1	RCT HD0801 FIL [75]	Serious[g]	Not serious	Not serious	Serious[e]	None	2/58 (3.4%)	1/58 (1.7%)	RR 2.00 (0.18 to 21.83)	17 more per 1000 (from 14 fewer to 359 more)	Low	IMPORTANT
Second primary malignancies: Patients in complete remission after chemotherapy (follow-up 79 months; assessed with RR)												
1	RCT EORTC 20884 [76]	Very serious[d]	Not serious	Not serious	Not serious	None	19/148	8/148	RR 2.375 (1.073 to 5.254)	74 more per 1000 (from 4 more to 230 more)	Low	IMPORTANT
Second primary malignancies: Patients with initial bulk or residual tumour												
1	RCT HD12 [59]	Very serious[d]	Not serious	Not serious	Not serious	None	44/755	32/765	RR 1.39 (0.89 to 2.17)	16 more per 1000 (from 5 fewer to 49 more)	Very low	IMPORTANT

NA=not applicable
a This study did not provide results for OS
b No information was provided on whether allocation was concealed.
c Very few events, very wide confidence intervals.
d Patients, clinicians and outcome assessors were not blinded to patients' assignment. Patients were not analyzed in the groups where they were randomized. There were missing data.
e The confidence interval of the point estimate crossed 1.
f The trials were open label, and no information was reported on whether the allocation was concealed.
g Patients, clinicians, and outcome assessors were aware of patients' assignment.
CI=Confidence interval; HR=Hazard ratio; Obs=Observation; PET=Positron emission tomography; RCT=Randomized controlled trial; RR=Relative risk; RT=Radiotherapy

Comparison 2: radiotherapy vs. chemotherapy

Two unique RCTs, the ECOG E1476 [78], and the Aviles et al. trial [79], were included for this comparison. These studies were clinically heterogeneous; therefore, statistical pooling of the results was inappropriate.

The ECOG E1476 [78] randomized patients to consolidation radiotherapy or to three cycles of ABVD. Radiotherapy dose was 15-200 Gy given at a rate of 15-20 Gy per fraction, and five fractions per week. Patients older than 60 years received 50% of the full dose of consolidation chemotherapy in the first cycle, then the full dose if tolerated.

Aviles et al. [79] compared chemotherapy alone (i.e., epirubicin, bleomycin, vinblastine and dacarbazine [EBVD]) with EBVD plus radiotherapy. Radiotherapy was given with a tumour dose of 35 Gy delivered in daily fractions of 1.25 Gy over a total of 4 weeks. Response was the primary outcome in both these studies.

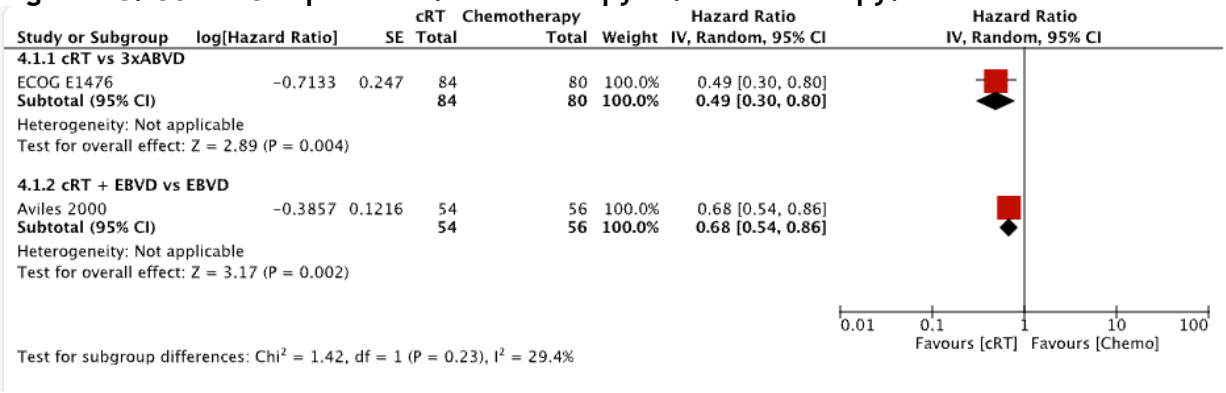
Critical outcomes

OS

The ECOG E1476 [78] reported that the 20-year estimate for OS was significantly better for patients allocated to chemotherapy versus radiotherapy (66% vs. 43%, p=0.002).

In the Aviles et al. trial [79], OS rates at five years were statistically significantly better for radiotherapy plus chemotherapy than with chemotherapy alone: 89% (48 of 54 patients) versus 61% (34 of 56 patients), p<0.01.

Figure 5G. OS for Comparison 2: Radiotherapy vs. chemotherapy.



ABVD=doxorubicin, bleomycin, vinblastine, and dacarbazine; CI=Confidence interval; cRT=Consolidation radiotherapy; EBVD=epirubicin, bleomycin, vinblastine and dacarbazine; OS=Overall survival

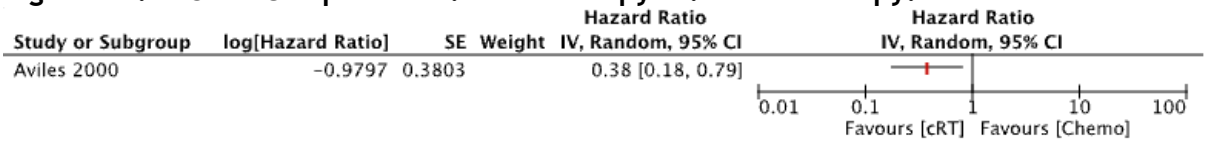
The certainty of the evidence for this comparison was moderate because of risk of bias, and because the included studies were inconsistent in the direction of their results.

Recurrence

FFS

Aviles et al. [79] at 66-month follow-up reported a statistically significant difference in favour of combination therapy (i.e., consolidation radiotherapy to anatomical sites of bulky disease plus 6 cycles of EBVD) compared with chemotherapy alone. FFS rates were 83% and 50% for consolidation radiotherapy plus chemotherapy and chemotherapy alone, respectively, p<0.01.

Figure 5H. FFS for Comparison 2: Radiotherapy vs. chemotherapy.



Chemo=chemotherapy; CI=Confidence interval; cRT=Consolidation radiotherapy; FFS=Failure-free survival

The certainty of the evidence for recurrence for the comparison radiotherapy consolidation versus chemotherapy was low because of risk of bias. Additionally one study [79] reported on this outcome and its sample was relatively small.

Adverse effects

Critical outcomes

TRD

The Aviles et al. trial [79] reported no TRD in either group. The ECOG E1476 trial [78] reported one patient per group had this adverse event.

Important outcomes

SPM

The Aviles et al. trial [79] reported no SPM in either group. The certainty of the evidence for critical and important adverse events for this comparison is low because of risk of bias and imprecision.

Outcomes that were not important

Pulmonary toxicity

The Aviles et al. trial [79] reported that seven of 54 patients in the radiotherapy plus chemotherapy group experienced pulmonary toxicity (12.9%) compared with none in the chemotherapy-alone group (p values were not reported).

Table 10B. Comparison 2. Radiotherapy versus chemotherapy

Certainty assessment							№ of patients		Effect		Certainty of the evidence	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	RT	Chemo	Relative (95% CI)	Absolute (95% CI)		
Overall survival cRT vs. 3×ABVD (follow-up: median 256 months; assessed with HR)												
1	RCT ECOG E1476 [78]	Very serious[a]	Serious[d]	Not serious	Not serious	None	84	80/66%	HR 0.49 (0.30 to 0.79)	249 fewer per 1000 (from 384 fewer to 6 fewer)	Low	CRITICAL
Overall survival cRT + Chemotherapy vs. Chemotherapy alone (follow-up: median 66 months; assessed with HR)												
1	RCT Aviles et al. trial [79]	Serious[b]	Serious[d]	Not serious	Not serious	None	54	56/39%	HR 2.65 (1.26 to 5.55)	340 more per 1000 (from 74 more to 546 more)	Low	CRITICAL
Failure-free survival												
Failure-free survival (follow-up: median 66 months; assessed with HR)												
1	RCT Aviles et al. trial [79]	Serious[b]	Not serious	Not serious	Serious[f]	None	54	56/50%	HR 0.38 (0.18 to 0.79)	268 fewer per 1000 (from 383 fewer to 78 fewer)	Low	CRITICAL
Adverse events												
Treatment-related death (follow-up 66 months; assessed with RR)												
1	RCT Aviles et al. trial [79]	Serious[b]	Not serious	Not serious	Serious[e]	None	0/54	0/56	RR Not estimable	Not estimable	Low	CRITICAL
1	RCT ECOG E1476 [78]	Very serious[a]	Not serious	Not serious	Serious[c]	None	1/84	1/80	RR 0.95 (0.06 to 14.97)	1 fewer per 1000 (from 12 fewer to 175 more)	Very low	CRITICAL

Secondary primary malignancies												
1	RCT Aviles et al. trial [79]	Serious[b]	Not serious	Not serious	Serious[e]	None	0/54	0/56	RR Not estimable	Not estimable	Low	IMPORTANT

a This was an open label trial; data were missing for 14% of patients. The ascertainment of the outcome was performed later for those who were initially in partial remission and converted to complete remission.
b Patients, clinicians, and outcome assessors were aware of patients' assignment.
c The confidence interval of the point estimate crossed 1.
d The studies results for OS pointed in different directions.
e With zero events in both group the confidence interval is wide, therefore imprecision is serious.
f Only one study report on this outcome, and the sample of this study is relatively small

ABVD=doxorubicin, bleomycin, vinblastine, and dacarbazine; CI=Confidence interval; cRT=Consolidation radiotherapy; FFS=Failure-free survival; HR=Hazard ratio; RCT=Randomized controlled trial; RR=Relative risk

Question 2

We included evidence for ABVD, brentuximab vedotin alone or in combination, and VEPEMB (vinblastine, cyclo- phosphamide, prednisolone, procarbazine, etoposide, mitoxantrone, and bleomycin). No studies that analyzed the effects of CHOP, CHLVPP, or any other treatments met our inclusion criteria.

Systematic reviews

One systematic review was identified which included younger and older patients [47], and it was excluded as considered at critical risk of bias. See Table 1-5A in Appendix 5.

Primary Literature

Literature Search Results

RCTs with a sample of ≥100 patients

No RCTs with a sample size ≥100 patients were identified for the population of older adults.

RCTs with a sample ≥30 patients

One RCT [124] with a sample of 54 older patients was identified comparing VEPEMB with ABVD. However, less than 80% (i.e., 69%) of participants had advanced disease, and the results were not presented separately; therefore, it was excluded.

Observational controlled trials with sample ≥30

We identified 915 citations from a search of MEDLINE, EMBASE, the reference lists of included trials, and from reviewing the trials that were identified by the search for RCTs, and that were excluded because observational and/or too small. After reviewing the titles and the abstracts against the selection criteria, the methodologist (FGB) included 47 articles of comparative trials. We retrieved the full text of these articles in the library and the methodologist screened them against the selection criteria and none of these studies was deemed eligible for inclusion. Reasons for exclusion are reported in the study flow-chart (Appendix 3C).

Subgroups analyses of RCTs included for Questions 1 and 3

Two studies, the ECHELON-1 [9] and the HD9 [54,64,81] reported subgroup analyses [61-63,88] of older patients with advanced-stage HL.

Among the ancillary studies of the ECHELON-1 trial [88], Evens et al. 2022 [61], in a non-preplanned analysis, reported outcomes of patients 60 years of age or older (median age 68 years, range 60-82 years) with stage III and IV HL, treated in the main study with A-AVD (n=84) versus ABVD (n=102).

The abstract by Hutchings et al. [62] reported the pre-specified five-year OS analysis for a subgroup of patients 60 years old or older. This ancillary publication [62] was a conference abstract of a pre-planned subgroup analysis. The full publication of this 2022 abstract was not identified as of our latest update search (January 12, 2024) because the final analysis for OS of the ECHELON-1 [88] has not been conducted as the 112 deaths pre-specified for the final analysis have not yet occurred. Therefore, this ancillary publication is considered an abstract of an ongoing trial, and we will not discuss it any further.

Among the companions of the HD9 [54,64,81], Ballova et al. 2005 [63] compared the standard COPP-ABVD regimen with bBEACOPP in 75 older patients: 26 treated with a standard COPP/ABVD, and 42 with baseline BEACOPP (bBEACOPP).

Additional details of these trials are reported in Table 3-7A in Appendix 7.

Critical outcomes

OS

Ballova et al. [63] reported no statistically significant difference between at 80-month follow-up in older patients between the COPP/ABVD and the bBEACOPP groups. Fourteen of 26 patients (54%) died in the COPP/ABVD and 23 of 42 (55%) died in the bBEACOPP group. The certainty of the evidence is low because of risk of bias and imprecision.

Recurrence

PFS

At 60.9-month follow-up Evens et al. [61] reported no statistically significant difference in PFS between the ABVD and the A-AVD groups in patients 60 years old or older. PFS rates were 61.6% for the ABVD compared with 67.1% for the A-AVD group (HR, 0.82; 95% CI, 0.494 to 1.362, p=0.443).

The certainty of the evidence for this outcome is low. The certainty of the evidence was low for this outcome in the main study. The Evens et al. trial [61] was the publication of a pre-specified subgroup analysis that included 102 in the ABVD and 84 older patients in the A-AVD groups respectively. The results for this subgroup pointed in a different direction than the results of the main study for this outcome.

FFTF

Ballova et al. [63] reported no statistically significant difference between the COPP/ABVD and the bBEACOPP groups at five-year follow-up (55% for COPP/ABVD, and 76% for bBEACOPP, p=0.13).

Quality of life

None of the included subgroup analyses reported on quality of life for older patients.

Adverse events

TRD

Evens et al., [61] reported TRD rates of 5.1% in the ABVD group and 4.4% in the A-AVD group (p values were not reported).

Important outcomes

Pulmonary toxicity

Evens et al., [61] reported on pulmonary adverse events of any grade in older patients included in the ECHELON-1 trial; in the ABVD group the rate was 13%, and in the A-AVD group it was 2%. In the ABVD group this adverse event was associated with three of the five TRD.

Certainty of the evidence

The Evens et al. 2022 [61] trial, for the outcomes that were considered critical or important for Question 2, including PFS per investigator assessment and safety, reported post-hoc, exploratory analyses. Considering this was not a pre-planned analysis, and considering that the original ECHELON-1 study [9,86-89] on which it is based was rated at some concerns for risk of bias for OS, and at high risk of bias for all other outcomes, we would consider this evidence of very low certainty.

The [63] subgroup analysis was based on the HD9 trial [54,64,81]. This was a pre-planned analysis of a study that we rated at high risk of bias for all outcomes. The standard therapy arm (COPP/ABVD) was closed prematurely at the second interim analysis, and from then on, the patients were assigned to the bBEACOPP arm. The results of this subgroup analysis are consistent with those of the parent study for OS. Ballova et al. [63] reported no statistically

significant difference at five-year follow-up for FFTF, while in the HD9 study FFTF was not reported at five years and it was better for bBEACOPP at 10 years for younger patients, but, consistent with Ballova et al. [63], no statistically significant difference was reported for a small subgroup of 64 patients aged 60 to 65 years. Ballova et al. [63] reported a toxicity-related death (2 patients [8%] in the ABVD group and 9 patients [21%] in the bBEACOPP group) and SPM rates (8% versus 10%, respectively) at 6.6 years, but p values and confidence intervals were not reported. It is unlikely that chance alone can explain this difference in the subgroup of older patients because older people were more susceptible to side effects that lead to dose reductions. According to the guidance provided by Sun et al. [38] we considered the certainty of the evidence provided by Ballova et al. [63] to be of low certainty.

In summary, the certainty of the evidence for older patients examined in subgroup analyses of included RCTs is low to very low. The main studies were at high risk of bias for the critical and important outcomes, and the samples of older people were small; therefore, the results were largely imprecise.

Phase 2 single-arm studies

The search of the existing database for phase 2 single-arm studies in a population of older patients identified 102 potentially relevant publications.

The methodologist (FGB) selected 11 studies for full-text review after reviewing the titles and the abstracts against the selection criteria. After full-text review, six studies were included [65-68,71,72]. Four additional studies [69,70,73,74], were initially excluded because in abstract form, and were subsequently fully published and included, for a total of 10 included publications.

Table 6-7A in Appendix 7 presents the general characteristics and summary results of the single-arm phase 2 studies included for Question 2 grouped in three intervention types: ABVD/ABVD-like, Brentuximab vedotin alone or in combination, and Other treatments. The excluded abstracts that have not been fully published after the date of our latest update (January 12, 2024) are reported in Table 7-7A in Appendix 7.

ABVD/ABVD-like treatment

A retrospective chart review by Yildiz et al. [65] reported on 51 patients; of these, 45 (88.2%) were treated with ABVD, five with AVD (5.9%), one with bendamustine/brentuximab (2%), one with ABD (2%), one with gemcitabine, cyclophosphamide, vincristine, dexamethasone (2%), and 12 with consolidation radiotherapy (23.56%).

Another retrospective analysis by Wrobel et al. 2019 [66] reported on outcomes of 149 patients with advanced-stage disease: 86% of these patients were treated with ABVD, 7% with CHOP/PVAG, 3% with BEACOPP, and 4% received palliative care. The results are not presented by treatment, and the numbers of patients on CHOP, PVAG, BEACOPP, and palliative care were <80% of the total sample. Therefore, we presented this trial under the ABVD treatment, and this evidence is partially indirect.

Cokgezer et al. [74] retrospectively reviewed the charts of 49 patients of age 60 or greater with AVD and mini-CHOP (400 mg/m^2 cyclophosphamide, 25 mg/m^2 doxorubicin, 1.4 mg/m^2 vincristine, and 40 mg/m^2 prednisolone)

OS

In the Yildiz et al. 2021 noncomparative study [65], at five-years follow-up, OS rate in the advanced-stage population was 54.4%. In the Wrobel et al. 2019 trial [66] the three-year OS rate for older patients treated with ABVD was 80% (95% CI, 71% to 90%). Cokgezer et al. [74] reported that age 60 years or more was a predictor of decreased OS.

Recurrence (PFS and other measures)

Wrobel et al. 2019 [66] reported, for patients treated with ABVD, a three-year EFS rate of 54% (95% CI, 43% to 66%).

Adverse events

In the Yildiz et al. [65] study, 2% of all patients died of treatment-related adverse events, 2% refused treatment, and 29.4% of patients had the dose reduced because of treatment-related toxicity. The results for adverse events in this study include both patients with early and advanced stage combined. In the Wrobel et al. 2019 study [66], seven patients (6%) died of treatment-related adverse events. It was not reported in this study which treatments caused the fatal toxicity.

Brentuximab vedotin alone or in combination

A phase 2 trial by Evens et al. [67] reported on outcomes of sequential brentuximab vedotin given before and after AVD in 48 older patients with advanced-stage HL. Another phase 2 trial by Gibb et al. 2021 [68], the BREVITY trial, tested brentuximab vedotin alone in a sample of 35 older patients with advanced stage HL. Friedberg et al. [69], in a phase 2 noncomparative trial, treated 22 patients with frontline brentuximab vedotin in combination with dacarbazine or with nivolumab. Cheson et al. [70] in a single arm, phase 2 multicentre trial evaluated the safety and efficacy of first-line brentuximab vedotin and nivolumab in 46 patients of 71.5 years of age.

OS

Evens et al., 2018 [67] reported a two-year OS of 93% (95% CI, 80% to 98%). In Friedberg et al. [69] median OS at 63.6-month follow-up was not reached; OS rate was 95% (95% CI, 77.2 to 99.9). In Cheson et al. [70], at 21.2-month follow-up median OS was not reached.

Recurrence (PFS and other measures)

Evens et al., 2018 [67] reported a two-year PFS rate of 84% (95% CI, 69% to 92%) and a two-year EFS of 80% (95% CI, 65% to 89%). Gibb et al. 2021 [68] reported a median PFS of 7.3 months (95% CI 5.2 to 9.0). Friedberg et al. [69] reported median PFS at 63.2-month follow-up was 47.2 months for the 22 patients treated with brentuximab vedotin in combination with dacarbazine. For patients in the cohort treated with the brentuximab vedotin/nivolumab combination, median PFS had not been reached at 63.6-month follow-up. In Cheson et al. [70], at 21.2-month follow-up, median PFS was 18.3 months (95% CI, 12.7-not reached).

Adverse events

The BREVITY trial [68] reported a treatment-related mortality of 9%. Fifty-one per cent of the patients experienced grade ≥3 adverse events, and 40% of the patients experienced a serious adverse event. In Evens et al., 2018 [67] treatment-related toxicity caused two (4.2%) patients to come off study. In Friedberg et al. [69] 45% (n=10) of patients treated with brentuximab-vedotin in combination with dacarbazine and 76% (n=16) of patients treated with brentuximab vedotin in combination with nivolumab experienced grade ≥3 treatment-related adverse events.

In Cheson et al. [70], at 21.2-month follow-up, 2% of patients died of a treatment-related adverse effect and 48% suffered from peripheral neuropathy.

CHOP, PVAG, BEACOPP and Palliative care

Wrobel et al. [66] examined eight patients treated with either CHOP or PVAG, three patients treated with BEACOPP, and three patients treated with palliative care. However, the

results were not presented by treatment; therefore, no evidence is available for any the individual treatments.

Other treatments

Among other treatments we found two studies of VEPEMB [71,72], and one study of prednisone, vinblastine, doxorubicin, bendamustine (PVAB) [73]. Proctor et al. [71], as part of the SHIELD study, reported a phase 2 single-arm trial where 103 older patients, 82.2% with advanced stage disease, were treated with VEPEMB. Levis et al. 2004 [72] reported a phase 2 single-arm trial of patients treated with VEPEMB. Of the 105 patients included, 57 (54%) had advanced disease, and their results were reported separately. Ghesquieres et al. 2024 [73] reported data from a prospective phase 2 multicentre noncomparative trial (LYSA study) that included a cohort of 89 patients of median age 68 years. The patients were treated with first-line PVAB.

OS

Levis et al. [72] reported an actuarial OS rate of 32%.

Recurrence (PFS and other measures)

In multivariate analysis, Proctor et al. [71] reported that complete response, age, and comorbidity assessment were predictors of PFS. Levis et al. [72] reported an actuarial RFS rate of 66%, and a disease-specific survival (DSS) rate of 37%. In multivariate analysis stage, systemic symptoms, and comorbidity were prognostic factors of OS, DSS and FFS. Ghesquieres et al. 2024 [73] reported a PFS rate at 48-month follow-up of 50% (95% CI, 39 to 61).

Adverse events

Proctor et al. [71] reported a 7% treatment-related death, and dose reductions required in 67% of the whole group (patients with early-stage disease included). Levis et al. [72] reported grade 3-4 neutropenia in all patients, interruption or modification of treatment in 26% of patients, and hospitalization during treatment in 21% of patients. Ghesquieres et al. 2024 [73] reported that 31.5% of patients reported ≥1 severe adverse event.

Certainty of the evidence

We considered the certainty of the evidence provided by the single arm studies included for Question 2 very low.

Ongoing, Unpublished, or Incomplete Studies

A search for ongoing trials was conducted on August 24, 2023, on https://clinicaltrials.gov/. We used the key term “Advanced Hodgkin Lymphoma” and limited for: Phase II, III, and IV, interventional studies, and for studies that are not yet recruiting, recruiting, and active not recruiting. The search resulted in 74 hits, from which we selected four relevant randomized studies. Table 11 reports a summary of the relevant ongoing trials. In addition, the abstracts of interim analyses identified during the search for noncomparative trials for Question 2, are reports of ongoing trials, and they are summarized in Table 7-7A in Appendix 7.

Table 11. Ongoing trials (RCTs)*

#	ID number	Trial name /title	Comparison	Design / Sample	Recruitment status	Expected completion date
1	NCT02661503	**HD21 for Advanced Stages** http://agmt.at/wp-content/uploads/2019/03/HD_21_OeGHO_2017_web.pdf	6×eBEACOPP vs. 6×BrECADD	Non-inferiority / 1500 pts	Recruiting	December 2022 (Part 1: Tolerability) completed September 2025 (Part 2: Efficacy)
2	NCT03159897	**FIL - Rouge** trial. Fondazione Italiana Linfomi Study on ABVD DD-DI as Upfront Therapy in HL.	PET-2 adapted strategy: [2ABVD then: if PET2+ either escalated BEACOPP or HDT plus ASCR: if PET - 34×ABVD] vs. straight dose- and time-intensified schedule ABVD [three cycles of a dose-dense/dose-intense ABVD (ABVD DD-DI)]	Open-label, Multicenter, Phase III / 500 pts	Active not recruiting	January 2023
3	NCT03907488	**SWOG S1826** Immunotherapy (Nivolumab or Brentuximab Vedotin) Plus Combination Chemotherapy in Treating Patients With Newly Diagnosed Stage III-IV Classic HL.	Nivolumab (Opdivo) Plus AVD or Brentuximab Vedotin (Adcetris) Plus AVD	Phase III / 987 pts	Active not recruiting	March 1 2024
4	NCT05949931	**PENAHL Study** Pianzumab Combined ith AVD Regimen in the Treatment of Newly-diagnosed Advanced Classic HL.	Concurrent penpulimab and AVD vs. Sequential penpulimab and AVD	RCT, Open label, Phase II / 108 pts	Not yet recruiting	December 2026

*The serach for ongoing trials was executed on August 22, 2023 and updated on December 23, 2023.

ABVD=doxorubicin, bleomycin, vinblastine, and dacarbazine; ABVD DD-DI=dose-dense/dose-intense ABVD; ASCR=Autologous stem cell rescue; AVD=doxorubicin, vinblastine, and dacarbazine; BEACOPP=bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, prednisone; BrECADD=brentuximab vedotin, etoposide, cyclophosphamide, doxorubicin, dacarbazine, dexamethasone; HDT=High-dose therapy; HL=Hodgkin lymphoma; PET=Positron emission tomography; RCT=Randomized controlled trial

DISCUSSION AND CONCLUSIONS

Question 1: patients younger than 60 years of age

Comparison 1: ABVD vs. BEACOPP

For OS, most studies showed no statistically significant between-groups difference. An exception to this trend was the HD9 [54] in which the high-intensity BEACOPP regimen was better than the ABVD regimen. It is possible that the higher-intensity BEACOPP regimen was more effective; however, the certainty of the evidence for the HD9 trial [54] was very low.

Although chemotherapy doses as well as the recurrence outcome measures were slightly different from study to study, most measures of recurrence showed a benefit of BEACOPP combinations compared with ABVD-based regimens and all the estimates lay above the minimally clinically significant difference of 10% in PFS.

The certainty of the evidence was moderate-to-very low for OS. For recurrence, all the studies except the GITIL trial [57] reported on PFS, and the certainty of the evidence was moderate to low. The certainty of the evidence for the other measures of recurrence, including FFFP, EFS, FFTF, and FFS was low.

The primary studies included showed no between-group statistically significant difference in any of the critical or important measures for adverse events. The certainty of the evidence for adverse events was very low. The systematic review by Amin et al. [43] reported worse fertility outcomes for young male patients treated with BEACOPP compared with those treated with ABVD, particularly for patients who received a higher number of cycles of the therapy.

Comparison 2: Higher- vs. lower-intensity/dosage BEACOPP

For OS, most of the comparisons examined showed no statistical significant difference between higher and lower BEACOPP dosage/intensity, except for the comparison 6×eBEACOPP versus the highest intensity (i.e., 8×eBEACOPP) in the HD15 trial [58], which showed better OS in patients treated with the lowest dosage regimen.

For recurrence, two non-inferiority randomized trials the HD15 [58], the HD12 [59], and a superiority analysis of the HD9 [54] were included. The two noninferiority trials, the HD15 [58] and the HD12 [59], although the drug combinations were slightly different, had similar results, showing that reduced intensity BEACOPP regimens (i.e., 8×BEACOPP$_{14}$ and 6×eBEACOPP) were noninferior to 8×eBEACOPP for FFTF and PFS, whereas the HD9 trial [54] reported a better PFS with the higher intensity BEACOPP regimen.

All the measures of adverse effects showed no statistically significant difference between low- and high-dosage/intensity BEACOPP, except for the HD15 trial [58] when comparing 6×eBEACOPP with 8×eBEACOPP for SPM. This difference may be due to a greater impact of higher-intensity regimen on the occurrence of subsequent hematological cancers. This study noted also less treatment-related toxic effects with the lower compared to the higher intensity/dosage regimen among those outcomes that we did not consider critical or important, but that can still contribute to a better patient quality of life; for example, grade 3 or 4 pain rates were 9.2% for the low dosage versus 13.2% in the higher dosage, anemia rates were 53.2% versus 62.1%, infection rates were 22.3% versus 24.7%, respiratory tract adverse events rates were 3.7% versus 6.4%, and nervous system toxicity rates were 5% versus 7.6%.

The overall certainty of the evidence across all outcomes for the five different doses and schedules was moderate-to-low across efficacy outcomes and very low for adverse effects because of risk of bias, imprecision, and inconsistency in the direction of the results.

Comparison 3: PET-adapted strategies

Among studies of PET-adapted strategies, the AHL2011 [17,84] randomized all patients to either a PET-directed strategy or to a standard strategy without PET intervention. The other

two studies in this section (i.e., the RATHL trial [49,85] and the HD18 trial [16]) randomized the subgroup of patients who were PET-2 negative after initial treatment with two cycles of ABVD [49,85] and two cycles of eBEACOPP [16], to a less intensive treatment, versus a high-intensity standard.

The AHL2011 trial [17] reported no statistically significant between-group difference in OS and serious adverse events; noninferiority for PFS was proven for the PET-directed strategy compared with the standard. Infertility was less severe in patients, both men and women, treated in the PET-directed arm.

Likewise, in the RATHL trial, no statistically significant difference was shown at 36-month for OS or recurrence (PFS) between PET-2 negative patients allocated to four cycles of ABVD compared with those allocated to four cycles of AVD. Noninferiority for PFS was proven at 87.2 months of follow-up.

On the other hand, for PET-2 negative patients treated in the HD18 [16], OS was better for those in the lower-intensity than for those in the higher-intensity regimen, while the noninferiority condition was met for PFS.

Very low adverse events rates and no between-groups statistically significant difference in the included studies were reported, except for pulmonary toxicity in the RATHL trial [49] where the DL_{CO} was statistically significant in better for AVD compared with ABVD. The AHL2011 reported a better fertility profile for patients allocated to the PET-directed strategy compared to standard.

In the AHL2011 trial [17], and in the RATHL trial [49] the certainty of the evidence was moderate for OS; for recurrence, and for serious adverse events the certainty of the evidence was low to very low because of serious risk of bias, inconsistency, and imprecision.

Comparison 4: ABVD versus A-AVD

The ECHELON-1, an industry-funded study, was the trial included for this comparison. For OS the final analysis was not yet available at the time of our latest update. The 78-month estimate showed a statistically significant better PFS for A-AVD compared with ABVD. Adverse event rates were very low. The certainty of the evidence was moderate for OS, low for recurrence, and very low for adverse events because of very serious risk of bias, and imprecision.

Comparison 5: Modified ABVD versus another modified ABVD and Comparison 6: BEACOPP versus modified BEACOPP

No fully published, final, results for these comparisons are available. We are aware of two studies that are still ongoing at the present time, and that may be practice changing when completed and published. The first is the SWOG1826 trial [24,123] that compares nivolumab in combination with AVD with brentuximab vedotin in combination with AVD. The second trial is the HD21 that compares four to six cycles of standard eBEACOPP with a modified version containing brentuximab vedotin (BrECADD). Those new agents are promising because they have a very favourable toxicity profile. This is the reason why the Working Group decided not to draft recommendations for this document, and to conclude it just an evidence summary.

Question 2: patients 60 years of age or older

The body of evidence for this population is very limited as it consists of two ancillary studies of the Echelon-1 [9] and of the HD9 [54] included for Question 1 [61,63] and 10 phase 2 single-arm studies [65-74]. Additionally, the authors of the SWOG S1826 trial [23,123], one of the studies for which we are awaiting publication, are planning to publish a subgroup analysis for 98 older patients. Evidence was available for ABVD, brentuximab vedotin alone or in

combination, and VEPEMB, while no studies of CHOP, CHLVPP, or any other treatments met our inclusion criteria.

The subgroup analysis of the HD9 study [54,64,81] by Ballova et al. [63] reported no statistically significant difference in OS and FFTF between older patients treated with COPP/ABVD and those treated with bBEACOPP; mortality rates were above 50% in both groups, and treatment failed in 45% in the COPP/ABVD group and 24% in the bBEACOPP group. Similarly, Evens et al. [61], in a non-preplanned analysis of the ECHELON-1 [9], detected no statistically significant difference in PFS between older patients treated with A-AVD and those treated with ABVD. Progression rates were above 30% in both groups, and deaths related to treatment were around 5% in both groups. Any grade pulmonary toxicity rates were higher in the ABVD compared with the A-AVD group. Rates for other outcomes, such as peripheral neuropathy, and febrile neutropenia, which we did not include as critical or important, were also in favour of the A-AVD regimen in this subgroup of older patients.

The certainty of the evidence was low to very low for both subgroup analyses, but, overall, the data confirmed the all-too-well-known poor outcomes for older patients across different treatments. The single-arm trials reported on ABVD versus ABVD-like treatments, brentuximab vedotin alone or in combination, and other treatments such as VEPEMB or PVAB. These studies were noncomparative; therefore, the certainty of the evidence in their results was very low.

The treatments that we were expecting to find for these patients were ABVD or modifications of ABVD, brentuximab vedotin alone or in combination, and other treatments, as well as CHOP, and CHLVPP. None of the studies that met our inclusion criteria treated older patients with advanced-stage HL with CHOP or CHLVPP. This can be interpreted as a further testimony of the heterogeneity of treatment patterns around the world for this population of older people.

Question 3: Radiotherapy consolidation

We included studies that compared radiotherapy consolidation either with observation, or with chemotherapy.

In the comparison of radiotherapy with observation, the four included studies were consistent in that they did not detect any between-group statistically significant difference for either OS or recurrence. Second primary malignancies were reported by three of the four studies (i.e., the HD0801 FIL [75], the EORTC 20884 [76], and the HD12 [59]) in different patients' subpopulations and the results were in favour of observation for studies of larger sample size where more events could be detected, while the smaller study reported no statistically significant difference.

The two studies that compared radiotherapy to chemotherapy had contrasting results: one of them [78] reported better OS at 20 years for patients allocated to chemotherapy than for those allocated to radiotherapy; the other [79], which had a much shorter, five-year follow-up, reported better OS and FFS for patients allocated to combination radiotherapy and chemotherapy than for those who received only chemotherapy.

We considered the certainty of the evidence for OS and recurrence moderate to low because of risk of bias, inconsistency, and imprecision and low to very low for SPM because of risk of bias, inconsistency and imprecision.

It is possible that with the accumulating of more evidence, when the clinical practice guideline or this comparison will be proposed, the certainty of the evidence for this question will be higher, and, once evidence matures, we will be in a better position to make a recommendation.

Study strengths and limitations

This review was very comprehensive and thorough. We used broad categories for comparisons, and we did not focus on a specific treatment, because many options are available to choose from for patients and clinicians depending on patients' circumstances, values and preferences. This approach has some limitations though as, for example, we may have lost granularity in detecting which type of secondary neoplasia can be caused by specific high-intensity treatments.

We chose a few outcomes that both the clinicians in the Working Group and the patients representatives considered critical or important, and we focussed our analyses on those outcomes. The included studies also reported on other outcomes (e.g., infections, neutropenia, peripheral neuropathy) that can contribute to some patients' quality of life, and we summarized them too.

We did not have well defined thresholds of intensity for the various regimens, and it was not easy to separate out $BEACOPP_{14}$ from 4× or 6×eBEACOPP in terms of beneficial and harmful effects on the patients, because they all were considered "lower-intensity" regimens compared to the 8×eBEACOPP.

We performed the risk of bias assessment for included studies by comparison, and outcome by outcome according to the Cochrane ROB2 tool [36] (see Table 3-5A-1). Most often this resulted in ratings that revealed a serious to very serious risk of bias. Most of the studies were open label trials and, if seen in a continuum, they can be considered more akin to pragmatic than explanatory trials. This explains the lack of blinding of clinicians, patients and outcome assessors that often reduced the certainty of the evidence for outcomes other than OS.

The treatments used for HL, even in its advanced stage attain a cure for many patients, and the major focus for patients and clinicians alike is reducing the number of adverse events and improve quality of life for those who survive, particularly for older patients. Therefore, serious treatment-related adverse events, were some of our critical or important outcomes. We did not include observational trials for adverse events. In the included studies the rate of those events was consistently, across all randomizations, very low. Infertility, an important outcome for younger patients, was not assessed in the included RCTs, but in ancillary, observational substudies of the original RCTs, and in the systematic review by Amin et al. [43].

This made the evidence for adverse events and infertility outcomes very imprecise and the certainty of this evidence very low thus impairing our ability to properly balance the beneficial effects with the harms of treatments and making the choice of a best treatment more difficult for patients and clinicians.

This difficulty is even increased for older patients: when we examined this population in our Question 2, we found consistent lack of data across all interventions and comparisons.

We also conducted a thorough search for ongoing and incomplete trials. This allowed us to identify at least two pivotal trials, the SWOG S1826 [23,123] and the HD21 [22], which may bring some certainty about adverse events and about older patients, at least from some of the new treatments. Of note, HD21 was recently published after our data cut-off date and will be included in the updated recommendations.

In the future, probably studies with larger sample sizes, observational studies or phase IV studies should be sought for a similar review. Additionally, since most treatment approaches are effective, although with different toxicity profiles, in the future it will be of interest to focus on clinical prediction models able to tell us which treatment is best suited for which specific patient, thus creating an approach akin to personalized medicine.

We hope that this piece of work will raise awareness among researchers in this field of the need for more studies focusing on underrepresented populations, particularly patients aged 60 and above.

INTERNAL REVIEW

The evidence summary was reviewed by Xiaomei Yao, and Jonathan Sussman. The Working Group was responsible for ensuring any necessary changes were made.

Acceptance by the Hematology Disease Site Group and to the PET Steering Committee

After internal review, the report was presented to the Hematology Disease Site Group and to the PET Steering Committee. The Hematology Disease Site Group and to the PET Steering Committee reviewed the document that was sent to them as an electronic file by email on October 16, 2024, and formally accepted the document.

ACKNOWLEDGEMENTS

The Hematology Disease Site Group and to the PET Steering Committee and the Working Group would like to thank the following individuals for their assistance in developing this report:

- Sheila McNair, Emily Vella, Xiaomei Yao, and Jonathan Sussman for providing feedback on draft versions.
- Wenjun Jiang for conducting a data audit.
- Sara Miller for copy editing.

References

1. Cancer Research UK. Hodgkin lymphoma incidence statistics 2015 [cited May 3, 2023]. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/hodgkin-lymphoma/incidence.
2. Brenner DR, Poirier A, Woods RR, Ellison LF, Billette JM, Demers AA, et al. Projected estimates of cancer in Canada in 2022. CMAJ. 2022;194(17):E601-E7.
3. Xie L, Semenciw R, Mery L. Cancer incidence in Canada: trends and projections (1983-2032). Health Promot Chronic Dis Prev Can [Internet]. 2015 [cited May 3, 2023]; Volume 35, (Suppl 1, Spring).
4. Nguyen J, Wellard C, Chung E, Cheah CY, Dickinson M, Doo NW, et al. Clinical characteristics of Australian treatment-naive patients with classical Hodgkin lymphoma from the lymphoma and related diseases registry. Eur J Haematol. 2023;110(4):386-95.
5. Bhurani D, Nair R, Rajappa S, Rao SA, Sridharan N, Boya RR, et al. Real-World Outcomes of Hodgkin Lymphoma: A Multi-Centric Registry From India. Front Oncol. 2021;11:799948.
6. Skoetz N, Will A, Monsef I, Brillant C, Engert A, von Tresckow B. Comparison of first-line chemotherapy including escalated BEACOPP versus chemotherapy including ABVD for people with early unfavourable or advanced stage Hodgkin lymphoma. Cochrane Database Syst Rev. 2017;5:CD007941.
7. Longley J, Johnson PWM. Current treatment paradigms for advanced stage Hodgkin lymphoma. Br J Haematol. 2019;184(1):60-71.
8. Canellos GP, Anderson JR, Propert KJ, Nissen N, Cooper MR, Henderson ES, et al. Chemotherapy of advanced Hodgkin's disease with MOPP, ABVD, or MOPP alternating with ABVD. N Engl J Med. 1992;327(21):1478-84.
9. Connors JM, Jurczak W, Straus DJ, Ansell SM, Kim WS, Gallamini A, et al. Brentuximab Vedotin with Chemotherapy for Stage III or IV Hodgkin's Lymphoma. N Engl J Med. 2018;378(4):331-44.
10. Ansell SM. PD-1 Blockade in Classic Hodgkin Lymphoma. JCO Oncol Pract. 2021;17(2):72-3.
11. Andre MPE, Carde P, Viviani S, Bellei M, Fortpied C, Hutchings M, et al. Long-term overall survival and toxicities of ABVD vs BEACOPP in advanced Hodgkin lymphoma: A pooled analysis of four randomized trials. Cancer Med. 2020;9(18):6565-75.
12. Schaapveld M, Aleman BM, van Eggermond AM, Janus CP, Krol AD, van der Maazen RW, et al. Second Cancer Risk Up to 40 Years after Treatment for Hodgkin's Lymphoma. N Engl J Med. 2015;373(26):2499-511.
13. Swerdlow AJ, Higgins CD, Smith P, Cunningham D, Hancock BW, Horwich A, et al. Second cancer risk after chemotherapy for Hodgkin's lymphoma: a collaborative British cohort study. J Clin Oncol. 2011;29(31):4096-104.
14. Borchmann S, Engert A, Boll B. Hodgkin lymphoma in elderly patients. Curr Opin Oncol. 2018;30(5):308-16.
15. Evens AM, Carter J, Loh KP, David KA. Management of older Hodgkin lymphoma patients. Hematol Am Soc Hematol Educ Program. 2019;2019(1):233-42.
16. Borchmann P, Goergen H, Kobe C, Lohri A, Greil R, Eichenauer DA, 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. Lancet. 2018;390(10114):2790-802.
17. Casasnovas RO, Bouabdallah R, Brice P, Lazarovici J, Ghesquieres H, Stamatoullas A, et al. PET-adapted treatment for newly diagnosed advanced Hodgkin lymphoma (AHL2011):

a randomised, multicentre, non-inferiority, phase 3 study. Lancet Oncol. 2019;20(2):202-15.
18. Lang N, Crump M. PET-adapted approaches to primary therapy for advanced Hodgkin lymphoma. Ther Adv Hematol. 2020;11:2040620720914490.
19. Trotman J, Barrington SF. The role of PET in first-line treatment of Hodgkin lymphoma. Lancet Haematol. 2021;8(1):e67-e79.
20. Johnson PW, Sydes MR, Hancock BW, Cullen M, Radford JA, Stenning SP. Consolidation radiotherapy in patients with advanced Hodgkin's lymphoma: survival data from the UKLG LY09 randomized controlled trial (ISRCTN97144519). J Clin Oncol. 2010;28(20):3352-9.
21. Trotman J, Fossa A, Federico M, Stevens L, Kirkwood A, Clifton-Hadley L, et al. Response-adjusted therapy for advanced Hodgkin lymphoma (RATHL) trial: Longer follow up confirms efficacy of de-escalation after a negative interim PET scan (CRUK/07/033). Hematol Oncol. 2017;35(Suppl 2):65-7.
22. Borchmann P, Moccia A, Greil R, Hertzberg M, Schaub V, Hüttmann A, et al. Treatment Related Morbidity in Patients with Classical Hodgkin Lymphoma: Results of the Ongoing, Randomized Phase III HD21 Trial By the German Hodgkin Study Group. Blood. 2022;140(Suppl 1):771-3.
23. Herrera AF, LeBlanc M, Castellino SM, Li H, Rutherford SC, Evens AM, et al. Nivolumab(N)-AVD Improves progression-free survival compared to brentuximab vedotin (BV)-AVD in advanced stage (as) classic Hodgkin lymphoma (HL): Results of SWOG S1826. Hematol Oncol. 2023;41(Suppl 2):33-5.
24. Borchmann P, Moccia AA, Greil R, Schneider G, Hertzberg M, Schaub V, et al. Brecadd Is Non-Inferior to Ebeacopp in Patients with Advanced Stage Classical Hodgkin Lymphoma: Efficacy Results of the Ghsg Phase Iii Hd21 Trial. Hematol Oncol. 2023;41(Suppl 2):881-2.
25. National Institute for Health and Care Excellence. Brentuximab vedotin for treating CD30-positive Hodgkin lymphoma. 2018.
26. McKay P, Fielding P, Gallop-Evans E, Hall GW, Lambert J, Leach M, et al. Guidelines for the investigation and management of nodular lymphocyte predominant Hodgkin lymphoma. Br J Haematol. 2016;172(1):32-43.
27. Bröckelmann PJ, Eichenauer DA, Jakob T, Follmann M, Engert A, Skoetz N. Hodgkin Lymphoma in Adults. Dtsch Arztebl Int. 2018;115(31-32):535-40.
28. National Institute for Health and Care Excellence. Brentuximab vedotin for untreated advanced Hodgkin lymphoma (terminated appraisal). Internet 2019.
29. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017;358:j4008.
30. Mendenhall W, Beaver RJ, Beaver BM. Introduction to Probability and Statistics 11th ed. Pacific Grove, CA: Thomson-Brooks/Cole; 2003.
31. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. J Clin Epidemiol. 2011;64(4):395-400.
32. Khan N, Feliciano J, Muller K, He M, Tao R, Korol E, et al. Patient preferences for first-line treatment of classical Hodgkin lymphoma: a US survey and discrete choice experiment. Leuk Lymphoma. 2020;61(11):2630-7.
33. Bröckelmann PJ, McMullen S, Wilson JB, Mueller K, Goring S, Stamatoullas A, et al. Patient and physician preferences for first-line treatment of classical Hodgkin lymphoma in Germany, France and the United Kingdom. Br J Haematol. 2019;184(2):202-14.

34. Schünemann HJ, Higgins JP, Vist G, Glasziou P, Akl EA, Skoetz N, et al. Chapter 14: Completing 'Summary of findings' tables and grading the certainty of the evidence. In: Higgins JTP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. Cochrane Handbook for Systematic Reviews of Interventions version 63 (updated February 2022). Internet: Cochrane; 2022.
35. Tierney JF, Stewart LA, Vale CL, Riley RD. A tool for the critical apprisal of IPD meta-analysis projects (CheckMAP). In: Riley RD, Tierney JF, Stewart LA, editors. Individual patient data meta-analysis: A handbook for healthcare research. Wiley Series in Statistics in Practice. Oxford, UK: John Wiley & Sons; 2021. p. 271-83.
36. Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:l4898.
37. Sterne JAC, Higgins JPT, Reeves BC, On behalf of the development group for ROBINS-I:. A tool for assessing Risk of Bias in Non-randomized Studies of Interventions 2016 [cited March 3, 2023]. Version 7 March 2016]. Available from: http://riskofbias.info.
38. Sun X, Ioannidis JP, Agoritsas T, Alba AC, Guyatt G. How to use a subgroup analysis: users' guide to the medical literature. JAMA. 2014;311(4):405-11.
39. Schünemann H, Brozek J, Guyatt G, Oxman, AD (editors). GRADE handbook for grading the quality of evidence and the strength of recommendations. Updated October 2013. The GRADE Working Group. 2013. Available from: http://gradepro.org
40. Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. Stat Med. 1998;17(24):2815-34.
41. Aldin A, Umlauff L, Estcourt LJ, Collins G, Moons KG, Engert A, et al. Interim PET-results for prognosis in adults with Hodgkin lymphoma: a systematic review and meta-analysis of prognostic factor studies. Cochrane Database Syst Rev. 2020;1:CD012643.
42. Amitai I, Gurion R, Vidal L, Dann EJ, Raanani P, Gafter-Gvili A. PET-adapted therapy for advanced Hodgkin lymphoma - systematic review. Acta Oncologica (Stockholm, Sweden). 2018;57(6):1-8.
43. Amin MSA, Brunckhorst O, Scott C, Wrench D, Gleeson M, Kazmi M, et al. ABVD and BEACOPP regimens' effects on fertility in young males with Hodgkin lymphoma. Clin Transl Oncol. 2020;23(6):1067-77.
44. Dalal M, Gupta J, Price K, Zomas A, Miao H, Ashaye A. Efficacy and safety of front-line treatments for advanced Hodgkin lymphoma: a systematic literature review. Expert Rev Hematol. 2020;13(8):907-22.
45. Zhang T, Yao Y, Feng F, Zhao W, Tian J, Zhou C, et al. Comparative effectiveness of different chemotherapy regimens of advanced-stage Hodgkin lymphoma in adults: a network meta-analysis. Cancer Manag Res. 2018;10:6017-28.
46. Franklin J, Eichenauer DA, Becker I, Monsef I, Engert A. Optimisation of chemotherapy and radiotherapy for untreated Hodgkin lymphoma patients with respect to second malignant neoplasms, overall and progression-free survival: individual participant data analysis. Cochrane Database Syst Rev. 2017;9(9):Cd008814.
47. Allen PB, Ayers A, Behera M, Evens AM, Flowers C. A systematic review of therapeutic regimens for older patients with newly diagnosed Hodgkin lymphoma. Leuk Lymphoma. 2020;61(7):1555-64.
48. Gallamini A, Tarella C, Viviani S, Rossi A, Patti C, Mulé A, 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(5):454-62.

49. Johnson P, Federico M, Kirkwood A, Fossa A, Berkahn L, Carella A, et al. Adapted treatment guided by interim PET-CT scan in advanced hodgkin's lymphoma. N Engl J Med. 2016;374(25):2419-29.
50. Borchmann P, Haverkamp H, Lohri A, Mey U, Kreissl S, Greil R, 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. 2017;18(4):454-63.
51. Borchmann P, Plütschow A, Kobe C, Greil R, Meissner J, Topp MS, et al. PET-guided omission of radiotherapy in early-stage unfavourable Hodgkin lymphoma (GHSG HD17): a multicentre, open-label, randomised, phase 3 trial. Lancet Oncol. 2021;22(2):223-34.
52. Carde P, Karrasch M, Fortpied C, Brice P, Khaled H, Casasnovas O, et al. Eight Cycles of ABVD Versus Four Cycles of BEACOPPescalated Plus Four Cycles of BEACOPPbaseline in Stage III to IV, International Prognostic Score ≥ 3, High-Risk Hodgkin Lymphoma: First Results of the Phase III EORTC 20012 Intergroup Trial. J Clin Oncol. 2016;34(17):2028-36.
53. Mounier N, Brice P, Bologna S, Briere J, Gaillard I, Heczko M, et al. ABVD (8 cycles) versus BEACOPP (4 escalated cycles >= 4 baseline): final results in stage III-IV low-risk Hodgkin lymphoma (IPS 0-2) of the LYSA H34 randomized trial. Ann Oncol. 2014;25(8):1622-8.
54. Diehl V, Franklin J, Pfreundschuh M, Lathan B, Paulus U, Hasenclever D, et al. Standard and increased-dose BEACOPP chemotherapy compared with COPP-ABVD for advanced Hodgkin's disease. N Engl J Med. 2003;348(24):2386-95.
55. Federico M, Luminari S, Iannitto E, Polimeno G, Marcheselli L, Montanini A, et al. ABVD compared with BEACOPP compared with CEC for the initial treatment of patients with advanced Hodgkin's lymphoma: results from the HD2000 Gruppo Italiano per lo Studio dei Linfomi Trial. J Clin Oncol. 2009;27(5):805-11.
56. Merli F, Luminari S, Gobbi PG, Cascavilla N, Mammi C, Ilariucci F, et al. Long-Term Results of the HD2000 Trial Comparing ABVD Versus BEACOPP Versus COPP-EBV-CAD in Untreated Patients With Advanced Hodgkin Lymphoma: A Study by Fondazione Italiana Linfomi. J Clin Oncol. 2016;34(11):1175-81.
57. Viviani S, Zinzani PL, Rambaldi A, Brusamolino E, Levis A, Bonfante V, et al. ABVD versus BEACOPP for Hodgkin's lymphoma when high-dose salvage is planned. N Engl J Med. 2011;365(3):203-12.
58. Engert A, Haverkamp H, Kobe C, Markova J, Renner C, Ho A, 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. 2012;379(9828):1791-9.
59. Borchmann P, Haverkamp H, Diehl V, Cerny T, Markova J, Ho AD, et al. Eight cycles of escalated-dose BEACOPP compared with four cycles of escalated-dose BEACOPP followed by four cycles of baseline-dose BEACOPP with or without radiotherapy in patients with advanced-stage hodgkin's lymphoma: final analysis of the HD12 trial of the German Hodgkin Study Group. J Clin Oncol. 2011;29(32):4234-42.
60. Gallamini A, Rossi A, Patti C, Picardi M, Romano A, Cantonetti M, et al. Consolidation Radiotherapy Could Be Safely Omitted in Advanced Hodgkin Lymphoma With Large Nodal Mass in Complete Metabolic Response After ABVD: Final Analysis of the Randomized GITIL/FIL HD0607 Trial. J Clin Oncol. 2020;38(33):3905-13.
61. Evens AM, Connors JM, Younes A, Ansell SM, Kim WS, Radford J, et al. Older patients (aged ≥60 years) with previously untreated advanced-stage classical Hodgkin lymphoma:

a detailed analysis from the phase III ECHELON-1 study. Haematologica. 2022;107(5):1086-94.
62. Hutchings M, Ansell SM, Straus DJ, Connors JM, Kim WS, Gallamini A, et al. Improved overall survival with first-line brentuximab vedotin plus chemotherapy in patients with advanced stage III/IV classical Hodgkin lymphoma: 6-year analysis of ECHELON-1. HemaSphere. 2022;6(Suppl 3):207-8.
63. Ballova V, Ruffer JU, Haverkamp H, Pfistner B, Muller-Hermelink HK, Duhmke E, et al. A prospectively randomized trial carried out by the German Hodgkin Study Group (GHSG) for elderly patients with advanced Hodgkin's disease comparing BEACOPP baseline and COPP-ABVD (study HD9elderly). Ann Oncol. 2005;16(1):124-31.
64. Diehl V, Franklin J, Hasenclever D, Tesch H, Pfreundschuh M, Lathan B, et al. BEACOPP, a new dose-escalated and accelerated regimen, is at least as effective as COPP/ABVD in patients with advanced-stage Hodgkin's lymphoma: interim report from a trial of the German Hodgkin's Lymphoma Study Group. J Clin Oncol. 1998;16(12):3810-21.
65. Yildiz J, Ulu BU, Yigenoglu TN, Basci S, Darcin T, Bakirtas M, et al. Is classical Hodgkin lymphoma a different disease in the elderly? A single-center retrospective cohort study. Eur Rev Med Pharmacol Sci. 2021;25(10):3655-60.
66. Wrobel T, Biecek P, Rybka J, Szulgo A, Sorbotten N, Giza A, et al. Hodgkin lymphoma of the elderly patients: a retrospective multicenter analysis from the Polish Lymphoma Research Group. Leuk Lymphoma. 2019;60(2):341-8.
67. Evens AM, Advani RH, Helenowski IB, Fanale M, Smith SM, Jovanovic BD, et al. Multicenter Phase II Study of Sequential Brentuximab Vedotin and Doxorubicin, Vinblastine, and Dacarbazine Chemotherapy for Older Patients With Untreated Classical Hodgkin Lymphoma. J Clin Oncol. 2018;36(30):3015-22.
68. Gibb A, Pirrie SJ, Linton K, Warbey V, Paterson K, Davies AJ, et al. Results of a UK National Cancer Research Institute Phase II study of brentuximab vedotin using a response-adapted design in the first-line treatment of patients with classical Hodgkin lymphoma unsuitable for chemotherapy due to age, frailty or comorbidity (BREVITY). Br J Haematol. 2021;193(1):63-71.
69. Friedberg JW, Bordoni R, Patel-Donnelly D, Larson T, Goldschmidt J, Boccia R, et al. Brentuximab vedotin with dacarbazine or nivolumab as frontline cHL therapy for older patients ineligible for chemotherapy. Blood. 2024;143(9):786-95.
70. Cheson BD, Bartlett NL, LaPlant B, Lee HJ, Advani RH, Christian B, et al. Phase II, multicenter trial of nivolumab (Nivo) and brentuximab vedotin (BV) in patients (Pts) with untreated Hodgkin lymphoma (HL) over the age of 60 years or unable to receive standard ABVD chemotherapy: Results of a study of Academic and Community Cancer Research United (ACCRU) RU051505I. J Clin Oncol. 2020;38(15 suppl):8014-.
71. Proctor SJ, Wilkinson J, Jones G, Watson GC, Lucraft HH, Mainou-Fowler T, et al. Evaluation of treatment outcome in 175 patients with Hodgkin lymphoma aged 60 years or over: the SHIELD study. Blood. 2012;119(25):6005-15.
72. Levis A, Anselmo AP, Ambrosetti A, Adamo F, Bertini M, Cavalieri E, et al. VEPEMB in elderly Hodgkin's lymphoma patients. Results from an Intergruppo Italiano Linfomi (IIL) study. Ann Oncol. 2004;15(1):123-8.
73. Ghesquieres H, Krzisch D, Nicolas-Virelizier E, Kanoun S, Gac AC, Guidez S, et al. The phase 2 LYSA study of prednisone, vinblastine, doxorubicin, and bendamustine for untreated Hodgkin lymphoma in older patients. Blood. 2024;143(11):983-95.
74. Cokgezer S, Elverdi T, Salihoglu A, Ar MC, Ongoren S, Baslar Z, et al. Treatment Responses, Toxicity, and Survival in Patients with Classical Hodgkin Lymphoma Aged >=50 Years: A Single-Center Experience Over Two Decades. Cancer Manag Res. 2022;14:1911-21.

75. Ricardi U, Levis M, Evangelista A, Gioia DM, Sacchetti GM, Gotti M, et al. Role of radiotherapy to bulky sites of advanced Hodgkin lymphoma treated with ABVD: final results of FIL HD0801 trial. Blood Advances. 2021;01:4504-14.
76. Aleman BM, Raemaekers JM, Tomisic R, Baaijens MH, Bortolus R, Lybeert ML, et al. Involved-field radiotherapy for patients in partial remission after chemotherapy for advanced Hodgkin's lymphoma. Int J Radiat Oncol Biol Phys. 2007;67(1):19-30.
77. von Tresckow B, Kreissl S, Goergen H, Bröckelmann PJ, Pabst T, Fridrik M, et al. Intensive treatment strategies in advanced-stage Hodgkin's lymphoma (HD9 and HD12): analysis of long-term survival in two randomised trials. Lancet Haematol. 2018;5(10):e462-e73.
78. Wiernik PH, Hong F, Glick JH, Bennett JM. Radiation therapy compared with chemotherapy for consolidation of chemotherapy-induced remission of advanced Hodgkin lymphoma: a study by the Eastern Co-operative Oncology Group (E1476) with >20 years follow-up. Leuk Lymphoma. 2009;50(10):1632-41.
79. Aviles A, Delgado S, Talavera A, Gonzalez JL, Diaz Maqueo JC. Adjuvant radiotherapy to initial bulky disease in patients with advanced stage Hodgkin's disease. Hematology. 2000;4(6):479-85.
80. Herrera AF, Li H, Castellino SM, Rutherford SC, Davison K, Evans AG, et al. SWOG S1826: A phase III, randomized study of nivolumab plus AVD or brentuximab vedotin plus AVD in patients with newly diagnosed advanced stage classical hodgkin lymphoma. Blood. 2020;136(Suppl 1):23-4.
81. Engert A, Diehl V, Franklin J, Lohri A, Dorken B, Ludwig WD, et al. Escalated-dose BEACOPP in the treatment of patients with advanced-stage Hodgkin's lymphoma: 10 years of follow-up of the GHSG HD9 study. J Clin Oncol. 2009;27(27):4548-54.
82. Engert A, Goergen H, Markova J, Pabst T, Meissner J, Zijlstra JM, 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(1):e5.
83. Kreissl S, Goergen H, Buehnen I, Kobe C, Moccia A, Greil R, et al. PET-guided eBEACOPP treatment of advanced-stage Hodgkin lymphoma (HD18): follow-up analysis of an international, open-label, randomised, phase 3 trial. Lancet Haematol. 2021;8(6):e398-e409.
84. Casasnovas RO, Bouabdallah R, Brice P, Lazarovici J, Ghesquieres H, Stamatoullas A, et al. Positron Emission Tomography-Driven Strategy in Advanced Hodgkin Lymphoma: Prolonged Follow-Up of the AHL2011 Phase III Lymphoma Study Association Study. J Clin Oncol. 2022;40(10):1091-101.
85. Luminari S, Fossa A, Trotman J, Molin D, d'Amore F, Enblad G, et al. Long-Term Follow-Up of the Response-Adjusted Therapy for Advanced Hodgkin Lymphoma Trial. J Clin Oncol. 2024;42(1):13-8.
86. Straus DJ, Dlugosz-Danecka M, Alekseev S, Illes A, Picardi M, Lech-Maranda E, et al. Brentuximab vedotin with chemotherapy for stage III/IV classical Hodgkin lymphoma: 3-year update of the ECHELON-1 study. Blood. 2020;135(10):735-42.
87. Radford J, Bartlett N, Straus D, Dlugosz-Danecka M, Alekseev S, Illes A, et al. Brentuximab vedotin plus chemotherapy for stage III/IV classical Hodgkin lymphoma: 4-year update of the ECHELON-1 study. Br J Haematol. 2020;189(Suppl 1):101-2.
88. Straus DJ, Dlugosz-Danecka M, Connors JM, Alekseev S, Illes A, Picardi M, et al. Brentuximab vedotin with chemotherapy for stage III or IV classical Hodgkin lymphoma (ECHELON-1): 5-year update of an international, open-label, randomised, phase 3 trial. Lancet Haematol. 2021;8(6):e410-e21.

89. Ansell SM, Radford J, Connors JM, Dlugosz-Danecka M, Kim W-S, Gallamini A, et al. Overall Survival with Brentuximab Vedotin in Stage III or IV Hodgkin's Lymphoma. N Engl J Med. 2022;387(4):310-20.
90. Demeestere I, Racape J, Dechene J, Dupuis J, Morschhauser F, De Wilde V, et al. Gonadal Function Recovery in Patients With Advanced Hodgkin Lymphoma Treated With a PET-Adapted Regimen: Prospective Analysis of a Randomized Phase III Trial (AHL2011). J Clin Oncol. 2021;39(29):3251-60.
91. Kobe C, Goergen H, Baues C, Kuhnert G, Voltin CA, Zijlstra J, et al. Outcome-based interpretation of early interim PET in advanced-stage Hodgkin lymphoma. Blood. 2018;132(21):2273-9.
92. Ferdinandus J, Muller H, Damaschin C, Jacob AS, Meissner J, Krasniqi F, et al. Impact of individualized treatment on recovery from fatigue and return to work in survivors of advanced stage Hodgkin Lymphoma: Results from the randomized international GHSG HD18 trial. Oncol Res Treat. 2023;46(Suppl 5):184-5.
93. Anderson RA, Remedios R, Kirkwood AA, Patrick P, Stevens L, Clifton-Hadley L, et al. Determinants of ovarian function after response-adapted therapy in patients with advanced Hodgkin's lymphoma (RATHL): a secondary analysis of a randomised phase 3 trial. Lancet Oncol. 2018;19(10):1328-37.
94. Ramchandren R, Advani RH, Ansell SM, Bartlett NL, Chen R, Connors JM, et al. Brentuximab Vedotin plus Chemotherapy in North American Subjects with Newly Diagnosed Stage III or IV Hodgkin Lymphoma. Clin Cancer Res. 2019;25(6):1718-26.
95. Grigg AP, Lacasce AS, Bartlett NL, Straus DJ, Savage KJ, Zinzani PL, et al. Brentuximab vedotin with chemotherapy in adolescents and young adults (ayas) with stage iii or iv hodgkin lymphoma: A subgroup analysis from the phase 3 echelon-1 study. HemaSphere. 2021;5(Suppl 2):351-2.
96. Behringer K, Mueller H, Goergen H, Thielen I, Eibl AD, Stumpf V, et al. Gonadal function and fertility in survivors after Hodgkin lymphoma treatment within the German Hodgkin Study Group HD13 to HD15 trials. J Clin Oncol. 2013;31(2):231-9.
97. Kreissl S, Müller H, Goergen H, Meissner J, Topp M, Sökler M, et al. Health-Related Quality of Life in Patients With Hodgkin Lymphoma: A Longitudinal Analysis of the German Hodgkin Study Group. J Clin Oncol. 2020;38(25):2839-48.
98. Eichenauer DA, Plütschow A, Kreissl S, Sökler M, Hellmuth JC, Meissner J, et al. Incorporation of brentuximab vedotin into first-line treatment of advanced classical Hodgkin's lymphoma: final analysis of a phase 2 randomised trial by the German Hodgkin Study Group. Lancet Oncol. 2017;18(12):1680-7.
99. Damaschin C, Goergen H, Kreissl S, Plutschow A, Breywisch F, Mathas S, et al. Brentuximab vedotin-containing escalated BEACOPP variants for newly diagnosed advanced-stage classical Hodgkin lymphoma: follow-up analysis of a randomized phase II study from the German Hodgkin Study Group. Leukemia. 2021;18:18.
100. Damaschin C, Goergen H, Kreissl S, Plutschow A, Breywisch F, Mathas S, et al. Brentuximab vedotin-containing escalated BEACOPP variants for newly diagnosed advanced-stage classical Hodgkin lymphoma: follow-up analysis of a randomized phase II study from the German Hodgkin Study Group. Leukemia. 2022;36(2):580-2.
101. Rigacci L, Puccini B, Broccoli A, Dona M, Gotti M, Evangelista A, et al. Clinical characteristics of interim-PET negative patients with a positive end PET from the prospective HD08-01 FIL study. Ann Hematol. 2020;99(2):283-91.
102. Zinzani PL, Broccoli A, Gioia DM, Castagnoli A, Ciccone G, Evangelista A, 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(12):1376-85.

103. Engel C, Loeffler M, Schmitz S, Tesch H, Diehl V. Acute hematologic toxicity and practicability of dose-intensified BEACOPP chemotherapy for advanced stage Hodgkin's disease. German Hodgkin's Lymphoma Study Group (GHSG). Ann Oncol. 2000;11(9):1105-14.
104. Borchmann S, Muller H, Haverkamp H, Baues C, Markova J, Huttmann A, et al. Symptomatic osteonecrosis as a treatment complication in Hodgkin lymphoma: an analysis of the German Hodgkin Study Group (GHSG). Leukemia. 2019;33(2):439-46.
105. Bröckelmann PJ, Goergen H, Kohnhorst C, Tresckow Bv, Moccia A, Markova J, et al. Late Relapse of Classical Hodgkin Lymphoma: An Analysis of the German Hodgkin Study Group HD7 to HD12 Trials. J Clin Oncol. 2017;35(13):1444-50.
106. Haverkamp H, Boll B, Eichenauer DA, Sasse S, Fuchs M, Borchmann P, et al. Impact of Bleomycin and Vincristine Dose Reductions in Patients With Advanced Hodgkin Lymphoma Treated With BEACOPP: An Analysis of the German Hodgkin Study Group HD12 and HD15 Trials. J Clin Oncol. 2015;33(22):2430-6.
107. Eichenauer DA, Thielen I, Haverkamp H, Franklin J, Behringer K, Halbsguth T, et al. Therapy-related acute myeloid leukemia and myelodysplastic syndromes in patients with Hodgkin lymphoma: a report from the German Hodgkin Study Group. Blood. 2014;123(11):1658-64.
108. Wongso D, Fuchs M, Plutschow A, Klimm B, Sasse S, Hertenstein B, et al. Treatment-related mortality in patients with advanced-stage hodgkin lymphoma: an analysis of the german hodgkin study group. J Clin Oncol. 2013;31(22):2819-24.
109. Gallamini A, Rambaldi A, Patti C, Romano A, Viviani S, Bolis S, et al. Baseline metabolic tumor volume and IPS predict ABVD failure in advanced-stage hodgkin lymphoma with a negative interim pet scan after 2 chemotherapy cycles. a retrospective analysis from the gitil/FIL HD0607 trial. Hematol Oncol. 2021;39(Suppl 2):47-8.
110. Straus D, Collins G, Walewski J, Zinzani PL, Grigg A, Sureda A, et al. Primary prophylaxis with G-CSF may improve outcomes in patients with newly diagnosed stage III/IV Hodgkin lymphoma treated with brentuximab vedotin plus chemotherapy. Leuk Lymphoma. 2020;61(12):2931-8.
111. Radford J, Connors JM, Younes A, Gallamini A, Ansell SM, Kim WS, et al. Exploratory biomarker analysis in the ph 3 echelon-1 study: Worse outcome with ABVD in patients with elevated baseline levels of SCD30 and TARC. Hematol Oncol. 2019;37(Suppl 2):290-1.
112. Suri A, Mould DR, Song G, Collins GP, Endres CJ, Gomez-Navarro J, et al. Population Pharmacokinetic Modeling and Exposure-Response Assessment for the Antibody-Drug Conjugate Brentuximab Vedotin in Hodgkin's Lymphoma in the Phase III ECHELON-1 Study. Clin Pharmacol Ther. 2019;106(6):1268-79.
113. Engert A, Josting A, Haverkamp H, Villalobos M, Lohri A, Sokler M, et al. Epoetin alfa in patients with advanced-stage Hodgkin's lymphoma: results of the randomized placebo-controlled GHSG HD15EPO trial. J Clin Oncol. 2010;28(13):2239-45.
114. Kobe C, Dietlein M, Franklin J, Markova J, Lohri A, Amthauer H, et al. Positron emission tomography has a high negative predictive value for progression or early relapse for patients with residual disease after first-line chemotherapy in advanced-stage Hodgkin lymphoma. Blood. 2008;112(10):3989-94.
115. Kreissl S, Mueller H, Goergen H, Mayer A, Brillant C, Behringer K, et al. Cancer-related fatigue in patients with and survivors of Hodgkin's lymphoma: a longitudinal study of the German Hodgkin Study Group. Lancet Oncol. 2016;17(10):1453-62.
116. Borchmann S, Muller H, Hude I, Fuchs M, Borchmann P, Engert A. Thrombosis as a treatment complication in Hodgkin lymphoma patients: a comprehensive analysis of

three prospective randomized German Hodgkin Study Group (GHSG) trials. Ann Oncol. 2019;30(8):1329-34.
117. Eich HT, Gossmann A, Engert A, Kriz J, Bredenfeld H, Hansemann K, et al. A Contribution to solve the problem of the need for consolidative radiotherapy after intensive chemotherapy in advanced stages of Hodgkin's lymphoma--analysis of a quality control program initiated by the radiotherapy reference center of the German Hodgkin Study Group (GHSG). Int J Radiat Oncol Biol Phys. 2007;69(4):1187-92.
118. Chevreux S, de Barros S, Laurent C, Durand A, Delpierre C, Robert P, et al. Influence of Sociodemographic Determinants on the Hodgkin Lymphoma Baseline Characteristics in Long Survivors Patients Enrolled in the Prospective Phase 3 Trial AHL2011. Cancers (Basel). 2022;15(1).
119. Ferdinandus J, van Heek L, Roth K, Dietlein M, Eich H-T, Baues C, et al. Patterns of PET-positive residual tissue at interim restaging and risk of treatment failure in advanced-stage Hodgkin's lymphoma: an analysis of the randomized phase III HD18 trial by the German Hodgkin Study Group. Eur J Nucl Med Mol Imaging. 2024;51(2):490-5.
120. Anonymous. Erratum: standard and increased-dose BEACOPP chemotherapy compared with COPP-ABVD for advanced Hodgkin's disease (New England Journal of Medicine (June 12, 2003) 348 (2386-2395)). N Engl J Med. 2005;353(7):744.
121. Aleman BM, Raemaekers JM, Tirelli U, Bortolus R, van 't Veer MB, Lybeert ML, et al. Involved-field radiotherapy for advanced Hodgkin's lymphoma. N Engl J Med. 2003;348(24):2396-406.
122. Behringer K, Goergen H, Muller H, Thielen I, Brillant C, Kreissl S, et al. Cancer-Related Fatigue in Patients With and Survivors of Hodgkin Lymphoma: The Impact on Treatment Outcome and Social Reintegration. J Clin Oncol. 2016;34(36):4329-37.
123. Rutherford SC, Li H, Herrera A, Leblanc M, Ahmed S, Davison Kl, et al. 181 Nivolumab-AVD Is Better Tolerated and Improves Progression-Free Survival Compared to Bv-AVD in Older Patients (Aged ≥60 Years) with Advanced Stage Hodgkin Lymphoma Enrolled on SWOG S1826. In: Hematology ASo, editor. ASH Annual Meeting and Exposition; San Diego, CA, December 9. OUT ABS OF INTERIM: ASH; 2023.
124. Zallio F, Tamiazzo S, Monagheddu C, Merli F, Ilariucci F, Stelitano C, et al. Reduced intensity VEPEMB regimen compared with standard ABVD in elderly Hodgkin lymphoma patients: results from a randomized trial on behalf of the Fondazione Italiana Linfomi (FIL). Br J Haematol. 2016;172(6):879-88.
125. Herrera AF, LeBlanc ML, Castellino SM, Li H, Rutherford SC, Evens AM, et al. SWOG S1826, a randomized study of nivolumab(N)-AVD versus brentuximab vedotin(BV)-AVD in advanced stage (AS) classic Hodgkin lymphoma (HL). J Clin Oncol. 2023;41(Suppl 17):LBA4-LBA.
126. Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G, et al. Preferred Reporting Items for Systematic Review and Meta-Analyses of individual participant data: the PRISMA-IPD Statement. JAMA. 2015;313(16):1657-65.
127. Casasnovas O, Racape J, Dechene J, Bouabdallah R, Brice P, Lazarovici J, et al. PET-Guided Strategy Improves the Safety of Beacopp-Based Treatment in Advanced Hodgkin Lymphoma: Prolonged Follow-up of the Lysa Ahl 2011 Phase 3 Study. Blood. 2020;136 (Suppl 1):23-4.
128. Forero-Torres A, Holkova B, Goldschmidt J, Chen R, Olsen G, Boccia RV, et al. Phase 2 study of frontline brentuximab vedotin monotherapy in Hodgkin lymphoma patients aged 60 years and older. Blood. 2015;126(26):2798-804.
129. Gibb A, Pirrie S, Linton K, Paterson K, Davies A, Collins G, et al. Results of a phase II study of brentuximab vedotin in the first line treatment of Hodgkin lymphoma patients

considered unsuitable for standard chemotherapy (BREVITY). Hematol Oncol. 2017;35(Suppl 2):80-1.
130. Yasenchak CA, Bordoni RE, Patel-Donnelly D, Larson T, Goldschmidt JH, Boccia RV, et al. Frontline brentuximab vedotin as monotherapy or in combination for older Hodgkin lymphoma patients. J Clin Oncol. 2020;38(15 suppl):8032-.
131. Ghesquieres H, Casasnovas R-O, Nicolas-Virelizier E, Kanoun S, Damaj GL, Delwail V, et al. Final Analysis of a Prospective Multicenter Phase II Trial of the Lymphoma Study Association (LYSA) Using Prednisone, Vinblastine, Doxorubicin and Bendamustine (PVAB) Regimen in First Line Therapy for Patients over 60 Years with Advanced-Stage Classical Hodgkin Lymphoma. Blood. 2021;138(Suppl 1):1370-.
132. Fiad NL, Prates MV, Fernandez I, Kurgansky N, Cerutti A, Aranguren FN, et al. Safety and efficacy analysis in older patients treated within the GATLA LH-05 protocol: PET-adapted therapy after 3 cycle of ABVD for all stages of Hodgkin Lymphoma. HemaSphere. 2022;6(Suppl 3):1872-3.
133. Makita S, Kusumoto S, Maeshima AM, Hashimoto H, Tsujimura H, Uchida T, et al. Clinical outcomes of elderly patients with advanced-stage classic Hodgkin lymphoma who received an ABVD regimen: A multi-center retrospective study in Japan (Horizon Study). Blood. 2021;138(Suppl 1):2451.
134. Gallamini A, Rambaldi A, Bijou F, Perrot A, Patti C, Gastaud L, et al. A PHASE 1/2 clinical trial of brentuximab-vedotin and bendamustin in elderly patients with previously untreated advanced hodgkin lymphoma (HALO STUDY. NCT identifier : 02467946): Preliminary report. Blood Conference: 58th Annual Meeting of the American Society of Hematology, ASH December 3-6, Sand Diego, CA. 2016;128(22).
135. Viviani S, Fontanet B, Rambaldi A, Perrot A, Patti C, Gastaud L, et al. Halo study: A phase 1/2 clinical trial of brentuximab-vedotin and bendamustin in elderly patients with previously untreated advanced Hodgkin lymphoma. Preliminary data. Haematologica. 2016;101(Suppl 5):24-5.
136. Gallamini A, Schiano de Colella J, Viotti J, Bijou F, Rambaldi A, Perrot A, et al. Brentuximab-vedotin and bendamustine is a feasible and effective drug combination as first-line treatment of Hodgkin lymphoma in in the elderly (HALO Trial). Hematol Oncol. 2019;37(S2):292-3.
137. De Colella JMS, Viviani S, Rapezzi D, Patti C, Filliatre LC, Rossi A, et al. Brentuximab vedotin and bendamustine as first-line treatment of Hodgkin lymphoma in the elderly (HALO Trial). J Clin Oncol. 2020;38(15).
138. LeBlanc ML, Castellino SM, Li H, Rutherford SC, Evens AM, Davison K, et al. SWOG S1826, a randomized study of nivolumab(N)-AVD versus brentuximab vedotin(BV)-AVD in advanced stage (AS) classic Hodgkin lymphoma (HL). J Clin Oncol. 2023;41(Suppl 17):LBA4-LBA.

Appendix 1: Affiliations and Conflict of Interest Declarations

In accordance with the PEBC Conflict of Interest (COI) Policy, the evidence summary authors and internal reviewers were asked to disclose potential conflicts of interest.

Name	Affiliation	Declarations of interest
Working Group		
Anca Prica	University Health Network, Toronto, Ontario	$500 or more in a single year to act in a consulting capacity (Astra Zeneca, Gilead)[1]
Matt Cheung	Sunnybrook Health Sciences Centre Toronto, Ontario	None declared[2]
Lee Mozessohn	Sunnybrook Health Sciences Centre, Toronto, Ontario	$500 or more in a single year to act in a consulting capacity (Advisory Board Abbvie, December 2023)[2]
Lisa Hicks	Saint Michael Hospital, Toronto, Ontario	None declared[2]
Carolyn Faught	The Ottawa Hospital, Ottawa, Ontario	None declared[2]
Amit Singnurkar	Sunnybrook Health Sciences Centre Toronto, Ontario	None declared[2]
Danielle Rodin	Radiation Medicine Program, University Health Network, Toronto, Ontario	Received a grant/research support from the Leukemia and Lymphoma society[2]
Adam Suleman	University Health Network, Toronto, Ontario	None declared[2]
Fulvia Baldassarre	Program in Evidence-Based Care, McMaster University Hamilton, Ontario	None declared[2]
Members of the Patient Consultation Group		
Lauri Petz	NA	
Lise Craig	NA	
Bob Tuck	NA	

NA=not applicable
1 Initial conflict of interest declaration (April 27, 2021)
2 Final Conflict of interest declaration (August 14, 2024)

Appendix 2: Literature Search Strategy

Guidelines

Table 1. Search of guideline developers' websites and repositories

Site	Date searched	Terms used	Guidelines documents of potential interest identified
Alberta Health services https://www.albertahealthservices.ca/info/cancerguidelines.aspx	March 18. 2021	Hodgkin lymphoma	Alberta CC. Lymphoma: Clinical Practice Guideline LYHE-002 V12. Available at www.ahs.ca/guru. 2019.
ASCO guidelines in development https://www.asco.org/practice-guidelines/quality-guidelines/guidelines-tools-resources/guidelines-development	March 18, 2021	NA	No relevant guidelines identified
ASCO http://ascopubs.org/jco/site/misc/specialarticles.xhtml	March 18, 2021	NA searched under hematological cancers	No relevant guidelines identified
ESMO European Society for Medical Oncology - ESMO	May 15, 2018	Hodgkin lymphoma	Hodgkin's Lymphoma: ESMO Clinical Practice Guidelines Eichenauer DA, Aleman BMP, André M, Federico M, Hutchings M, Illidge T, et al. Hodgkin lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2018;29:iv19-iv29
NICE Evidence search https://www.evidence.nhs.uk//guidance	March 18, 2021	Hodgkin lymphoma in Title	Brentuximab vedotin for untreated advanced Hodgkin lymphoma (terminated appraisal) Ref N. TA594; Published: 14 August, 2019; last updated: 14 August, 2019
NICE Evidence search https://www.evidence.nhs.uk//guidance	March 18, 2021	Hodgkin lymphoma in Title	Pembrolizumab for treating relapsed or refractory classical Hodgkin lymphoma. Ref N. TA540; Published: 03 Sept 2018; last updated: : 03 Sept 2018
NICE Evidence search https://www.evidence.nhs.uk//guidance	March 18, 2021	Hodgkin lymphoma in Title	Brentuximab vedotin for treating CD30-positive Hodgkin lymphoma Ref N. TA524; Published 13 June, 2018; last updated: 13 June 2018
NICE UK https://www.nice.org.uk/search?q=hodgkin	March 18, 2021	Hodgkin lymphoma in Title	Nivolumab for treating relapsed or refractory classical Hodgkin lymphoma (TA462) Evidence-based recommendations on nivolumab (Opdivo) for treating relapsed or refractory classical Hodgkin lymphoma in adults Technology appraisal guidance Published July 2017 Last updated November 2017
ECRI ECRI Guidelines Trust® **Former National Guideline Clearinghouse**	March 19, 2021	Hodgkin lymphoma	Hodgkin lymphoma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Guideline ID: 19552018 Oct 1 • European Society for Medical Oncology (ESMO)
ECRI ECRI Guidelines Trust® **Former National Guideline Clearinghouse**	March 19, 2021	Hodgkin lymphoma	Hodgkin lymphoma Lymphoma. Guideline ID: 11192018 Jul LAST REVIEWED 2019 • Cancer Care Alberta
CPAC Database: https://www.partnershipagainstcancer.ca/tools/cancer-guidelines-database/	March 19, 2021	Hematological malignancies general search	No guidelines identified
SIGN (UK) - SIGN Guidelines Home (sign.ac.uk)	March 19, 2021	Hodgkin lymphoma	No guidelines identified

Site	Date searched	Terms used	Guidelines documents of potential interest identified
Clinical Practice Guidelines Portal | Australian Clinical Practice Guidelines (clinicalguidelines.gov.au)	March 19, 2021	Hodgkin lymphoma	No guidelines identified
Cancer Council Australia - Cancer Guidelines Wiki	March 19, 2021	Hodgkin lymphoma	No guidelines identified
MEDLINE	April 7, 2021	See Table 2	47 documents
EMBASE	April 7, 2021	See Table 2	16 documents

Table 2 Search for guidelines in MEDLINE and EMBASE

Database: OVID Medline Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to Present
Search Strategy: Executed on April 7, 2021
---
1 exp clinical pathway/
2 exp clinical protocol/
3 exp consensus/
4 exp consensus development conference/
5 exp consensus development conferences as topic/
6 critical pathways/
7 exp guideline/
8 guidelines as topic/
9 exp practice guideline/
10 practice guidelines as topic/
11 health planning guidelines/
12 (guideline or practice guideline or consensus development conference or consensus development conference, NIH).pt.
13 (position statement* or policy statement* or practice parameter* or best practice*).ti,ab,kf,kw.
14 (standards or guideline or guidelines).ti,kf,kw.
15 ((practice or treatment* or clinical) adj guideline*).ab.
16 (CPG or CPGs).ti.
17 consensus*.ti,kf,kw.
18 consensus*.ab. /freq=2
19 ((critical or clinical or practice) adj2 (path or paths or pathway or pathways or protocol*)).ti,ab,kf,kw.
20 recommendat*.ti,kf,kw.
21 (care adj2 (standard or path or paths or pathway or pathways or map or maps or plan or plans)).ti,ab,kf,kw.
22 (algorithm* adj2 (screening or examination or test or tested or testing or assessment* or diagnosis or diagnoses or diagnosed or diagnosing)).ti,ab,kf,kw.
23 (algorithm* adj2 (pharmacotherap* or chemotherap* or chemotreatment* or therap* or treatment* or intervention*)).ti,ab,kf,kw.
24 or/1-23

25 ((((systematic or state-of-the-art or scoping or literature or umbrella) adj (review or overview- or assessment)) or "review- of reviews" or meta-analy-OR metaanaly or (systematic or evidence)) adj1 assess).mp. or ("research evidence" or metasynthe or meta-synthe).tw. or exp Review Literature as Topic/ or exp Review/ or Meta-Analysis as Topic/ or Meta-Analysis/ or "systematic review"/ (2912475)
26 24 and 25
27 exp Hodgkin Disease/
28 Hodgkin$.tw,kf,ot.
29 ((malignan$ or advanced) adj2 (lymphoma and Hodgkin$)).tw.
30 27 or 28 or 29
31 ("non-Hodgkin" or "mantle cell").tw.
32 30 not 31
33 26 and 32
34 limit 33 to english language
35 limit 34 to yr="2000 -Current"
36 (comment or letter or editorial or news or newspaper article or case reports or historical article).pt.
37 35 not 36
38 exp animal/ not (exp human/ or humans/)
39 37 not 38

Database: Embase <1996 to 2021 April 07>
Search Strategy:
---
1 exp clinical pathway/
2 exp clinical protocol/
3 exp consensus/
4 exp consensus development conference/
5 exp consensus development conferences as topic/
6 critical pathways/
7 guidelines as topic/
8 exp practice guideline/
9 practice guidelines as topic/
10 health planning guidelines/
11 (position statement* or policy statement* or practice parameter* or best practice*).ti,ab,kw.
12 (standards or guideline or guidelines).ti,kw.
13 ((practice or treatment* or clinical) adj guideline*).ab.
14 (CPG or CPGs).ti.
15 consensus*.ti,kw.
16 consensus*.ab. /freq=2
17 ((critical or clinical or practice) adj2 (path or paths or pathway or pathways or protocol*)).ti,ab,kw.
18 recommendat*.ti,kw.
19 (care adj2 (standard or path or paths or pathway or pathways or map or maps or plan or plans)).ti,ab,kw.
20 (algorithm* adj2 (screening or examination or test or tested or testing or assessment* or diagnosis or diagnoses or diagnosed or diagnosing)).ti,ab,kw.
21 (algorithm* adj2 (pharmacotherap* or chemotherap* or chemotreatment* or therap* or treatment* or intervention*)).ti,ab,kw.
22 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21

23 ((((systematic or state-of-the-art or scoping or literature or umbrella) adj (review or overview- or assessment)) or "review- of reviews" or meta-analy-OR metaanaly or (systematic or evidence)) adj1 assess).mp. or ("research evidence" or metasynthe or meta-synthe).tw. or systematic review/ or "systematic review (topic)"/ or meta analysis/ or "meta analysis (topic)"/ [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word]
24 22 and 23
25 hodgkin lymphoma.mp. or exp Hodgkin disease/
26 ((malignan$ or advanced) adj2 (lymphoma and Hodgkin$)).tw.
27 ("Non-Hodgkin" or "mantle cell").tw.
28 25 or 26
29 28 not 27
30 24 and 29
31 limit 30 to english language

Systematic Reviews

Table 3. Search of electronic databases

Site	Date searched	Terms used	Systematic review documents of potential interest identified
MEDLINE	August 12, 2021	See Table 4	340 documents published from 2017 to 2021
EMBASE	August 12, 2021	See Table 4	15 documents published from 2017 to 2021
Cochrane Database of Systematic Reviews	August 13, 2021	Hodgkin lymphoma	6 documents published from 2017 to 2021, and 15 documents published from 2005 to 2015 (as these are subject to updates they are listed here).
EPISTEMONIKOS	August 13, 2021	Hodgkin lymphoma	6 documents published from 2017 to 2021; 5 documents published in 2016
AHRQ	August 16, 2021	Hodgkin lymphoma	No documents found
PROSPERO	August 16, 2021	Hodgkin lymphoma	No documents found
CADTH	August 17, 2021	Hodgkin lymphoma	No documents found
NIHR HTA (UK)	August 17, 2021	Hodgkin lymphoma	No documents found

AHRQ=Agency for Healthcare Research and Quality; CADTH=Canadian Agency for Drugs and Technologies in Health; EMBASE=Excerpta Medica Database; EPISTEMONIKOS=Health evidence database; MEDLINE=Medical Literature Analysis and Retrieval System Online; NIHR HTA=National Institute for Health and Care Research, Health Technology Assessment; PROSPERO=International prospective register of systematic reviews

Table 4. Search for systematic reviews in MEDLINE and EMBASE

Database: Embase <1996 to 2021 August 12>, OVID Medline Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to Present
Search Strategy:
---
1 exp Hodgkin Disease/
2 hodgkin.mp.
3 1 or 2
4 (systematic adj (review: or overview:)).mp.
5 (meta-analy: or metaanaly:).mp.
6 (pooled analy: or statistical pooling or mathematical pooling or statistical summar: or mathematical summar: or quantitative synthes?s or quantitative overview:).mp.
7 (exp review literature as topic/ or review.pt. or exp review/) and systematic.tw.
8 (cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinhal or cinahl or science citation index or scisearch or bids or sigle or cancerlit or pubmed or pub-med or medline or med-line).ab.
9 4 or 5 or 6 or 7 or 8
10 (selection criteria or data extract: or quality assess: or jadad score or jadad scale or methodologic: quality).ab.
11 (stud: adj1 select:).ab.
12 (10 or 11) and review.pt.

13 9 or 12
14 (comment or letter or editorial or news or newspaper article or case reports or historical article).pt.
15 13 not 14
16 exp animals/ not humans.mp. [mp=ti, ab, hw, tn, ot, dm, mf, dv, kw, fx, dq, nm, kf, ox, px, rx, ui, sy]
17 15 not 16
18 3 and 17
19 limit 18 to (english language and yr="2017 -Current")
20 hodgkin.mp.
21 1 or 20
22 17 and 21
23 limit 22 to english language
24 limit 23 to yr="2017 -Current"

Primary studies

Table 5. Search strings for randomized controlled trials in MEDLINE and EMBASE

Database: OVID Medline Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to Present. Search executed on November 5, 2021

1 exp Hodgkin Disease/
2 (hodgkin adj2 lymphoma).mp.
3 1 or 2
4 exp randomized controlled trials as topic/ or exp clinical trials, phase III as topic/ or exp clinical trials, phase IV as topic/
5 (randomized controlled trial or clinical trial, phase III or clinical trial, phase IV).pt.
6 random allocation/ or double blind method/ or single blind method/
7 (randomi$ control$ trial? or rct or phase III or phase IV or phase 3 or phase 4).tw.
8 4 or 5 or 6 or 7
9 (phase II or phase 2).tw. or exp clinical trial/ or exp clinical trial as topic/
10 (clinical trial or clinical trial, phase II or controlled clinical trial).pt.
11 (9 or 10) and random$.tw.
12 (clinic$ adj trial$1).tw.
13 ((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3 or dummy)).tw.
14 placebos/
15 (placebo? or random allocation or randomly allocated or allocated randomly).tw.
16 (allocated adj2 random).tw.
17 12 or 13 or 14 or 15 or 16
18 8 or 11 or 17
19 (comment or letter or editorial or news or newspaper article or patient education handout or case reports or historical article).pt.
20 18 not 19
21 exp animals/ not humans/
22 20 not 21
23 3 and 22
24 limit 23 to (english language and yr="1998 -Current")

Database: Embase <1996 to 2021 November 05>

---

1 exp Hodgkin disease/
2 (hodgkin adj2 lymphoma).mp.
3 1 or 2
4 exp randomized controlled trial/ or exp phase 3 clinical trial/ or exp phase 4 clinical trial/
5 randomization/ or single blind procedure/ or double blind procedure/

6 (randomi$ control$ trial? or rct or phase III or phase IV or phase 3 or phase 4).tw.
7 4 or 5 or 6
8 (phase II or phase 2).tw. or exp clinical trial/ or exp prospective study/ or exp controlled clinical trial/
9 8 and random$.tw.
10 (clinic$ adj trial$1).tw.
11 ((singl$ or doubl$ or treb$ or tripl$) adj (blind$3 or mask$3 or dummy)).tw.
12 placebo/
13 (placebo? or random allocation or randomly allocated or allocated randomly).tw.
14 (allocated adj2 random).tw.
15 10 or 11 or 12 or 13 or 14
16 7 or 9 or 15
17 3 and 16
18 (editorial or note or letter or short survey).pt. or abstract report/ or letter/ or case study/
19 17 not 18
20 exp animal/ not human/
21 19 not 20
22 limit 21 to (english language and yr="1998 -Current")

Table 6 – Search strings for observational trials in MEDLINE and EMBASE

Database: OVID Medline Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to Present. Search executed on September 23, 2022

1. Hodgkin Disease/
2. (hodgkin adj2 lymphoma).mp.
3. (malignant or advanced or (("stage IIB" or "stage III" or "stage IV") adj Hodgkin)).tw.
4. (1 or 2) and 3
5. ("Non-Hodgkin" or "mantle cell").tw.
6. 4 not 5
7. exp Aged/
8. elderly.tw,kw.
9. geriatric*.tw,kw.
10. senior.tw,kw.
11. (older adj (adult? or m#n or wom#n or person? or people)).tw,kw.
12. Geriatrics/
13. (Age? adj3 (over or older) adj2 (5# or 6# or 7# or 8# or 9#)).tw.
14. sexagenarian.tw,kw.
15. septuagenarian.tw,kw.
16. octogenarian.tw,kw.
17. nonagenarian.tw,kw.
18. centenarian.tw,kw.
19. gerontolog*.tw,kw.

20. (">=5# years old" or ">5# years old").tw.
21. (">=6# years old" or ">6# years old").tw.
22. (">=7# years old" or ">7# years old").tw.
23. (">=8# years old" or ">8# years old").tw.
24. (">=9# years old" or ">9# years old").tw.
25. 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24
26. 6 and 25
27. Epidemiologic studies/
28. exp case control studies/
29. exp cohort studies/
30. Case control.tw.
31. (cohort adj (study or studies)).tw.
32. Cohort analy$.tw.
33. (Follow up adj (study or studies)).tw.
34. (observational adj (study or studies)).tw.
35. Longitudinal.tw.
36. Retrospective.tw.
37. 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36
38. 26 and 37
39. (animals/ not humans/) or comment/ or editorial/ or exp review/ or meta analysis/ or consensus/ or exp guideline/ or hi.fs. or case report.mp.
40. 38 not 39
41. limit 40 to (english language and yr="1998-Current")
42. limit 41 to (address or autobiography or bibliography or biography or case reports or clinical conference or clinical trial, veterinary or comment or congress or consensus development conference or consensus development conference, nih or dataset or dictionary or directory or editorial or festschrift or guideline or historical article or interactive tutorial or interview or lecture or legal case or legislation or letter or meta analysis or news or newspaper article or observational study, veterinary or overall or patient education handout or periodical index or personal narrative or portrait or practice guideline or randomized controlled trial, veterinary or "review" or "scientific integrity review" or "systematic review" or validation study)
43. 41 not 42

Database: Embase <1996 to 2022 September 23>

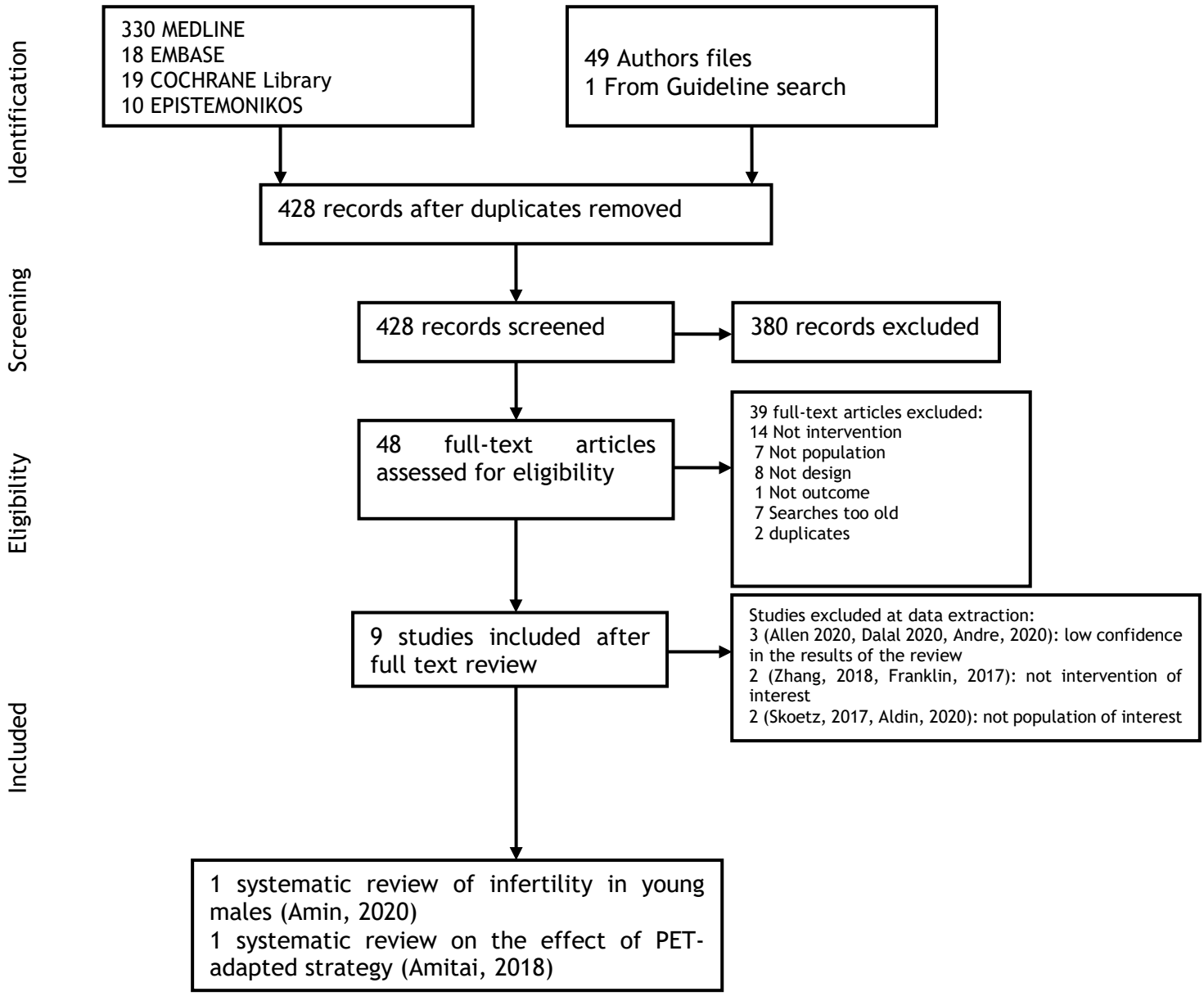
1. exp Hodgkin disease/

2. (Hodgkin adj2 lymphoma).mp.
3. (malignant or advanced or (("stage IIB" or "stage III" or "stage IV") adj Hodgkin)).tw.
4. (1 or 2) and 3
5. ("Non-Hodgkin" or "mantle cell").tw.
6. 4 not 5
7. exp aged/
8. elderly.tw,kw.
9. geriatric*.tw,kw.
10. senior.tw,kw.
11. (older adj (adult? or m#n or wom#n or person? or people)).tw,kw.
12. exp geriatrics/
13. (Age? adj3 (over or older) adj2 (5# or 6# or 7# or 8# or 9#)).tw.
14. sexagenarian.tw,kw.
15. septuagenarian.tw,kw.
16. octogenarian.tw,kw.
17. nonagenarian.tw,kw.
18. centenarian.tw,kw.
19. gerontolog*.tw,kw.
20. (">=5# years old" or ">5# years old").tw.
21. (">=6# years old" or ">6# years old").tw.
22. (">=7# years old" or ">7# years old").tw.
23. (">=8# years old" or ">8# years old").tw.
24. (">=9# years old" or ">9# years old").tw.
25. 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24
26. 6 and 25
27. clinical study/
28. case control study/
29. longitudinal study/
30. retrospective study/
31. prospective study/
32. "randomized controlled trial (topic)"/
33. 31 not 32
34. cohort analysis/

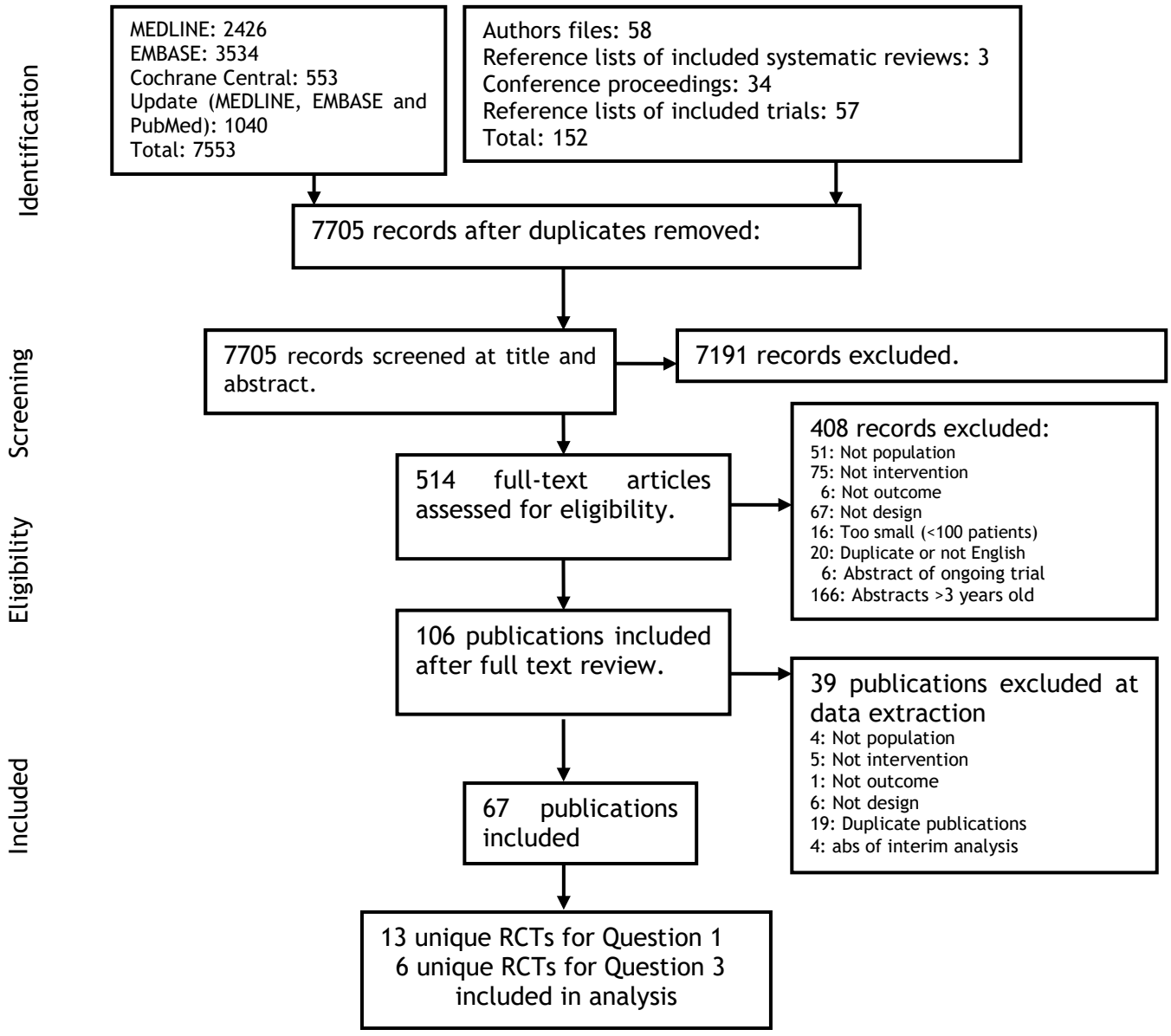
35. (Cohort adj (study or studies)).mp.
36. (Case control adj (study or studies)).tw.
37. (follow up adj (study or studies)).tw.
38. (observational adj (study or studies)).tw.
39. (epidemiologic$ adj (study or studies)).tw.
40. 27 or 28 or 29 or 30 or 33 or 34 or 35 or 36 or 37 or 38 or 39
41. 26 and 40
42. exp animal/ not human/
43. 41 not 42
44. limit 43 to (english language and yr="1998 -Current")
45. limit 44 to ((consensus development or meta analysis or "systematic review") and (books or chapter or conference abstract or conference paper or "conference review" or editorial or letter or note or "review" or short survey or tombstone) and (book or book series or conference proceeding or trade journal))
46. 44 not 45

Appendix 3: PRISMA Flow Diagram

A) Systematic reviews

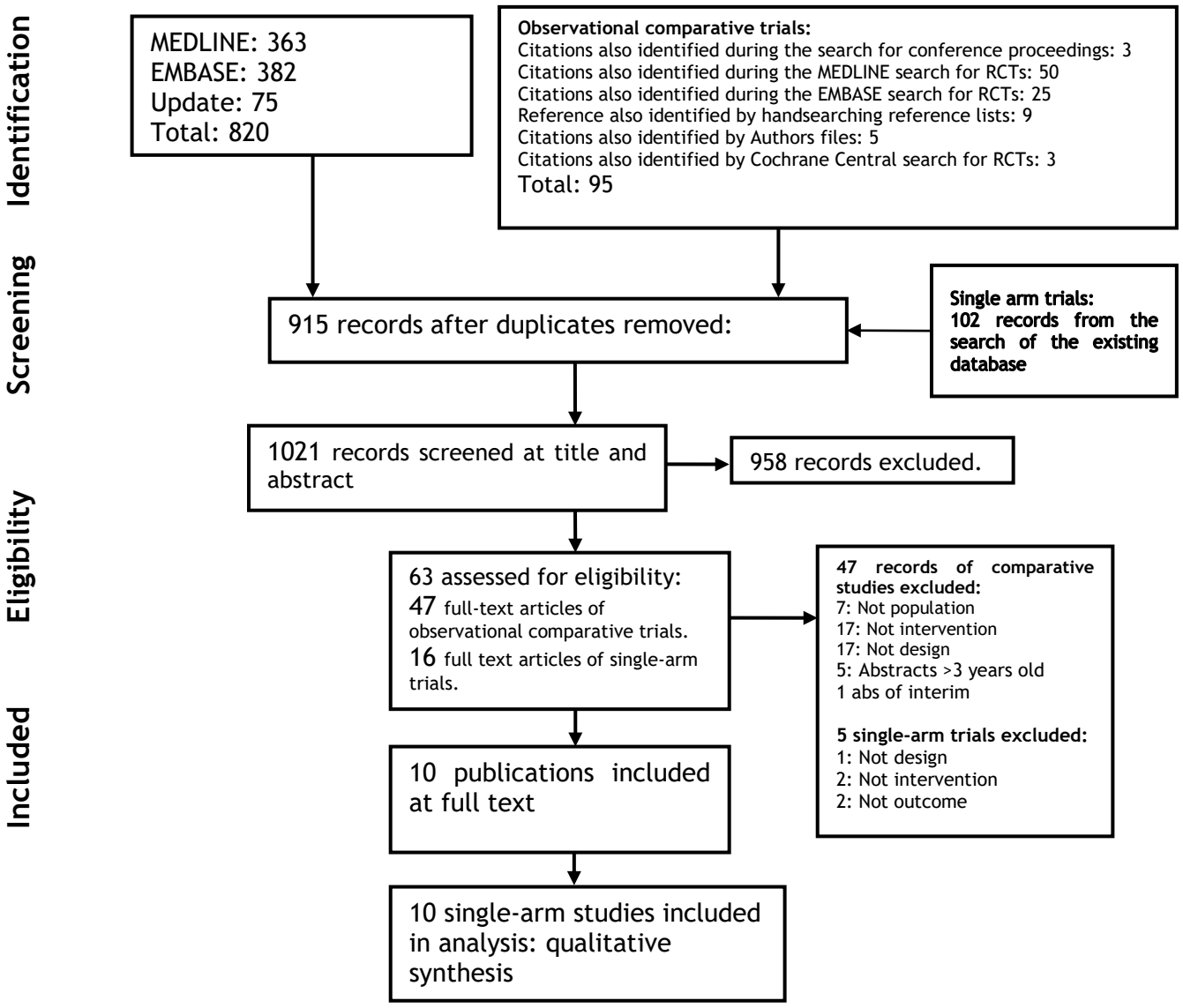


B) Primary studies – RCTs included for Questions 1 and 3 This is updated as of February 7, 2024



Note: only fully published studies are accounted for in this chart, because not enough information is in the abstract publication to evaluate the studies' quality.

C) Observational trials: Question 2 only; sample size n=30, randomized and observational comparative design and phase 2 single arm trials.



Appendix 4 Selection Criteria

Inclusion	
Population	Adult patients (≥18 years of age) with advanced HL (stages II with adverse features, III and IV) who need first-line treatment. For question 1 patients <60 years of age; for question 2 patients ≥60 years of age. For studies that included a mixed population (i.e., containing patients with advanced-stage and early-stage HL, or with NHL), and where data for the advanced HL cohort are not reported separately, only study with a sample of at least 80% of patients with advanced-stage HL will be included. For studies that included a mixed population (i.e., patients 60 years old and older and patients younger than 60 years of age) and where results are not reported separately for the different age groups: we will include in Question 1 when ≥80% of patients are younger of 60 years of age or if the median age is <60 years. We will include in Question 2 if ≥80% of patients are ≥60 years of age or if the median age is ≥60 years. Systematic reviews of adverse events will be included even if the population includes patients with HL, and there is no separate analysis for patients with advanced HL. Studies of patients with nodular lymphocyte predominant HL will also be excluded.
Intervention	Included treatments: Question 1: ABVD eBEACOPP A-AVD Any other novel / experimental treatment Question 2: ABVD A-AVD Brentuximab alone or in combination PVAG CHOP CHLVPP Other / experimental novel treatments Question 3 Consolidation radiotherapy: 30 Gy in 20 fractions
Comparison	Included comparisons: **Questions 1 and 2**: Alternative treatments among the interventions listed above **Question 3:** No consolidation radiotherapy
Outcomes	Included outcomes: Question 1:

	• OS* • Recurrence* (includes PFS and other measures)* • AE* (i.e., hematological toxicity and secondary cancers, infertility) • QOL**, Time: 5 years for OS, and 20 years for AE (second cancers, including solid tumours and AML and MDS) Clinical threshold: >10% difference in PFS would be a clinically significant difference Question 2: • OS* • Recurrence* (includes PFS and other measures)* • AE (hematological toxicity and secondary cancers)* • QOL* Time: 5 years for OS, and 20 years for AE (second cancers, including solid tumours and AML and MDS) Clinical threshold: >10% difference in PFS would be a clinically significant difference Question 3: • OS* • Recurrence (includes PFS and other measures)* • AE* • QOL** Time: 5 years for OS, and 20 years for AE (second cancers, including solid tumours and AML and MDS) Clinical threshold: >10% difference in PFS would be a clinically significant difference
Design	Systematic reviews of low risk of bias (as measured with the ROBIS tool) or a moderate/high CERTAINTY overall rating as assessed with the AMSTAR 2. RCTs with a sample ≥100 patients for questions 1 and 3. RCTs with a sample of ≥30 patients for question 2. Observational comparative trials that controlled for confounding with a sample ≥30 patients for the comparisons for which RCTs are not available. Phase 2 single arm studies for the comparisons for which comparative studies were not available.
Cut off date	Systematic reviews published from 2016 to 2021 RCTs and observational trials published from 1998 to 2021
Exclusion	
• Letters, comments, editorials, non-English publications, non-RCTs IF RCTS ARE AVAILABLE, abstracts of ongoing trials, abstracts of systematic reviews • Studies of > 80% pediatric patients • Studies that did not report on the population of interest • Studies that did not include the interventions of interest; this includes studies of interventions that are now obsolete • Studies that did not report on the outcomes of interest	

- Studies published before the cut-off date, or systematic reviews with a search date before 2016

*Critical outcomes
**Important outcomes

A-AVD=Brentuximab vedotin, doxorubicin, vinblastine, dacarbazine; ABVD=doxorubicin, bleomycin, vinblastine, and dacarbazine; AE=Adverse events; AML=Adult acute myeloid leukemia; AMSTAR=Assessing the methodology quality of Systematic Reviews; CHLVPP=chlorambucil, vinblastine, procarbazine, prednisolone; CHOP=Cyclophosphamide, doxorubicin, vincristine and prednisolone; eBEACOPP=escalated BEACOPP; HL=Hodgkin lymphoma; MDS=Myelodysplastic syndromes; NHL=Non-Hodgkin lymphoma; OS=Overall survival; PFS=Progression-free survival; PVAG=prednisone, vinblastine, doxorubicin, gemcitabine; QOL=Quality of life; RCT=Randomized controlled trial

Appendix 5: Quality assessment

Table 1-5A. Systematic reviews assessed with the AMSTAR 2 tool [29]

Study	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12 yes (only low-risk)	Q13	Q14	Q15	Q16	Overall confidence in the results of the review
Skoetz 2017 [6]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	high
Aldin, 2020 [41]	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	high
Amitai, 2018 [42]	yes	no	yes	yes	yes	yes	yes	yes	yes	no	NA	NA	yes	yes	na	yes	high
Amin, 2020 [43]	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	NA	NA	yes	NA	NA	yes	high
Dalal, 2020 [44]	yes	yes	no	yes	yes	yes	yes	yes	yes	no	NA	na	no	NA	NA	yes	Low
Allen 2020 [47]	yes	no	yes	no	no	no	no	yes	no	no	NA	NA	no	NA	NA	yes	Critically low

NOTES
AMSTAR 2 questions.
Q1 Did the research questions and inclusion criteria for the review include the components of PICO?
Q2 Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?
Q3 Did the review authors explain their selection of the study designs for inclusion in the review?
Q4 Did the review authors use a comprehensive literature search strategy?
Q5 Did the review authors perform study selection in duplicate?
Q6 Did the review authors perform data extraction in duplicate?
Q7 Did the review authors provide a list of excluded studies and justify the exclusions?
Q8 Did the review authors describe the included studies in adequate detail?
Q9 Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?
Q10 Did the review authors report on the sources of funding for the studies included in the review?
Q11 If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?
Q12 If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?
Q13 Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?
Q14 Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?
Q15 If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?
Q16 Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?

Questions 2, 4, 7, 9, 11, 14, and 15 are considered critical items. The Overall Confidence in the results of the systematic reviews is considered High if: No weakness or one non-critical weakness is present. Moderate if: more than one non-critical weakness is present. Low if one critical flaw with or without non-critical weaknesses is present. Critically low if: more than one critical flaw with or without non-critical weaknesses are present.

Table 2-5A. Risk of bias of individual patient data meta-analyses assessed according to the CheckMAP tool [35]

CheckMAP signaling questions	Answer and relevant supporting information Franklin, 2017 [46]	Answer and relevant supporting information Andre, 2020 [11]
1) Was the IPD meta-analysis done within a systematic review framework?	Yes It is part of a Cochrane review. Both title and introduction show that this work is part of a systematic review.	Yes Page 6566-6567
2) Were all of the methods pre-specified in a publicly available protocol?	Yes Protocol of the Cochrane reviews are always published, and although the full protocol is not currently available and does not appear in PROSPERO, as the review is now completed, at the end of the review there is a paragraph called Differences between protocol and review.	No No PROSPERO number is cited, and no other mention of a protocol. I searched the internet for a protocol of this study, and I could not find it.
3) Did it have a clear research question qualified by explicit eligibility criteria for trials and participants?	Yes Study question (SQ) 1) Chemotherapy alone versus same chemotherapy plus radiotherapy SQ 2) Chemotherapy plus involved-field radiation versus same chemotherapy plus extended-field radiation SQ 3) Chemotherapy plus lower-dose radiation versus same chemotherapy plus higher-dose radiation SQ 4) Fewer versus more courses of chemotherapy SQ 5) Dose-intensified chemotherapy versus ABVD-like chemotherapy Types of studies, participants, interventions, and outcomes are specified on page 9.	Yes In paragraph 2.2, page 6567, the study objectives purpose is stated. Not a true research question(s)
4) Did it use a systematic and comprehensive search to identify trials? • Were fully published trials sought? • Were trials in the grey literature sought? • Were unpublished trials sought?	Yes MEDLINE, Cochrane Central, conference abstracts, reference lists of included trials, and international trial registries were searched. Yes to all three sub-questions.	Yes The Authors searched the Cochrane Library, MEDLINE, and 3 conference proceedings.
5) Was the approach to data collection consistent and thorough? • Were key baseline variables defined and sought? • Were key outcomes defined and sought?	Yes The authors followed the Cochrane Handbook. Type of studies, type of participants, type of interventions and types of outcomes are thoroughly described.	Yes Key outcomes are defined and sought (OS and PFS) "OS and PFS were defined as the time from the date of randomization to the date of death from any cause and to the date of disease progression or death from any cause, respectively." Baseline variables are reported Table 1 for each study and in total.
6) Were IPD obtained for most trials for most of the eligible trials and their participants? • Were IPD obtained for a large proportion of eligible trials? • Were these trials likely to be representative of all high-quality trials? • Were IPD obtained for a large proportion of eligible participants?	Yes IPD were obtained from 16 out of 21 eligible trials (76%), including a total of 9498 pts (85%). Trials that were smaller in size were not included (<50 pts per arm), and trials with recruitment beyond 2007 (as there would not be enough time for sufficient follow-up for the primary outcome – second cancer). The included trials were generally of high quality, except for blinding. Reasons for unavailability of data are reported. Reason for not obtaining IPD were not receiving data from authors. Primary outcome was occurrence of second malignant neoplasm, Secondary outcomes were OS and PFS.	No It is not stated how many trials were included by the systematic review. Page 6567: "After this literature review, four individual patient data corresponding to the selection criteria were identified."

CheckMAP signaling questions	Answer and relevant supporting information Franklin, 2017 [46]	Answer and relevant supporting information Andre, 2020 [11]
• Were the reasons for not obtaining IPD provided?		
7) Was the quality of the IPD checked for each trial? • Were the data checked for missing, invalid, out-of-range or inconsistent items? • Were the data checked for discrepancies with the trial report? • Were any issues queried with trial investigators, and either resolved or addressed?	Yes Page 10: "As suggested in Chapter eight of the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2011), we checked IPD for completeness and consistency. As a preparatory step, we analysed each trial separately, comparing the treatment arms with respect to recruitment times, patient characteristics, complete remission rate, length of follow-up, PFS, OS and occurrence of secondary malignant neoplasms. This step investigates the comparability of the treatment arms and the consistency of the data with previous publications of the trial. We assessed each trial according to the Cochrane Risk of Bias tool. Since this was an IPD analysis and all our targeted outcomes could be obtained from all included trials, selective reporting bias did nor occur. Further, the following special aspects were assessed. • Reliability of secondary malignant neoplasm follow-up methods: we assessed the method of follow-up as described by the trialists for likely completeness and accuracy. • Completeness of follow-up: we calculated the median follow-up time, using the Kaplan-Meier method, to indicate average length of follow-up. We quantified the distribution of last information dates. Both high variability (large interquartile range), in relation to the median follow-up time, and significant differences between treatment arms indicate less reliable follow-up. We also compared completeness of follow-up between patients with and without secondary malignant neoplasms. • We compared the secondary malignant neoplasm rate with that expected in an age- and sex-matched cohort from the general population, using data from various US, European and Australasian cancer registries as appropriate to the trial. These special aspects were expected to be the most problematic since secondary malignant neoplasm events were not a major endpoint for most trials. Risk of bias was considered when interpreting review results by qualitative and quantitative description as well as by a sensitivity analysis excluding incomplete follow-up periods."	No • Were data checked for missing, invalid, out-of-range or inconsistent items? It is stated that (page 6567): "OS and PFS were analyzed according to randomized treatment arm for the intent-to-treat populations. Intent-to-treat population was defined as: all randomized patients (even if later found to be ineligible) with informed consent in IIL, H34 low risk, and EORTC200012; and all patients excluding patients for whom an exclusion criterion was discovered after random assignment and patients who had insufficient follow-up information to determine treatment outcome in HD2000." • Were data checked for discrepancies with the trial report? Not reported • Were any issues queried with trial investigators and resolved or addressed? Not reported.
8) Was the risk of bias assessed for each trial and informed by checks of the associated IPD? • Was the randomization process checked based on the IPD? • Was the IPD checked to ensure that all (or most) trial participants were included? • Were all important outcomes included in the IPD?	Yes All but 2 studies, for which no publication was available as of at May 2014, were assessed for risk of bias (allocation and attrition bias). Figures 3, and 4. IPD was obtained for all other studies for the outcomes of interest, therefore, selective reporting was not considered. This included direct checking of the IPD, see "Dealing with missing data" section on page 11. The evidence from the IPD was graded according to the GRADE methods and a summary of finding table was provided.	Yes

CheckMAP signaling questions	Answer and relevant supporting information Franklin, 2017 [46]	Answer and relevant supporting information Andre, 2020 [11]
• Were the outcomes measured/defined appropriately? • Was the quality of outcome data checked?		
9) Were the methods of meta-analysis appropriate? • Were the analysis methods pre-specified in detail and the key estimands (target parameters) defined (e.g., average treatment effect; treatment co-variate interaction)? • Were the methods of summarizing overall effects of treatments appropriate? - Did researchers account for clustering of participants within trials in their meta-analysis methods? - Was the choice of one- or two-stage meta-analysis framework appropriate, and was the choice specified in advance and/or were results for both approaches provided? - Was between-trial heterogeneity in treatment effects examined and accounted for? - Were other statistical intricacies accounted for (e.g., non-proportional hazards, clustering of trials with a cluster randomized design, etc.)? • Were the methods of assessing whether effects of treatments varied by trial-level characteristics appropriate? - Did researchers estimate a treatment-covariate interaction separately for each trial and combine these across trials in a two-stage common-effect or random-effects meta-analysis? Or	Yes / Probably yes This review had a protocol that most likely pre-specified the analyses, because Cochrane reviews do have a protocol. This is not available as the review has now been completed. The Authors did a subgroup analysis to check "if certain types of patients or treatment types show different treatment effects" (i.e., they checked for clustering). Authors used a two-step approach (i.e. first, analysis of each trial separately and second, meta-analysis combination of estimates) based on the fixed-effect model (p. 11 second column), and they a Cox proportional hazards regression when they used a one-step approach (see Sensitivity analysis paragraph on page 12). In the sensitivity analysis they analyzed data for all trials making a comparison of modalities together. The Authors summarize within-trial covariate interactions. Pre-specified sensitivity analyses were conducted. The Authors investigated heterogeneity (see page 12 "Subgroup analysis and investigation of heterogeneity" paragraph). The Authors examined the treatment effects by study question for each outcome (see subgroup analysis) for age and gender (page 12 in methods, and page 21 in the results). They combined treatment covariate interaction across trials using a random effects meta-analysis.	Yes/ probably yes i. Analysis stratified by trial for the IPI score Page 6567: "Multivariate analysis was performed using a Cox regression model stratified by trial." The Authors used a one-step-method, but they did the correct analysis.

CheckMAP signaling questions	Answer and relevant supporting information Franklin, 2017 [46]	Answer and relevant supporting information Andre, 2020 [11]
- Did researchers incorporate treatment co-variate interaction terms in a one-stage regression model, whilst also accounting for clustering of participants, and separating these participant-level interactions from any trial-level associations? - Were continuous covariates analysed on their continuous scale? Were potential non-linear relationships examined? • Was the robustness of conclusions checked using relevant sensitivity or other analysis? - E.g., removing trials at high risk of bias, including aggregate data from trials that did not supply IPD, examining asymmetry on funnel plots		
10) Did the project's report cover the items described in PRISMA-IPD, [126] or explain why they were not relevant?	Yes This study was planned before the publication of the PRISMA IPD, and the Authors used the PRISMA 2009 version of that tool.	No

IPD=Individual participant data; IPI=International Prognostic Index; OS=Overall survival; PFS=Progression-free survival; PRISMA-IPD=Preferred reporting items for systematic reviews and meta-analyses for individual patient data systematic reviews; PROSPERO=International Prospective Register of Systematic Reviews; Pts=patients; SQ=Study question.

Table 3-5A.1. Risk of bias summary table, assessed with the Cochrane ROB2 [36] Questions 1 and 3

Study	Comparison	Outcomes	ROB from the randomization process	ROB from deviation from the intended interventions	Missing outcome data	ROB in measurement of the outcome	ROB in selection of the reported results	Overall ROB	Comment/ Concerns
QUESTION 1									
ABVD vs. BEACOPP									
EORTC 20012 [52]	4×eBEACOPP + 4×bBEACOPP vs. 8×ABVD	PFS, EFS, DFS	Low	Some concerns[a]	Low	Low	Low	Low for OS, PFS and TRD because the imaging was centrally reviewed and blinded Some concerns for TRD, and SPM because these were local investigator reported and subjective.	These two studies appear pooled in meta-analysis in the SOF table. OS was not affected by lack of blinding, and for other outcomes, there was no evidence or no information about deviation from the intended interventions, therefore, we scored some concerns for them.
		OS	Low	Low[b]	Low	Low	Low		
		TRD	Low	Some concerns[ac]	Low	Low	Low		
		SPM	Low	Some concerns[ac]	Low	Low	Low		
LYSA H34 [53]	4×eBEACOPP + 4×bBEACOPP vs. 8×ABVD	PFS, EFS,	Low	Some concerns[a]	Low	Some concerns[d]	Low	Low for OS, Some concerns for all other outcomes	
		OS	Low	Low[b]	Low	Low	Low		
		SPM	Low	Some concerns[a]	Low	Some concerns[d]	Low		
		TRD	nr	nr	nr	nr	nr		
		Pulmonary toxicity	Low	Some concerns[a]	Low	Some concerns[d]	Low		
HD9 [54,64,77,81] (only HD9 data)	8×COPP/ABVD vs. 8×bBEACOPP vs. 8×eBEACOPP (G-CSF)	PFS	Low	High[a]	Low	High[dce]	Low	High	Clinicians, pts and outcome assessors aware of assignment; trial stopped early for benefit and there were deviations from the intended interventions because of the trial context.
		OS	Low	Low[b]	Low	High[dce]	Low		
		TRD	Low	High[a]	Low	HIgh[dce]	Low		
		SPM	Low	High[a]	Low	High[dce]	Low		
HD2000 [55,56]	6×ABVD vs 4eBEACOPP + 2bBEACOPP vs 6× COPPEBVCAD -CEC	PFS	High[f]	High[ag]	Low	High[d]	Some concerns[h]	High for all outcomes	The study was initially designed to compare myelotoxicity referred to leukopenia of CEC and BEACOPP regimens with respect to that of ABV, then there was a change. There were baseline differences between groups; patients excluded after randomization were not
		OS	High[f]	High[ag]	Low	High[d]	Some concerns[h]		
		TRD	High[f]	High[ag]	Low	High[d]	Some concerns[h]		
		SPM	High[f]	High[ag]	Low	High[d]	Some concerns[h]		

Study	Comparison	Outcomes	ROB from the randomization process	ROB from deviation from the intended interventions	Missing outcome data	ROB in measurement of the outcome	ROB in selection of the reported results	Overall ROB	Comment/ Concerns
									included in the ITT analysis.
GITIL (Viviani) [57]	6 or 8×ABVD vs 4eBEACOPP+ 4bBEACOPP	FFFP	Low	Some concerns[a]	Low	Some concerns[d]	Low	Low for OS, Some concerns for all other outcomes	Participants, clinicians and outcome assessors were aware of assignment to interventions. This may have no impact on OS, and we rated some concerns for other outcomes because there was no evidence of deviations from intended interventions.
		OS	Low	Low	Low	Low	Low		
		TRD	Low	Some concerns[a]	Low	Some concerns[d]	Low		
		SPM	Low	Some concerns[a]	Low	Some concerns[d]	Low		
Higher-dose (more cycles) vs. Lower-dose (less cycles, lower dose) eBEACOPP									
HD15 [58,82]	8×eBEACOPP vs. 6×eBEACOPP vs. 8×BEACOPP$_{14}$	PFS	Low	Some concerns[b]	Low	High[d]	Low	Low for OS, High for all other outcomes	Patients, investigators, and outcome assessors were not blinded to assignment of interventions. This would not have an impact on the risk of bias of OS, and we rated some concerns for all other outcomes since there was no evidence of deviation from planned interventions.
		OS	Low	Low[b]	Low	Low	Low		
		TRD	Low	Some concerns[b]	Low	High[d]	Low		
		SPM	Low	Some concerns[b]	Low	High[d]	Low		
HD12 Borchmann, 2011 [59]; von Tresckow, 2018 [77] (only HD12 data)	Chemotherapy comparison (RT comparison is in Question 3): 8×eBEACOPP ± RT (n=787) vs. 4×eBEACOPP + 4×bBEACOPP ± RT (n=787)	PFS	Low[j]	Some concerns	Low	Low	Low	Low for OS, some concerns for all outcomes	Participants and clinicians were not blinded to the intervention assignment. This would not impact OS, and there was no evidence of deviations from the intended interventions. Therefore, we rated some concerns for all other outcomes.
		FFTF	Low[j]	Some concerns	Low	Low	Low		
		OS	Low[j]	Low	Low	Low	Low		
		TRD	Low[j]	Some concerns	Low	Low	Low		
		SPM	Low[j]	Some concerns	Low	Low	Low		

Study	Comparison	Outcomes	ROB from the randomization process	ROB from deviation from the intended interventions	Missing outcome data	ROB in measurement of the outcome	ROB in selection of the reported results	Overall ROB	Comment/ Concerns
PET-driven strategies									
HD18 [16,50,83]	2×eBEACOPP then PET PET-2 positive pts: 8×eBEACOPP (Standard) vs. 8×R-eBEACOPP PET-2 negative pts: 8×eBEACOPP (Standard) + 6×eBEACOPP (New Standard) vs. 4×eBEACOPP	PFS	Low	High[k]	Low	High[d]	Low	High	Pts and clinicians were aware of the assignment to interventions. No information is provided if outcome assessors were blinded. An amendment in 2011 changed the standard treatment from 8×eBEACOPP to 6×eBEACOPP, therefore, there were deviation from the intended interventions due to the trial context, and they were imbalanced between arms. Therefore, we rated this study at high risk of bias.
		OS	Low	High[k]	Low	Low	Low		
		TRD	Low	High[k]	Low	High[d]	Low		
		SPM	Low	High[k]	Low	High[d]	Low		
AHL2011 [17,127]	6×eBEACOPP vs. PET-driven treatment: All pts received 2×eBEACOPP, then PET-2.	OS	Low	Low	Some concerns[l]	Low	Low	Some concerns for OS, High for other outcomes	Pts and clinicians were aware of the intervention assignment. It is not reported if outcome assessors were blinded, and 4% of patients did not receive the allocated treatment because of clinician decision; these reasons, potentially other deviations from intended interventions, and a low percentage of missing data led us to rate this study at some concerns for OS and high risk of bias for all other outcomes.
		PFS	Low	High[a]	Some concerns[l]	High[d]	Low		
		TRD	Low	High[a]	Some concerns[l]	High[d]	Low		
		SPM	Low	High[a]	Some concerns[l]	High[d]	Low		
RATHL [49]		PFS	Low	High[ai]	Low	High[d]	Low	High	

Study	Comparison	Outcomes	ROB from the randomization process	ROB from deviation from the intended interventions	Missing outcome data	ROB in measurement of the outcome	ROB in selection of the reported results	Overall ROB	Comment/ Concerns
	PET negative after 2×ABVD: ABVD vs. AVD PET positive after 2×ABVD: BEACOPP14 vs. eBEACOPP This arm of the study was not randomized, and not considered here.	OS	Low	Low[a]	Low	High[d]	Low		Pts, clinicians and outcome assessors were aware of intervention allocation. The study was slightly underpowered. There were baseline differences in the pt population with older patients assigned to AVD. For these reasons we rated this study at high risk of bias
		Pulmonary toxicity	Low	High[ai]	Low	High[d]	Low		
		TRD	Low	High[ai]	Low	High[d]	Low		
GITIL/FIL HD0607 [48]	PET-2-positive pts only: Ritux (375 mg/m^2) + 4×eBEACOPP +4×bBEACOP P vs. 4×eBEACOPP +4×bBEACOP P PET negative pts were randomized to RT vs. no RT, see Question 3.	OS	Some concerns	High	High[m]	Low	Low	High	PET Positive: Pts and clinicians were not blinded to intervention assignment. There were deviations from the intended interventions for 7.4% of the patients. These deviations most likely affected the outcomes. There were missing data, and this might have skewed the outcomes (see Table 3-5A.1 in Appendix 5). Therefore, we rated this study at high risk of bias.
		PFS	Some concerns	High	High[m]	Low	Low		
ABVD vs. modified ABVD with brentuximab instead of bleomycin									
ECHELON-1 [9,86-89]	ABVD vs. A-AVD	PFS	Some concerns[n]	Some concerns[o]	Low	High	Low	Some concerns for OS High for other outcomes	None of the published reports and the protocol we reviewed reported

Study	Comparison	Outcomes	ROB from the randomization process	ROB from deviation from the intended interventions	Missing outcome data	ROB in measurement of the outcome	ROB in selection of the reported results	Overall ROB	Comment/ Concerns
		OS	Some concerns[n]	Some concerns[o]	Low	Low	Low		how the randomization was conducted. Deviations from the intended interventions were unlikely, but not enough information was provided to answer definitively whether there had been or not such deviations. For this reason, we rated Some concerns for OS. Pts, clinicians and outcome assessors were aware of the intervention assignment. The assessment of the outcome could have been influenced by knowledge of the intervention received. For these reasons we rated high risk of bias for other outcomes.
		TRD	Some concerns[n]	Some concerns[o]	Low	High	Low		
		Peripheral neuropathy	Some concerns[n]	Some concerns[o]	Low	High	Low		
		SPM	Some concerns[n]	Some concerns[o]	Low	High	Low		
Modified ABVD vs. another modified ABVD									
No fully published studies included									
BEACOPP vs. modified BEACOPP									
No fully published studies included									
QUESTION 3									
Consolidation RT vs. observation									
HD0801 FIL [75]	RT with the dose of 30 Gy in 2-Gy fractions vs. Observation	PFS	Low	Some concerns[a]	Low	Some concerns	Low	Some concerns for all outcomes	Clinicians and pts were aware of assignment to intervention. No information was reported on whether outcome assessors were blinded. No data were reported for OS.
		OS	Low	Low	High	Low	Low		
		SPM	Low	Some concerns[a]	Low	Some concerns	Low		
GITIL/FIL HD0607 [48,60]	Consolidation RT over large masses	PFS	Some concerns	Some concerns	Low	Low	Low	Some concerns for all outcomes	No information was reported on how the sequence allocation was

Study	Comparison	Outcomes	ROB from the randomization process	ROB from deviation from the intended interventions	Missing outcome data	ROB in measurement of the outcome	ROB in selection of the reported results	Overall ROB	Comment/ Concerns
	vs. No further treatment	OS	Some concerns	Low	Low	Low	Low		generated and if allocation was concealed. Patients and clinicians were aware of intervention allocation.
EORTC 20884 [76]	IF-RT vs. no further treatment	EFS	Low	High[aq]	High	High	High	High	Pts, clinicians, and outcome assessors were aware of intervention allocation. Pts were not analyzed in the groups where they were randomized. Data were available for 649 pts out of 739 enrolled.
		OS	Low	High[aq]	High	High	High		
		SPM	Low	High[aq]	High	High	High		
HD12 [59,77]	RT vs. BEACOPP	PFS	Low	Some concerns	Low	Low	Low	Low for OS, Some concerns for all other outcomes.	Open label trial
		FFTF	Low	Some concerns	Low	Low	Low		
		OS	Low	Low	Low	Low	Low		
		TRD	Low	Some concerns	Low	Low	Low		
		SPM	Low	Some concerns	Low	Low	Low		
RT vs. Chemotherapy									
ECOG E1476 [78]	RT vs. ABVD	OS	Low	Low	High	High	Low	High for all outcomes	Open label trial. Missing outcome data on 14% of pts. The ascertainment of outcomes was performed later for those who were initially in partial remission and converted to Complete remission.
		PFS	Low	High	High	High	Low		
		DFS	Low	High	High	High	Low		
Aviles, 2000 [79]	Chemo alone vs. RT+ Chemo	OS	Some concerns	Low	Low	Some concerns	Some concerns	Some concerns for OS, High for FFS	Pts, clinicians, and outcome assessors were aware of the assignment to the intervention. No information was available on whether allocation sequence was concealed. No protocol was available for this study.
		FFS	Some concerns	Some concerns	Low	High	Some concerns		

A-AVD=Brentuximab vedotin plus doxorubicin, vinblastine, and dacarbazine; ABV=doxorubicin, bleomycin and vinblastine; ABVD=doxorubicin, bleomycin, vinblastine, and dacarbazine; AVD=doxorubicin, vinblastine, and dacarbazine; bBEACOPP=baseline BEACOPP; BEACOPP=bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine,

procarbazine, prednisone; $BEACOPP_{14}$=BEACOPP for 14 days; COPP=cyclophosphamide, vincristine, procarbazine, and prednisone; COPPEBVCAD-CEC=cyclophosphamide, lumustine, vindesine, melphalan, prednisone, epidoxirubicin, vincristine, procarbazine, vinblastine, and bleomycin; DFI=disease-free interval; DFS=disease-free survival; eBEACOPP=escalated BEACOPP; EFS=event-free survival; FFFP=Freedom from first progression; FFS=failure-free survival; FFTF=freedom from treatment failure; G-CSF=Granulocyte-colony stimulating factor; Gy=Gray; IF-RT=Involved field radiotherapy; ITT=Intention to treat; N-AVD=nivolumab-AVD; nr=Not reported; OS=overall survival; PET=Positron emission tomography; PFS=progression-free survival; pts=patients; Ritux=Rituximab; ROB=Risk of bias; RT=Radiotherapy; SOF=Summary of findings; SPM=Secondary primary malignancies; TRD=Treatment-related deaths.

NOTES
a The study was open label
b Lack of blinding does not lead to the same uncertainty in the evidence for OS as for other outcomes
c Deviations from the planned treatment occurred more in the BEACOPP 4+4 than in the ABVD (control) arm
d Outcome assessors were not blinded
e The study was stopped early for benefit
f Allocation was not concealed
g The intention-to-treat analysis did not include all patients randomized
h The plan for the study was different than what is reported in the publication
i Deviations from the planned interventions that arose because of the trial context: population was older in the AVD arm.
j Randomization was stratified
k The 2011 amendment affected the analysis of PET-2-negative patients; standard treatment was reduced from 8×eBEACOPP to 6×eBEACOPP.
l 10% and 13% of data were missing in the standard and PET-adapted groups
m Not all randomized patients were included in the analysis
n No description of how randomization was conducted was presented
o Not enough information was provided to answer question 2.3: Were there deviations from the intended intervention that arose because of the trial context?
p The first publication used a modified PFS (composite outcome)
q Patients were not analyzed as they were randomized

Appendix 6: Importance of outcomes for decision-making: assessment and rating

According to the GRADE methodology, outcomes are assessed at three stages during the development of a guideline: First at the project plan stage, then after reviewing the evidence included by the systematic review, and finally when judging the balance between the desirable and undesirable effects of an intervention [31].

On an initial meeting for this project, at the project plan stage, the patient representatives were asked by two PEBC methodologists (FGB and CZ) which outcomes they would rate as critical, which important, and which not important for the patient populations of concern to Questions 1, 2 and 3. The members of the working group agreed with the choices and rankings of outcomes that the patients did during this first brainstorming session.

The outcomes that are listed in the first column of Table 1-6A below are those chosen at the initial stage for this project. When reviewing the evidence from the systematic reviews it appeared evident that many different measures were used for recurrence, PFS being one of them, and that many different adverse events were reported (column 3 of the table below).

Table 1-6A. Importance of outcomes as rated at the initial brain-storming session at the Project Plan stage.

Initial list of outcomes	Initial rating of the importance of the outcomes	Measures of these outcomes reported in the studies
Question 1		
OS	CRITICAL	OS
Recurrence PFS	CRITICAL	PFS EFS Duration of response DFS FFS FFTF
Adverse events (hematological toxicity and second cancers)	CRITICAL	Treatment-related death Infections Febrile neutropenia Pulmonary toxicity Cardiovascular toxicity Peripheral neuropathy SPM
Infertility	CRITICAL	Number of pregnancies Hormones levels (males and females)
Quality of life	IMPORTANT	Quality of life
Question 2		
OS	CRITICAL	OS
Recurrence PFS	CRITICAL	PFS EFS Duration of response DFS FFS FFTF
Adverse events (hematological toxicity and second cancers)	CRITICAL	Treatment-related death Infections Febrile neutropenia

Initial list of outcomes	Initial rating of the importance of the outcomes	Measures of these outcomes reported in the studies
		Pulmonary toxicity Cardiovascular toxicity Peripheral neuropathy SPM
Infertility	NOT IMPORTANT	Not reported
Quality of life	CRITICAL	Quality of life measures (data not available)
Question 3		
OS	CRITICAL	OS
Recurrence PFS	CRITICAL	PFS EFS Duration of response DFS FFS FFTF
Adverse events (hematological toxicity and second cancers)	CRITICAL	Treatment-related death Infections Febrile neutropenia Pulmonary toxicity Cardiovascular toxicity Peripheral neuropathy SPM
Infertility	NOT IMPORTANT	Not reported
Quality of life	IMPORTANT	Quality of life

DFS=disease-free survival; EFS=event-free survival; FFS=Failure-free survival; FFTF=Freedom from treatment failure; OS=Overall survival; PFS=Progression-free survival; SPM=Secondary primary malignancies.

Members of the Working Group were asked to rank the outcomes on a 1 to 9 scale for each question, and the average ratings of two of the members with the meaning of the rating is reported in Table 2-A6 below.

Questions:
Question 1. For patients younger than 60 years of age, what is the ideal treatment strategy, among A-AVD, ABVD, ABVD-PET adapted, eBEACOPP, and eBEACOPP–PET adapted, as part of the first-line therapy, to improve patient outcomes, and how does it affect adverse events?
Question 2. For patients 60 years old and older what is the ideal treatment strategy, among ABVD, A-AVD, Brentuximab alone or in combination, PVAG, CHOP, CHLVPP, or other agents, as part of the first-line therapy to improve patient outcomes, and how does it affect adverse events?
Question 3. Does consolidation radiotherapy after first-line chemotherapy improve outcomes such as OS, PFS, recurrence, adverse events, and quality of life in patients with advanced HL?

Table 2-6A. Average rating of outcomes by two Working Group members (LM and MC) on a scale of 1-9.

Outcomes	Q 1	Q 2	Q 3
OS	9	9	9
PFS	8	8	8
DOR	6.5	6.5	6.5
EFS	7	7	6.5
DFS	7	7	6.5
FFS	7.5	7	6.5
FFTF	7.5	7.5	7
QOL	6	6.5	6.5
Death (b/c of treatment)	8.5	8.5	5.5
Infections	5.5	5.5	3.5
Feb Neutropenia	5.5	5.5	3
Pulmonary tox	6.5	6	6
Cardiac tox	5	5.5	6.5
Peripheral neuropathy	4.5	5.5	4
SPM	5.5	3.5	7.5
Impaired Fertility	4.5	1	2

b/c=Because; DFS=Disease-free survival; DOR=duration of response; EFS=event-free survival; Feb=Febrile; FFS=Failure-free survival; FFTF=Freedom from treatment failure; OS=overall survival; PFS=Progression-free survival; QOL=Quality of life; SPM=Secondary primary malignancies; tox=toxicity

Meaning of the scoring:

1 to 3 **Limited or no importance** for decision making
4 to 6 **Important but not critical** for decision making
7 to 9 **Critical** for decision making

Additionally, a search of the literature for patients' preference for outcomes in patients treated with first-line therapy for advanced-stage HL was undertaken. We searched PubMed with the terms: "*Hodgkin disease AND *Patient preference AND first-line". Ten records were identified of which two were relevant [32,33]. These were two cross-sectional surveys with a discrete choice experiment offering patients trade-offs between outcomes. Khan et al. [32] was based in the US, and only included patients' preferences, while Bröckelmann et al. [33] interviewed both patients and physicians in France, Germany, and UK. Although these studies focused on patients with advanced-stage HL, they also included patients with early-stage disease. Median age of the included patients in these two surveys was respectively 35.0 years (range 19.0 to 69.0 years), and 36 years (19 to 74 years), but data are reported also on the older patients' subgroup. A summary of patients', and clinicians' preferences reported by these two studies is below:

- Patients, particularly the advanced-stage subgroup, rated survival outcomes as more important than drug-related toxicity.
- PFS had the highest importance to younger patients, followed by OS, pulmonary toxicity, and peripheral neuropathy.
- Older patients put a higher value on OS than on PFS and infertility was seen as not important; these patients put a greater importance in pulmonary toxicity than younger ones.
- Physicians preferred OS to PFS (except for older patients for whom this was reversed), followed by decreased risk of short-term side effects, peripheral neuropathy, infertility, and pulmonary toxicity. Their rating varied depending on the patient profile, e.g., side effects was considered more important for older patients, and pulmonary toxicity more important if the hypothetical patient was a smoker.

The Working Group **discussed the** results of the survey and of the data on patient preferences retrieved from the literature during the March 10, 2023, meeting. Considering that seven is the maximum number of outcomes recommended by the GRADE group [34], we re-classified the outcomes as follows:

Critical outcomes:
Q1, Q2, Q3: OS. Recurrence (as measured by PFS, and other measures such as EFS, DFS, FFS, FFTF). Treatment-related death.
Q2: Quality of life.

Important outcomes:
Q1, Q2: Pulmonary toxicity
Q1: SPM, Infertility
Q3: SPM.

Table 2-6B. Critical and important outcomes

Question	Critical outcomes	Important outcomes	Outcomes that are not important
Question 1	OS Recurrence[a] Treatment-related death	Pulmonary toxicity SPM Infertility	Quality of life Other adverse events
Question 2	OS Recurrence[a] Treatment-related death Quality of life	Pulmonary toxicity	Infertility Other adverse events
Question 3	OS Recurrence[a] Treatment-related death	SPM	Other adverse events

[a]As measured by PFS, and other measures such as for instance EFS, DFS, FFS, FFTF
OS=overall survival; SPM=second primary malignancies

We showed this re-classification to the patient representatives who participated to the initial brainstorming session, and we received the following comment: “I think if the adverse events are recorded and well explained that is the important point and there is a mention that these will be shared with the patient, I am comfortable with that approach and I don’t need to re-review.”

Appendix 7: Results of included trials

Table 1-7A. Results of studies included for Question 1 (patients in the younger age group)

Study name, Reference, Treatment comparisons	OS	Recurrence	AE (Includes: Treatment-related death, SPM, Pulmonary toxicity, infertility)
Comparison 1. ABVD vs. BEACOPP			
EORTC 20012 Intergroup Trial Carde, 2016 [52] 8×ABVD vs. 4×eBEACOPP+4×bBEACOPP.	8×ABVD vs. 4e+4bxBEACOPP at 4 yrs (48 mos): OS: 86.7% vs. 90.3% (HR, 0.71; 95% CI, 0.42 to 1.21; p=0.208	8×ABVD vs. 4×e+4×bBEACOPP at 4 yrs (48 mos): PFS*: 72.8% vs. 83.4% (HR, 0.58; 95% CI, 0.39 to 0.85; p=0.005 EFS* 63.7% vs. 69.3% (HR, 0.86; 95% CI, 0.64 to 1.15; p=0.313) The events that defined EFS were: premature discontinuation (10.5% vs. 13.9%), no CR or CRu at cycle 8 (9.1% vs. 4.7%), progression or relapse (12.7% vs. 7.7%), and death (3.3% v 3.3%), respectively. DFS (only pts who reached CR or CRu, 83% of the total sample) 85.8% vs. 91.0% (HR, 0.59; 95% CI, 0.33 to 1.06; p=0.076)	8×ABVD vs. 4e+4bxBEACOPP at 4 yrs (48 mos): Treatment toxicity-related death (TRD): A) Deaths due to toxicity: 9 (3.2%) vs. 6 (2.2%) Deaths due to bleomycin-induced pulmonary toxicity: 6 (2.2%) vs. 0 (0%) p values nr Deaths due to septic shock or severe intercurrent infection: 4 (1.5%) vs. 8 (2.9%) Deaths due to liver failure: 1 (0.36%) vs. 2 (0.73%) Deaths due to cardiomyopathy 2 (0.73%) vs. 1 (0.36%) Deaths from SPM: 2 (0.73%) vs. 4 (1.5%); p values not provided for the above B) Late effects of treatment: SPM: 8 (2.9%) vs. 10 (3.6%) Cumulative incidence of SPM: 3.4% vs. 4.7%, p=NS
LYSA H34 Mounier, 2014 [53] 8×ABVD vs. 4×eBEACOPP+4×bBEACOPP	OS estimated at 5 yrs (60 mos): 92% (95% CI, 81-97) vs. 99% (95% CI, 90-100), HR 0.18 (95% CI, 0.02- 1.53), p=0.06.	8×ABVD vs. 4e+4bxBEACOPP at 5 yrs (60 mos) PFS 75% (95% CI, 63-83) vs. 93% (95% CI, 83-97), HR=0.3 (95% CI, 0.12-0.77), p=0.007. EFS* 62% (95% CI, 50-71) vs. 77% (95% CI, 65-85), HR=0.6 (95% CI, 0.33-1.06), p=0.07	Death because of SPM: 2 (2.5%) vs. 1 (1.4%) p value nr Death for any cause 6 (7.5%) vs. 1 (1.4%) p value nr Grade 3-4 toxicity: Pulmonary: 1% vs. 1%
HD9 Diehl, 2003 [54] 5-yr follow up; Diehl, 1998 [64] 1st and 2nd Interim analyses	5-*yrs follow-up* OS rate: bBEACOPP vs. COPP/ABVD: 87% vs. 81.2%, p=0.16 eBEACOPP vs. COPP/ABVD:	PFS *1st Interim analysis (at 2 yrs – 23 mos)* COPP/ABVD vs. bBEACOPP+eBEACOPP FFTF* at 2 yrs 23 mos: 70% vs. 83% Test for a difference p=0.0007	*5-yrs follow-up* [54]: Treatment-related death: <2% in all groups p=NS: COPP/ABVD: 5 of 260 pts (1.9%) bBEACOPP: 7 of 469 pts (1.5%) eBEACOPP: 8 of 466 pts (1.7%)

Study name, Reference, Treatment comparisons	OS	Recurrence	AE (Includes: Treatment-related death, SPM, Pulmonary toxicity, infertility)
Engert, 2009 [81] 10-yr follow-up; Von Treskow, 2018 [77], 15-yr follow-up [120] Erratum 8×COPP/ABVD vs. 8×bBEACOPP vs. 8×eBEACOPP (G-CSF)	91% vs. 81.2%, p=0.002 *10 yrs follow-up* 2009 *[81]:* OS rate: COPP/ABVD vs. bBEACOPP vs. eBEACOPP: 75% vs. 80% vs. 86%, p=0.0005 COPP/ABVD vs. bBEACOPP: p=0.19 COPP/ABVD vs. eBEACOPP: p<0.001 *15 yrs follow-up:* OS COPP/ABVD vs. bBEACOPP vs. eBEACOPP 72.3% (95% CI, 66.5–78.1) vs. 74.5% (95% CI, 70.1–78.9) vs. 80.9% (95% CI, 76.7–85.0) The difference in OS between eBEACOPP and COPP/ABVD: 8.6% (95% CI,1.4–15.7), HR 0.68 (95% CI, 0.50–0.93), p=0.015.	*5-yrs follow-up [54]* Progression rate: COPP-ABVD: 10% bBEACOPP: 8% eBEACOPP: 2% Difference between COPP-ABVD and eBEACOPP: p<0.001 Difference between COPP-ABVD and bBEACOPP; p=NS (value *nr*) *5-yrs follow up:* FFTF* rate: bBEACOPP vs. COPP/ABVD: 76% (95% CI, 72%-80%) vs. 69% (95% CI, 63%-75%), p=0.04 eBEACOPP vs. COPP/ABVD and bBEACOPP: 87% (95% CI, 83%-91%), p<0.001. Subgroups for FFTF: Pts aged 60-65 yrs old (n=64): p=0.93 Note: More data (larger sample) for this subgroup are in the companion study by Ballova et al. [63]. *10 yrs follow-up* FFTF* rate: COPP/ABVD vs. bBEACOPP vs. eBEACOPP: 64% vs. 70% vs. 82%, p<0.0001 COPP/ABVD vs. bBEACOPP: p=0.04 *15 yrs follow-up:* PFS: COPP/ABVD vs. bBEACOPP vs. eBEACOPP 57.0% (95% CI, 50.0–64.0) vs. 66.8% (95% CI, 61.9–71.8) vs. 74.0% (95% CI, 69.0–79.0) The difference in PFS between eBEACOPP and COPP/ABVD: 17.0% (95% CI ,8.3–25.6), HR 0.53 (95% CI, 0.41–0.69), p<0.0001.	SPM: AML: 0.4% vs. 0.6% vs. 2.5%, p=0.03 Solid tumours: 1.2% vs. 1.7% vs. 0.43% p values nr Non-HL: 2.7% vs. 0.85% vs. 1.1% Pulmonary toxicity: Respiratory tract effects Grade ≥3: 2% vs. 5% vs. 4%, p nr Infertility nr *10 yrs follow-up* Cumulative incidence rate of AML/MDS: 0.4% vs. 2.2% vs. 3.2%, p=0.03 Cumulative incidence rate of all SPM: 5.3% vs. 7.9% vs. 6.5%, p=0.82 Death due to toxicity from salvage: 1.9% vs. 1.5% vs. 0.6% p values nr Death due to acute toxicity of first-line treatment, secondary malignancies, and other factors: no difference between treatment arms. *15 yrs follow- up:* SPM Cumulative incidence of second cancers: 15 yrs cumulative incidence of secondary acute leukemia or MDS: incidence in the entire group: 117 (10%) COPP/ABVD vs. eBEACOPP: 0.4% (95% CI: 0.0-1.1) vs. 4.0% (95% CI: 1.9-6.1), p=0.0018 15 yrs of all other secondary neoplasms: COPP/ABVD vs. bBEACOPP vs. eBEACOPP: 7.2% (95% CI: 3.7 -10.7) vs. 13.0% (95% CI: 9.1-16.9) vs. 11.4% (95% CI: 7.6-15.1); p=0.41 for COPP/ABVD vs. bBEACOPP, p=0.53 for COPP/ABVD vs. eBEACOPP.
HD2000 Federico, 2009 [55]; Merli, 2016 [56] (10-yr follow-up post-hoc analysis) 6×ABVD vs. 4eBEACOPP + 2bBEACOPP vs. 6× COPPEBVCAD-CEC	Estimated at 5-yr [55]: OS rate at 41 mos [55]: ABVD vs. BEACOPP vs. COPPEBVCAD-CEC 84% vs. 92% vs. 91% eBEACOPP vs. ABVD, p=0.893 COPPEBVCAD-CEC vs. ABVD, p=0.821 OS rate at 120 mos [56]:	Estimated at 5-yr [55]: FFS* rate 65.7% vs. 79.6% vs. 72.4% eBEACOPP vs. ABVD, p=0.036 COPPEBVCAD-CEC vs. ABVD, p=0.423 PFS rate 68.7% vs. 82.7% vs. 79.6%	Estimated at 5-yr [55]: SPM 1% vs. 1% vs. 2% p values nr At 10-yrs follow-up [56] Treatment-related death First-line: 0 vs. 2% vs. 0 Salvage: 2% vs. 3% vs. 2%

Study name, Reference, Treatment comparisons	OS	Recurrence	AE (Includes: Treatment-related death, SPM, Pulmonary toxicity, infertility)
	85% (95% CI, 75-91) vs. 84% (95% CI, 74-90) vs. 86% (95% CI, 77-92), p=0.892;	eBEACOPP vs. ABVD, p=0.038 COPPEBVCAD-CEC vs. ABVD, p=0.247 Subgroups: IPS 0-2 eBEACOPP vs. ABVD, p=0.125 COPPEBVCAD-CEC vs. ABVD, p=0.676 IPS 3-7 eBEACOPP vs. ABVD, p=0.038 COPPEBVCAD-CEC vs. ABVD, p=0.056 RFS rate 76.8% vs. 87.8% vs. 88.8% eBEACOPP vs. ABVD, p=0.101 COPPEBVCAD-CEC vs. ABVD, p=0.190 At 10-yrs follow-up [56] PFS rate (regression analysis with ABVD as reference): 69% (95% CI, 58-77) vs. 75% (95% CI, 64-83) vs. 76% (95% CI, 66-84) FFS rate: 65% (95% CI, 54-73), 73% (95% CI, 62-81), and 71% (95% CI, 61-79)	SPM rate: 0.88% vs. 6.7% vs. 6.45% Cumulative incidence SPM rate: 0.9% (95% CI, 0.1-4.5), 6.6% (95% CI, 2.4-13.8), and 6% (95% CI, 1.8-13.9), p=0.02
GITIL Viviani, 2011 [57] 6 or 8ABVD vs. 4eBEACOPP+4bBEACOPP	OS at 7-yr (61 mos) follow-up 84% (95% CI, 77-91) vs. 89% (95% CI, 84-95), HR 0.75 (95% CI,0.39-1.45). p=0.39	7-yr (61 mos) follow-up: FFFP* rate in the ITT population: 73% (95% CI, 66-80) vs. 85% (95% CI, 78-91), HR 0.46, (95% CI, 0.27-0.78), p=0.004 EFS rate: 71% (95% CI, 64-78) vs. 78% (95% CI, 70-85), HR 0.72, (95% CI, 0.46-1.13), p=0.15	≥1 episode of severe toxic effect (hematologic): 43% vs. 81%, p<0.001 or (nonhematologic): 7% vs. 19%, p=0.001 Treatment-related death: 1/166 (1%) vs. 5/156 (3%), p values nr SPM Leukemia during follow-up: 1 (1%) vs. 2 (1%) Pulmonary toxicity: nr Infertility: nr
Comparison 2. Higher-intensity (i.e., higher dose, more cycles) vs. Lower-intensity (lower dose, less cycles) eBEACOPP			
HD15 Engert, 2012 [58], Engert, 2017 [82] 8×eBEACOPP vs. 6×eBEACOPP vs. 8×$BEACOPP_{14}$	Survival at 48 mos: 8×eBEACOPP vs. 6×eBEACOPP vs. 8×$BEACOPP_{14}$ OS rates (ITT analysis): 91.9% (97.5% CI, 89.4-94.4) vs 95.3% (97.5% CI, 93.4-97.2) vs. 94.5% (97.5% CI, 92.5-96.6)	Recurrence at 48 mos follow-up [58]: 8×eBEACOPP vs. 6×eBEACOPP vs. 8×$BEACOPP_{14}$ FFTF* rates (ITT analysis): 84.4% (97.5% CI, 81.0-87.7) vs. 89.3% (97.5% CI, 86.5-92.1) vs. 85.4% (97.5% CI, 82.1-88.7)	Adverse events at 48 mos [58]: Treatment-related death: 2.1% vs. 0.8% vs. 0.8% SPM: 4.7% vs. 2.4% vs. 3.1% Adverse events at 102 mos follow-up [82] Treatment-related death rate: 2% vs. 1% vs. 1%

Study name, Reference, Treatment comparisons	OS	Recurrence	AE (Includes: Treatment-related death, SPM, Pulmonary toxicity, infertility)
	6×eBEACOPP vs. 8×eBEACOPP: HR 0.60 (97.5% CI, 0.39-0.92), p=0.019. *From Figure 2* 8×BEACOPP$_{14}$ vs. 8×eBEACOPP: HR 0.68 (97.5% CI, 0.42-1.10), p=0.07 6×eBEACOPP vs. 8×BEACOPP$_{14}$ HR 0.88 (97.5% CI, 0.51-1.51), p=0.6 OS rates (per-protocol analysis) 91.8% (97.5% CI, 89.0-94.6) vs. 96.2% (97.5% CI, 94.4-98.0) vs. 94.8% (97.5% CI, 92.6-97.0) Difference 6×eBEACOPP from 8×eBEACOPP: 4.4 (97.5% CI, 1.1-7.7) Difference 8×BEACOPP$_{14}$ from 8×eBEACOPP: 3.0 (97.5% CI, -0.5-6.6) Difference 8×BEACOPP$_{14}$ from 6×eBEACOPP: -1.4 (97.5% CI, -4.2-1.5) Survival at 112 mos follow-up [82] OS rates (ITT analysis) 87.6% (97.5% CI, 84.6-90.6) vs 90.4% (97.5% CI, 87.5-93.2) vs 91.6% (97.5% CI, 89.1-94.1) 6×eBEACOPP vs. 8×eBEACOPP: HR 0.7 (97.5% CI, 0.5-0.999), p=0.0245 8×BEACOPP$_{14}$ vs. 8×eBEACOPP: HR 0.7 (97.5% CI, 0.5-1.04) 8×BEACOPP$_{14}$ vs. 6×eBEACOPP: HR 1.0 (97.5% CI, 0.7-1.6) OS rates (per-protocol analysis) 87.2% (97.5% CI, 83.7-90.7) vs. 91.9% (97.5% CI, 89.0-94.8) vs 91.9% (97.5% CI, 89.2-94.6) 6×eBEACOPP vs. 8×eBEACOPP: HR 0.6 (97.5% CI, 0.4-0.9) 8×BEACOPP$_{14}$ vs. 8×eBEACOPP: HR 0.7 (97.5% CI, 0.5-1.1), p=0.05 8×BEACOPP$_{14}$ vs. 6×eBEACOPP: HR 1.3 (97.5% CI, 0.8-2.0)	6×eBEACOPP vs. 8×eBEACOPP: HR 0.67 (97.5% CI, 0.47-0.95), p=0.009 8×BEACOPP$_{14}$ vs. 8×eBEACOPP: HR 0.92 (97.5% CI, 0.67-1.26), p=0.5 6×eBEACOPP vs. 8×BEACOPP$_{14}$ HR 0.73 (97.5% CI, 0.51-1.03), p=0.042 6×eBEACOPP and 8×BEACOPP$_{14}$ were non-inferior when compared with 8×eBEACOPP as their repeated 97.5% CI for the HR excluded the noninferiority margin of 1.51 (6×eBEACOPP vs. 8×eBEACOPP 0.44 to 1.02; 8×BEACOPP$_{14}$ vs. 8×eBEACOPP: 0.62 to 1.36). FFTF rates (per-protocol analysis): 84.9% (97.5% CI, 81.2-88.6) vs. 90.8% (97.5% CI, 88.1-93.6) vs. 86% (97.5% CI, 82.4-89.6) Difference 8×BEACOPP$_{14}$ from 8×eBEACOPP: 1.1% (97.5% CI, -4.1-6.3) Difference 8×BEACOPP$_{14}$ from 6×eBEACOPP: -4.8% (97.5% CI, -9.4- —0.3) PFS rate (ITT analysis) at 48 mos: 85.6% (97.5% CI, 82.3-88.9) vs. 90.3% (97.5% CI, 87.6-93.0) vs. 85.8% (97.5% CI, 82.4-89.2) Time to progression (ITT analysis) 6×eBEACOPP vs. 8×eBEACOPP: HR 0.74 (97.5% CI, 0.48-1.13), p=0.11 8×BEACOPP$_{14}$ vs. 8×eBEACOPP: HR 1.15 (97.5% CI, 0.78-1.68), p=0.4 6×eBEACOPP vs. 8×BEACOPP$_{14}$ HR 0.64 (97.5% CI, 0.43-0.97), p=0.016 PFS rate (per-protocol analysis): 85.3% (97.5% CI, 81.7-89.0) vs. 91.1% (97.5% CI, 88.4-93.8) vs. 86.4% (97.5% CI, 82.8-90.0) Difference 6×eBEACOPP from 8×eBEACOPP: HR 5.8 (97.5% CI, 1.2-10.4) Difference 8×BEACOPP$_{14}$ from 8×eBEACOPP HR 1.1 (97.5% CI, -4.1-6.2) Difference 8×BEACOPP$_{14}$ from 6×eBEACOPP: HR -4.7 (97.5% CI, -9.2- —0.2) Recurrence at 102 mos follow-up [82]	SPM rate: 8% vs. 6% vs. 6% Cumulative incidence of SPM 10% (97.5% CI, 7%–13%), 7% (97.5% CI, 4%–9%), and 7% (97.5% CI, 4%–10%) 6×eBEACOPP vs. 8×eBEACOPP: HR 0.7 (97.5% CI,0.4-1.1), p=0.1 8×BEACOPP$_{14}$ vs. 8×eBEACOPP: HR 0.7 (97.5% CI, 0.4-1.1), p=0.1

Study name, Reference, Treatment comparisons	OS	Recurrence	AE (Includes: Treatment-related death, SPM, Pulmonary toxicity, infertility)
		PFS rates (ITT analysis): 8×eBEACOPP 80.9% (97.5% CI, 77.3-84.6) 6×eBEACOPP 83.7% (97.5% CI, 79.9-87.4) 8×BEACOPP$_{14}$ 83.6% (97.5% CI,80.2-87.0) 6×eBEACOPP vs. 8×eBEACOPP: HR 0.7 (97.5% CI, 0.5-1.01) 8×BEACOPP$_{14}$ vs. 8×eBEACOPP: HR 0.9 (97.5% CI, 0.7-1.2) 8×BEACOPP$_{14}$ vs. 6×eBEACOPP: HR 1.2 (97.5% CI, 0.9-1.6) Both 6×eBEACOPP and 8×BEACOPP$_{14}$ were noninferior to 8×eBEACOPP and both excluded the non-inferiority margin of 1.51. PFS (per-protocol analysis, at 103 mos): 8×eBEACOPP 79.9% (97.5% CI, 75.7-84.1) 6×eBEACOPP 85.1% (97.5% CI, 81.1-89.0) 8×BEACOPP$_{14}$ 84.1% (97.5% CI, 80.5-87.8) 6×eBEACOPP vs. 8×eBEACOPP: HR 0.7 (97.5% CI, 0.5-0.9) 8×BEACOPP$_{14}$ vs. 8×eBEACOPP: HR 0.8 (97.5% CI, 0.6-1.2) 8×BEACOPP$_{14}$ vs. 6×eBEACOPP: HR 1.3 (97.5% CI, 0.9-1.8)	
HD12 Borchmann, 2011 [59]; von Tresckow, 2018 [77] Chemotherapy comparison: 8×eBEACOPP ± RT (n=787) vs. 4×eBEACOPP + 4×bBEACOPP± RT (n=787) *NOTE: This study also reported a radiotherapy comparison, and it is reported in Question 3.*	*At 5 yrs (78 mos):* OS rate 92% vs. 90.3% (difference, −1.7%, (95% CI, −4.6%-1.1%) HR, 1.14 (95% CI, 0.83-1.56) Estimated 10 yrs rates (97 mos follow-up) [77]: OS 8×eBEACOPP vs. 4×eBEACOPP+4×bBEACOPP 87.3% (95% CI, 84.7-89.9) vs. 86.8% (95% CI, 84.2-89.4) HR, 1.02 (95% CI, 0.77-1.36).	*At 5 yrs (78 mos):* FFTF* rate 86.4% vs. 84.8%, (difference −1.6, (95% CI, −5.2% - −1.9%) HR, 1.07 (95% CI, 0.83-1.38) PFS 87.5% vs. 85% (difference, −2.5%, (95% CI, −6% – 1%) HR, 1.16 (95% CI, 0.89 - 1.50) *At 8 yrs follow-up (97 mos) [77]:* PFS (Reduced intensity chemotherapy was not inferior to standard within the non-inferiority margin of 1.50) HR, 1.13 (95% CI, 0.89 – 1.43) Estimated 10 yrs rates: 8×eBEACOPP vs. 4+4 82.6% (95% CI, 79.6-85.6) vs. 80.6 (95% CI, 77.4-83.7)	Chemotherapy comparison at 78 mos: 8×eBEACOPP ± RT vs. 4×eBEACOPP +4×bBEACOPP ± RT Treatment-related death: 19 (2.4%) vs. 27 (3.4%) p=NS, too small number of events to detect a difference SPM: 43 (5.5%) vs. 33 (4.2%) p=NS Estimated 10 yrs rates (78 mos follow- up) [77]: Treatment-related death: 8×eBEACOPP vs. 4+4 8 (2%) vs. 11 (3%) vs. 17 (4%) vs. 10 (3%) SPM: 36 (9%) vs. 25 (6%) vs. 23 (6%) vs. 24 (6%) Estimated 10 yr cumulative incidence: 4 + 4 vs. eBEACOPP: 6.4% (95% CI 3.3–9.5) vs. 8.8% (5.2–12.4)

Study name, Reference, Treatment comparisons	OS	Recurrence	AE (Includes: Treatment-related death, SPM, Pulmonary toxicity, infertility)
HD9 Objective 2: bBEACOPP vs. eBEACOPP Diehl, 1998 [64]; Interim analysis. Diehl, 2003 [54]; Engert, 2009 [81]; [120] Erratum; Von Treskow, 2018 [77]	At 72-mos follow-up [54]: bBEACOPP vs. eBEACOPP: 88% vs. 91%, p=0.06. At 111-mos follow-up [81]: bBEACOPP vs. eBEACOPP: 80% vs. 86%, p=0.0053	*5-yrs follow-up [54]* Progression rate: bBEACOPP: 8% eBEACOPP: 2% Difference between bBEACOPP and eBEACOPP: p=0.001	See Comparison 1.
Comparison 3. PET-adapted strategies			
HD18 Borchmann, 2018 [16] (final analysis) Borchmann, 2017 [50] (PET-2 positive pts) Kreissl, 2021 [83] (5-yr follow-up) 2×eBEACOPP then PET PET-2 positive pts: n=217 8×eBEACOPP (Standard) + vs. n=217 8×R-eBEACOPP PET-2 negative pts: (n=244) 8×eBEACOPP (Standard) + 6×eBEACOPP (n=202 New Standard)[c] vs. (n=474) 4×eBEACOPP	OS *PET-2 positive pts* 2nd interim analysis at 36 mos [50]: 8×eBEACOPP vs. 8×ReBEACOPP OS rates: 96.5% (95% CI, 93.6-99.3) vs. 94.4% (95% CI, 90.8-97.9) p=0.31 Estimated OS at 66 mos [16]: Pts assigned before the amendment: 96.4% (95% CI, 93.8-99.0) vs. 93.0% (95% CI, 90.6-97.3), HR 1.62 (95% CI, 0.70-3.75), p=0.25 OS at 78 mos [83], pts assigned before the amendment: 96.5% (95% CI, 94.0-99.1) vs. 93.4% (95% CI, 90.0-96.9), p=0.15 *PET negative pts* Per-protocol analysis (b/c noninferiority) 8×eBEACOPP or 6×eBEACOPP vs. 4×eBEACOPP at 55 mos [16]: 95.4% (95% CI, 93.4-97.4) vs. 97.7% (95% CI, 96.2-99.3), HR 0.32 (95% CI, 0.14-0.72), p=0.0037 8×eBEACOPP or 6×eBEACOPP vs. 4×eBEACOPP at 66 mos [83] 95.3% (95% CI, 93.3-97.3) vs. 98.1% (95% CI, 96.8-99.4), HR 0.36 (95% CI, 0.17-0.74), p=0.0038	PFS* PET-2 positive pts 2nd interim analysis at 3 yrs [50] 8×eBEACOPP vs. 8×ReBEACOPP PFS* rates: 91.4% (95% CI, 87.0-95.7) vs. 93.0% (95% CI, 89.4-96.6) p=0.99 Estimated PFS at 66 mos [16]: Pts assigned before the amendment: 89.7% (95% CI, 85.4-94.0) vs. 88.1% (95% CI, 83.5-92.7), HR 1.25 (95% CI, 0.69-2.26), p=0.46 PFS at 78 mos [83], pts assigned before the amendment: 89.9% (95% CI, 85.7-94.1) vs. 87.7% (95% CI, 83.1-92.4), p=0.40 PET negative pts Per-protocol analysis 8×eBEACOPP or 6×eBEACOPP vs. 4×eBEACOPP at 55 mos [16]: 90.8% (95% CI, 87.9-93.7) vs. 92.2% (95% CI, 89.4-95.0). Difference: −1.4 (95% CI, −2.7%-5.4%) which excluded the predefined non-inferiority margin of -6%. HR 0.79 (95% CI, 0.50-1.24) 8×eBEACOPP or 6×eBEACOPP vs. 4×eBEACOPP at 66 mos [83] 91.2% (95% CI, 88.4-93.9) vs. 93% (95% CI, 90.6-95.4). Difference 1.9% (95% CI, −1.8-5.5), HR 0.70 (95% CI, 0.45-1.08) *Note to*	AE PET-2 positive pts 2nd interim analysis at 36 mos [50] 8×eBEACOPP vs. 8×ReBEACOPP Treatment-related death: <1% vs. 1%, p values nr Grade 4 respiratory tract toxicity: 1% vs. 2%, p values nr SPM at 5 yrs: 66 mos cumulative incidence of SPM [83], pts assigned before the amendment: 4.0% (95% CI: 1.3-6.7) vs. 3.3% (95% CI: 0.7-5.9), p=0.44 PET negative pts 8×eBEACOPP or 6×eBEACOPP vs. 4×eBEACOPP at 78 mos [83] Treatment-related death: 1% vs. 0, p values nr Cumulative incidence of SPM: 3.7% (95% CI, 2.0-5.4) vs. 3.3% (95% CI, 1.6-5.0), HR 0.87 (95% CI, 0.46-1.63), p=0.66 Lung function (mean (SD) diffusion capacity of the lung at 5 yr: 8×eBEACOPP vs. 6×eBEACOPP vs. 4×eBEACOPP Data are reported on a very small number of participants (33 female and 37 male, no comparison is reported) Fertility: pregnancies are reported, and data are not complete.
AHL2011	Per protocol analysis: OS at 50.4 mos:	PFS: Standard treatment vs. PET-driven treatment	Death from treatment-related AE at 50.4 mos: 6 (1%) vs. 2 (<1%)

Study name, Reference, Treatment comparisons	OS	Recurrence	AE (Includes: Treatment-related death, SPM, Pulmonary toxicity, infertility)
Casasnovas, 2019 [17] Casasnovas, 2022 [84] (for updated late treatment -related AE such as SPM, and fertility) 6×eBEACOPP vs. PET-driven treatment: All pts received 2×eBEACOPP, then PET2. Standard group: 2×eBEACOPP (irrespective of PET results) then PET4. If PET-4-negative: 2×eBEACOPP; if PET-4 positive: salvage treatment. PET-driven treatment: PET2-positive pts: 2×eBEACOPP PET2-negative pts: 2×ABVD Then PET-4: Consolidation If PET-4 negative: 2×ABVD If PET-4 positive: salvage treatment	95.6% (95% CI: 91.2-97.8) vs. 95.9% (95% CI: 92.5-97.8), HR 1.248 (95% CI: 0.53-2.88, p=0.69) OS at 67.25 mos: 93.6% vs. 97.1% (HR 1.20 (95% CI:0.55- 2.60, p=0.69) ITT analysis: OS at 50.4 mos: 95.2% (95% CI: 91.1-97.4) vs. 96.4% (95% CI: 93.3-98.1), HR 0.936 (95% CI: 0.427-2.051), p=0.43 OS at 67.25 mos: 97.7% vs. 97.7%, HR 1.012 (95% CI: 0.50-2.10, p=0.53) Subgroups: PET-2-negative at 67.25 mos: Standard vs. PET-driven treatment 97.6% (95% CI:95.2-98.8) vs. 97.8% (95% CI:95.5-99.0)	Per-protocol analysis: *PFS at 50.4 mos: 86.7% (95% CI: 81.9-90.3) vs. 85.4% (95% CI: 80.7-89.0); HR 1.144 (95% CI: 0.75 -1.72, p=0.74 PFS at 67.25 mos: 88.1% vs. 86.5% (HR 1.01 (95% CI: 0.75-1.66, p=0.75) ITT analysis: *PFS at 50.4 mos: 86.2% (95% CI: 81.6-89) vs. 85.7% (95% CI: 81.4-89.1), HR 1.04 (95% CI: 0.73-1.59, p=0.65) *PFS at 67.25 mos: 87.5% vs. 86.7% (HR 1.07 (95% CI: 0.74-1.57, p=0.67) PET-2-negative at 67.25 mos: Standard vs. PET-driven treatment 89.9% (95% CI: 86.2-92.7) vs. 90.5% (95% CI:86.9-93.2) Subgroups at 50.4 mos: PET-2-positive vs. PET-2-negative: 70.7% (95% CI: 60.7-78.6) vs. 88.9% (95% CI: 85.7-91.4), HR 3.59 (95% CI: 2.32-5.56), p<0.0001 PET-4-positive vs. PET-4-negative at 67.25 mos: 46.5% (95% CI: 31.2-60.4) vs. 89.6% (95% CI: 86.5-92.0), HR 10.9 (95% CI: 6.75-17.71), p<0.0001	SPM: At 67.25 mos: 13 (3.2%) vs. 9 (2.2%) **Infertility:** Pregnancies: At 50.4 mos: 28 (7%, 95% CI: 5-10) vs. 45 (11%, 95% CI: 8-15), p=0.036 At 67.25 mos: 35 (8.5%) vs. 51 (12.5%) at 5.6 yrs From Demeestere, 2021 [90]: Hormone status favoured the PET-driven strategy: Premature ovarian insufficiency (FSH >24 IU/L): OR, 0.20 (95% CI, 0.08-0.50), p=0.001, adjusted for age: (aOR, 0.09, 95% CI, 0.03-0.32; p<0.001). Low ovarian reserve (anti-müllerian hormone, <0.5 ng/mL) OR, 0.15; 95% CI, 0.04 to 0.56, p=0.005 Ovarian function recovery: HR 2.52, (95% CI, 1.73-3.67), p<0.0001.
RATHL Johnson, 2016 [49,85] PET negative after 2×ABVD: 4×ABVD vs. 4×AVD PET positive after 2×ABVD: BEACOPP14 vs. eBEACOPP (not randomized, and not reported here)	PET-2 negative pts: OS rate at 36 mos: ABVD vs. AVD 97.2% (95% CI, 95.1-98.4) with ABVD and 97.6% (95% CI, 95.6-98.7); HR 0.90, (95% CI, 0.47-1.74), p=0.76 At 87.2 mos (IQR 63-104) 7.3 yrs [85] ABVD vs. AVD: OS: 93.2% (95% CI, 90.2 - 95.3) vs. 93.5% (95% CI, 90.5 - 95.5), HR, 0.84 (95% CI, 0.51 - 1.37).	PET-2 negative pts: PFS rate at 3 yrs: ABVD vs. AVD ITT analysis: 85.7% (95% CI, 82.1-88.6) vs. 84.4% (95% CI, 80.7-87.5) with a difference (ABVD minus AVD) at 3 yrs of 1.6% (95% CI, −3.2-5.3); HR 1.13 (95% CI, 0.81-1.57), p=0.48. The noninferiority upper boundary margin of the 95% CI was initially 5. Per-protocol analysis: difference: 1.3% (95% CI, −3.7-5.1), HR 1.10 (95% CI, 0.79-1.53), p=0.58.	PET-2 negative pts: Diffusing capacity of the lung for carbon monoxide: ABVD vs. AVD: Absolute difference from baseline to the completion of therapy: −7.4% (95% CI, 5.1-9.7; p<0.001). At 1 yr: −4.6% (95% CI, 1.6-7.5; p=0.003). TRD (due to initial therapy): 4 pts (0.85%) vs. 0 (0%)

Study name, Reference, Treatment comparisons	OS	Recurrence	AE (Includes: Treatment-related death, SPM, Pulmonary toxicity, infertility)
		At 87.2 mos (IQR 63-104) 7.3 yrs [85] ABVD vs. AVD PFS: 81% (95% CI, 76.9 - 84.4) vs. 79.2% (95% CI, 75.1 - 82.8); The 1.3% difference in 3-year PFS (95% CI, –3.0 to 4.7) between ABVD and AVD now falls within the predefined noninferiority margin (exclude a difference of 5 or more).	Infertility: See companion trial by Anderson et al. 2018 [93] At 7 yrs follow- up [85]: SPM: 52 in the entire population of 1201 pts. Cumulative incidence of SPM at 7 yrs: ABVD/AVD vs. BEACOPP ABVD: 5.1% (95% CI, 3.2 - 8.1), AVD 5.8% (95% CI, 3.8 to 9.0), and 2.5% (95% CI, 0.8 to 7.7) for eBEACOPP.
GITIL/FIL HD0607 Gallamini, 2018 [48] (Long-term results) PET-2-positive pts only: Ritux (375 mg/m^2) + 4×eBEACOPP+4×bBEACOPP vs. 4×eBEACOPP+4×bBEACOPP	OS at 3 yrs PET-2-positive pts: OS rate: 89% (95% CI, 79-95) vs. 90% (95% CI, 78-95), p=0.895	PFS rate at 3 yrs* PET-2-positive pts: 63% (95% CI, 50-74) vs. 57% (95% CI, 45-68), p=0.534 *Note: PET-negative pts are reported in Question 3* Subgroups: DS score 4 vs. DS 5 73% (95% CI, 62-81) vs. 35% (95% CI, 22-49), p<0.001	AE Not reported by chemotherapy treatment group.
Comparison 4. ABVD vs. modified ABVD			
ECHELON-1 Connors, 2018 (2-yr interim analysis) [9]; Straus, 2020 (3-yr) [86]; Radford, 2020 (abs, 4-yr)[87]; Straus, 2021 (5-yr) [88]; Ansell, 2022 (6-yr) [89] ABVD vs. A+AVD.	OS** rate 24.6 mos [9] (pre-planned interim analysis); Events: 39 vs. 28, HR, 0.728 (95% CI, 0.448-1.184), p=0.199 94.2% (95% CI, 92.0-95.9) vs. 96.6% (95% CI, 94.8-97.7) HR 0.73 (95%CI, 0.45-1.18), p=0.20 37 mos; [86]; OS nr because data not mature at this stage 48.4 mos [87]; OS nr 60.9 mos [88] OS nr 73 mos [89] 64 vs. 39 deaths, HR 0.59 (95% CI, 0.40-0.88), p=0.009); 6-yr estimate: ABVD vs. A-AVD	mPFS* rate *24.6 mos (2-yr) [9]* (independently determined outcome); 77.2% (95% CI, 73.7-80.4) vs. 82.1% (95% CI, 78.8-85.0), HR for progression, death, or modified progression, 0.77 (95% CI, 0.60-0.98) p=0.04. mPFS Events 146 vs. 117 of which: Disease progression: 70% vs. 77% Death from any cause: 15% vs. 15% Subsequent therapy when complete response not achieved at completion of chemotherapy: 15% vs. 8%. *NOTE: mPFS is not one of the outcomes that we considered important or critical* Subgroups: Age: (number of events/ at risk) <60 yrs old: 117/568 (20.6%), vs. 93/580 (16%), HR 0.73 (95% CI, 0.56-0.96) ≥60 yrs old: 29/102 (28.4%) vs. 24/84 (28.6%), HR 1.00 (95% CI,0.86 (0.58-1.72)	AE *Connors, 2018 (2-yr interim analysis)* Dose reduced prescribed: Bleomycin: 3% vs. 26% p values nr Drug-related serious adverse events: 19% vs. 36% Adverse events resulting in drug discontinuation: 16% vs. 13% Adverse events resulting in dose modification: 44% vs. 64% Death due to drug-related adverse events: 1% vs. 1% Hospitalizations: 28% vs. 37% Serious adverse events 27% vs. 43% Neutropenia:45% vs. 58% Subgroup: Age: <60 yrs old: 6% vs. 17% ≥60 yrs old: 17% vs. 37% Febrile neutropenia: 8% vs. 19% Peripheral neuropathy Grade ≥3: <1% vs. 4%

Study name, Reference, Treatment comparisons	OS	Recurrence	AE (Includes: Treatment-related death, SPM, Pulmonary toxicity, infertility)
	89.4% (95% CI, 86.6-91.7) vs. 93.9% (95% CI, 91.6-95.5), p value nr 6-yr OS estimates in subgroups: PET-2 negative: 94.9% vs. 90.6%, HR for death: 0.54 (95% CI, 0.34-0.86), p value nr. PET-2 positive: 95% vs. 77%, HR 0.16 (95% CI, 0.04-0.72), p value nr.	Location: North America: 57/247 (23.1%) vs. 38/250 (15.2%), HR 0.60 (95% CI, 0.40-0.90) Europe: 74/336 (22.0%) vs. 62/333 (18.6%) HR 0.83 (95% CI, 0.59-1.17) PFS (defined as time from randomization to first occurrence of disease progression or death from any cause) rate per investigator in the ITT population: *37 mos; [86];* 76.0% (95% CI, 72.4-79.2) vs. 83.1% (95% CI, 79.9-85.9), HR 0.704 (95% CI, 0.550-0.901), p=0.005 Subgroups: <60 yrs old: 116/568 (20.4%) vs. 86/580 (14.8%), HR 0.689 (95% CI, 0.521-0.910) ≥60 yrs old: 35/102 (34.3%) vs. 23/84 (27.4%), HR 0.785 (95% CI 0.464-1.329) North America: 56/247 (22.7%) vs. 30/250 (12%), HR 0.490 (95% CI, 0.315-0.764) Europe:81/336 (24.1%) vs. 59/333 (17.7%), HR 0.707 (95% CI, 0.505-0.989) *48.4 mos [87]* PFS rate (per investigator assessment): 75.1% vs. 81.7%, HR 0.691, (95% CI, 0.542-0.881), p=0.003 *60.9 mos [88]* PFS rate (per investigator assessment): 75.3% (95% CI, 71.7–78.5) vs. 82.2% (95% CI, 79.0–85.0), HR 0·68 (95% CI, 0.53–0.87), p=0.0017. Subgroups: PET-2 negative pts: 78.9% (95% CI, 75.2-82.1) vs. 84.9% (95% CI, 81.7-87.6), HR 0.66 (95% CI, 0.50-0.88), p=0.0035 PET positive pts: 45.9% (95% CI, 32.7-58.2) vs. 60.6% (95% CI, 45-73.1). HR 0.70 (95% CI, 0.39–1.26), p=0.23 Pts <60 yrs old: 77.8% (95% CI, 74.0–81.1) vs. 84.3% (95% CI, 81.0–87.1), HR 0.67 (0.51–0.88), p<0.0034 Pts ≥60 yrs old:	Fertility: not assessed *37 mos; [86];* SPM: 3% vs. 2%, p values nr Peripheral neuropathy: *60.9 mos [88]* Peripheral neuropathy (PN): 286 of 659 pts (43%) vs. 443 of 662 (67%) pts PN still ongoing at 5 yrs: 59 of 659 (9%) pts vs. 127 of 662 (19%) pts SPM at 5 yrs Median time to resolution: 29 of 659 (4%) pts vs. 19 of 662 (3%) pts p values nr PN at 5 yrs [88]: 9% vs. 19% p values nr NOTE: assessment of this outcome for Grade 3-4 was confounded because in 12 of 15 pts in the A-AVD group 3 died before resolution, 4 were lost to follow-up, and 5 withdrew from the study. Median time to PN resolution [88] 16 wks (IQR 11-78 wks) vs. 34 wks (IQR13-71 wks) p values nr 72.6 mos [89] On-treatment or treatment-related deaths: 7 (1%) vs. 8 (1%) Second cancers: 4.9% vs. 3.5% Death due to second cancer: 11 (1.6%) vs. 1 (0.2%) Peripheral neuropathy: 288/659 (43.7%) vs. 443/662 (66.9%) Of these: 79.4% vs. 71.8% resolved and 7.7% vs. 13.8% improved. Median time to improvement was: 72.5 wks vs. 42 wks, p value nr for all the above.

Study name, Reference, Treatment comparisons	OS	Recurrence	AE (Includes: Treatment-related death, SPM, Pulmonary toxicity, infertility)
		61.6% (95% CI, 50.9–70.7) vs. 67.1% (95% CI, 55.1–76.5), HR 0.82 (95% CI, 0.49–1.36), p=0.44 72.6 mos [89] 6-yr PFS rates estimates for ABVD vs. A-AVD: 74.5% vs. 82.3%, HR for disease progression: 0.68 (95% CI,0.53-0.86) EFS: nr DFS: nr	
Comparison 5. Modified ABVD vs. another modified ABVD			
No fully published trials of final results are available for this comparison as of February 16, 2024. See Table 11 for abs of interim analyses of ongoing trials in this category.	NA	NA	NA
Comparison 6. BEACOPP vs. Modified BEACOPP			
No fully published trials of final results are available for this comparison as of February 16, 2024. See Table 11 for abs of interim analyses of ongoing trials in this category.	NA	NA	NA

*Primary outcome
** For OS the final analysis after 112 deaths has not been published yet as of our latest update (January 12, 2024

a The results presented in this table are from the per-protocol analysis unless marked intention to treat (ITT).
b Outcomes from 1005 patients (all randomized patients – ITT population)
c The publication is either one of the companion articles or not yet appeared – awaiting for a longer follow-up at the time of this publication
d Data for this outcome are from Borchman 2017 [50]; eBEACOPP n=219, R-BEACOPP n=220; follow-up: 33 months
A-AVD=brentuximab vedotin plus doxorubicin, vinblastine, and dacarbazine; ABVD=doxorubicin, bleomycin, vinblastine, and dacarbazine; AE=adverse events; AML=Acute myeloid leukemia; AVD=doxorubicin, vinblastine, and dacarbazine; bBEACOPP=baseline BEACOPP; BEACOPP=bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, prednisone; b/c=because; CI=Confidence interval; COPP=cyclophosphamide, vincristine, procarbazine, and prednisone; COPP/ABVD=cyclophosphamide, vincristine, procarbazine, and prednisone alternating with doxorubicin, bleomycin, vinblastine, and dacarbazine; COPPEBVCAD-CEC=cyclophosphamide, lumustine, vindesine, melphalan, prednisone, epidoxirubicin, vincristine, procarbazine, vinblastine, and bleomycin; CR=complete remission; CR(u)=complete remission (unconfirmed); DFS=disease-free survival; DS=Deauville five-point scale; eBEACOPP=escalated BEACOPP; EFS=Event-free survival; FFS=Failure-free survival; FFTF=Freedom from treatment failure; G-CSF=granulocyte-colony stimulating factor; HL=Hodgkin lymphoma; HR=Hazard Ratio; IPS=International prognostic score; MDS= myelodysplastic syndromes; mos=months; mPFS=Median progression-free survival; *nr*=not reported; NS=Not significant; ITT=intention to treat; OS=overall survival; PET=Positron emission tomography; PFS=progression-free survival; PN=Peripheral Neuropathy; pts=patients; QOL=quality of life; R-eBEACOPP=bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone in escalated dose; RFS=Relapse-free survival; RT=Radiotherapy; SD=Standard deviation; SPM=Secondary primary malignancies; SPMN=Second primary malignant neoplasms; TRD=Treatment-related deaths; wks=weeks; yr(s)=year(s)

Table 2-7A. Results of studies included for Question 3 (RT consolidation)

Reference, Treatment	OS	Recurrence	AE (Includes: Treatment-related death, SPM)
Comparison 1. Consolidation RT vs. observation			
HD0801 FIL Ricardi, 2021 [75], Rigacci, 2020 [101] RT vs. Observation	nr	PFS at 2 yrs (ITT analysis) 91.3% (95% CI, 80.3-96.3) vs. 85.8% (95% CI, 73.6-92.6), HR, 1.2; 95% CI, 0.59-3.07, p=0.7 EFS at 2 years (ITT analysis) 87.8% (95% CI, 76.0-94.0) vs. 85.8% (95% CI, 73.6-92.6), HR 1.5 (95% CI, 0.6-3.5), p=0.34	SPM: cRT vs. NFT 4 of 118 (3.4%) vs. 2 of 115 (1.7%) p values *nr*
GITIL/FIL HD0607 Gallamini, 2020 [60] (final analysis); Gallamini, 2018 [48] (Long-term results); PET2-negative pts with a baseline mass ≥5 cm: RT consolidation vs. no further treatment	OS at 6 yrs RT vs. No further treatment: 99% (95% CI, 97%-100%) vs. 98% (95% CI, 96% to 100%), p=0.61	Results at 6 yrs, RT vs. no further treatment: PFS at 70.8 mos: 92% (95% CI, 88% to 97%) vs. 90% (95% CI, 85% to 95%), p=0.48; DFS rate for NFT and cRT: 91% (95% CI, 86% to 96%) and 94% (95% CI, 89% to 98%)	nr
EORTC trial No. 20884 Aleman, 2003 [121], 2007 [76] No treatment vs. IF-RT	OS at 8 yrs: PR-RT (n=227): 84% (95% CI: 78-89) CR-no RT (n=246): 85% (95% CI: 78-89) CR-RT (n=176): 78% (95% CI: 70-84) p values comparing the groups: NS OS at 5 yrs: 91% (84-94) vs. 85% (78-90), HR 0.57, (95% CI: 0.31 to 1.06), p=0.07	EFS at 8 yrs: PR-RT (n=227): 76% (95% CI: 69-81) CR-no RT (n=246): 77% (95% CI: 70-82) CR-RT (n=176): 73% (95% CI: 65-80) p values comparing the groups: NS EFS at 5 yrs: 84% (77-89) vs. 79% (72-85), HR 1.27 (95% CI:0.77 to 2.11), p=0.35	Grade 3-4 radiation-related: hematologic toxicity: PR-RT: 15% CR-no RT: - CR-RT: 18% Pulmonary toxicity PR-RT: <1% CR-no RT: - CR-RT: 1% Digestive toxicity: PR-RT: 5% CR-no RT: - CR-RT: 8% Cumulative SPM rate at 8 yrs: PR-RT: 5.7 (95% CI: 3.0-10.8)% CR-no RT: 5.6 (95% CI: 3.0 – 10.0) CR-RT: 12.9 (95% CI: 8.0 – 20.3) Comparison among the three groups: p=0.0177
HD12 Borchmann, 2011 [59], von Tresckow, 2018 [77]	OS at 78 mos: HR, 1.09 (95% CI, 0.74-1.60) OS at 97 mos: Estimated 10 yr OS rates: No RT vs. RT	FFTF* rates at 78 mos: 90.4% vs. 87% (difference, −3.4%, 95% CI, −6.6% - 0.1%, p=0.08) HR, 1.29 (95% CI, 0.97-1.73) PFS	RT comparison: Chemotherapy + RT (n=755) vs. Chemotherapy – RT (n=765) AE Treatment-related death:

Reference, Treatment	OS	Recurrence	AE (Includes: Treatment-related death, SPM)
RT comparison: Chemotherapy + RT (n=755) vs. Chemotherapy – RT (n=765)	91.5% (95% CI, 89.3-93.7) vs. 88.7% (95% CI, 86.0-91.3) HR, 1.21 (95% CI, 0.87-1.70)	HR, 1.28 (95% CI, 0.95-1.73) At 97 mos follow-up: Estimated 10 yrs PFS rates: no RT vs. RT 82.2% (95% CI, 79.0-85.4) vs. 86.8% (84.1-89.6) HR 1.34 (95% CI, 1.02-1.75) (RT was better than no RT: the upper limit of the 95% CI lay just above the non-inferiority margin of 1.74	2 (0.3%) vs. 5 (0.7%) SPM: 44 (5.8%) vs. 32 (4.2%)
RT vs. Chemotherapy			
ECOG E1476 Wiernik, 2009 [78] RT vs. 3×ABVD	The 20 yrs estimate: 43% vs. 66%; HR=0.49 (95% CI, 0.30–0.79) p=0.002	Data on 164 pts (per protocol analysis) Disease-free interval (DFI) Pts with complete response: ABVD vs. RT: HR=0.54; 95% CI, 0.29–1.0; p=0.05) At 20 yrs follow- up: DFI: 70% vs. 52% Relapse: ABVD vs. RT: pts with a CR at randomization 33% vs. 54%, HR=0.49 (95% CI, 0.28–0.89, p=0.02) pts with PR at randomization 56% vs. 48%, p=0.58	Toxicity data analyzed on 199 pts who received consolidation Thrombocytopenia: 10% vs. 1%, p=0.004 Treatment-related death: 1% vs. 1%, p values *nr*
Aviles trial Aviles, 2000 [79] Chemo (EBVD) alone vs. Chemo +RT	OS at 5 yrs: Chemo alone vs. Combined therapy 60% vs. 88%, p<0.01	FFS at 5 yrs: Chemo alone vs. Combined therapy 50% vs. 83%, p<0.01	Pulmonary toxicity (asymptomatic): Combined therapy 12.9% vs. Chemo alone: 0% TRD: 0 vs. 0 SPM: 0 vs. 0

*Critical Outcomes; ‡Important outcomes

ABVD=doxorubicin, bleomycin, vinblastine, and dacarbazine; AE=adverse events; Chemo=chemotherapy; CI=Confidence interval; CR=complete remission; cRT=Consolidation radiotherapy; DFS=disease-free survival; EBVD=epirubicine, bleomycin, vinblastine, and dacarbazine); EFS=event-free survival; FFS=failure-free survival; HR=hazard ratio; IF-RT=involved field radiotherapy; ITT=intention-to-treat; mos=months; NFT=no further treatment; nr=not reported; OS=overall survival; PET=positron emission tomography; PFS=progression-free survival; PR=partial remission; QOL=quality of life; RT=radiotherapy; SPM=secondary primary malignancies; TRD=Treatment-related deaths; yrs=years.

Table 3-7A. Companion studies of the unique trials included for Questions 1 and 3 that addressed important or critical outcomes.

Main study; Author(s); Comparison; Objectives	Companion publications; Author(s); Objectives; Design; Notes	Companion publication related question	Summary results of the companion publication
AHL2011 Casasnovas, 2019 [17] Casanovas, 2020 [127] (for updated late treatment - related AE such as SPM, and fertility) Comparison: 6×eBEACOPP vs. PET-driven treatment: 2×eBEACOPP +4×ABVD in PET negative and +4×eBEACOPP in PET-2-positive pts Objectives: To assess a PET-driven strategy after 2 cycles of eBEACOPP with PET-2-negative pts switching to ABVD, and PET-2-positive pts continuing with eBEACOPP compared with standard care of 6 cycles of eBEACOPP.	Demeestere, 2021 [90] Objectives: To study if switching regimen (BEACCOPP to ABVD) in early responders could reduce the risk of ovarian damage and infertility compared to standard escalated BEACOPP regimen in pts <45 years old Design: cohort study.	Question 1 PET-adapted strategies	Sample: 145 women and 424 men Standard vs. PET-directed therapy: Premature ovarian insufficiency: FSH >24IU/L 46.1% vs. 14.5% OR 0.20 (95% CI: 0.08-0.50, p=0.001); adjusted for age: aOR 0.09 (95% CI: 0.03-0.32, p<0.001) Low ovarian reserve: AMH <0.5ng/mL OR 0.15 (95% CI: 0.04-0.56, p=0.005) Ovarian function recovery: HR 2.52 (95% CI: 1.73-3.67, p<0.0001) Testicular dysfunction: OR 0.26 (95% CI 0.13-0.50), p<0.0001 Pregnancies: Women: 37 vs. 39 Men: 7 vs. 22, OR 3.7 (95% CI: 1.4-9.3, p=0.004)

Main study; Author(s); Comparison; Objectives	Companion publications; Author(s); Objectives; Design; Notes	Companion publication related question	Summary results of the companion publication
HD18 Borchmann, 2018 [16] Comparison: PET-2-positive pts: 8×eBEACOPP (Standard) + 6×eBEACOPP (New Standard) (not included in randomization) vs. 8×R-eBEACOPP PET-2-negative pts: 8×eBEACOPP (Standard) + 6×eBEACOPP (New Standard) vs. 4×eBEACOPP Objectives: PET-2-negative pts: To compare 8×eBEACOPP (Standard) + (after June 2011 amendment) 6×eBEACOPP (New Standard) vs. 4×eBEACOPP and exclude the inferiority of the shorter regimen of at least 6% PET-2-positive pts: 8×eBEACOPP (Standard) + 6×eBEACOPP (New Standard) vs. 8×R-eBEACOPP and show superiority of R-eBEACOPP	Kobe, 2018 [91] Objectives: To identify the best cutoff to separate risk groups after 2 cycles of eBEACOPP by analyzing PFS and OS in pts who underwent 8×eBEACOPP according to their PET-2 results. Design: Subgroup analysis of HD18	Question 1 High vs. low dose	Pts treated with 8×eBEACOPP at 3 yrs: Deauville score 1-2: PFS=92.2% OS=97.6% Deauville score 3: PFS=95.9% OS=99.0% Deauville score 4: PFS=87.6% OS=96.8% Deauville score 4 was indicated as the cut-off in multivariate Cox regression models including large mediastinal mass, extranodal involvement, and high IPS. For Deauville score 4 vs. 1-3: PFS: HR, 2.4 (95% CI, 1.4-4.1), p=0.002) OS: HR, 3.2 (95% CI, 1.3-8.4), p=0.02

Main study; Author(s); Comparison; Objectives	Companion publications; Author(s); Objectives; Design; Notes	Companion publication related question	Summary results of the companion publication
	Ferdinandus, 2023 ABS [92] Objectives: to examine the impact of treatment de-escalation on long-term QOL (as measured with the, EORTC QLQ-C30), TTR-F and TTR-W. Design: cohort	Question 1 PET-adapted treatment	N=1632 pts (83.9%) of all randomized pts. PET-2-negative pts: Treatment reduction from eight to four chemotherapy cycles led to a significantly shorter TTR-F (HR 1.41, p=0.008). PET-2-positive pts: Rituximab caused significantly slower TTR-F (HR 0.70, p=0.0163) and TTR-W (HR 0.64, p=0.0017). HRQoL at baseline and age were the main determinants of 2y HRQoL.
RATHL Johnson, 2016 [49] Comparison: PET negative after 2×ABVD: 4×AVD vs. 4×ABVD PET positive after 2×ABVD: 4×BEACOPP14 vs. 3×eBEACOPP Note: The BEACOPP part of the study was not randomized, therefore not further details are reported) Objectives: to test noninferiority of AVD vs. ABVD in pts who were negative at PET-2	Anderson, 2018 [93] Objectives: To determine the effect of response-adapted chemotherapy regimens on ovarian function in women treated with chemotherapy for advanced HL Design: Prospective cohort	Question 1 PET-adapted treatment	36% received ABVD,49% AVD, 6% received BEACOPP-14, and 9% eBEACOPP; results were grouped by ABVD-AVD vs. BEACOPP AMH decreased during chemotherapy with ABVD-AVD and BEACOPP. Before vs. after treatment: ABVD-AVD group: 9.8 pmol/L (IQR 5.9–18.1) before vs. 1.7 pmol/L (IQR 0.4–4.3) after (p<0·0001) BEACOPP group: 6.8 pmol/L (IQR 2.2–12.8) before vs. 0.08 pmol/L (0.07–0.24) after (p<0·0001) At 1 yr after chemotherapy, AMH concentrations recovery ABVD-AVD group: median of 10.5 pmol/L (IQR 4.3–17.3); BEACOPP group: median 0.11 pmol/L (0.07–0.20). FSH: data from 321 participants with median follow-up of 59.3 mos (IQR 41.9-60.9) FSH recovery to <25 IU/L at 2 yrs: ABVD-AVD group vs. BEACOPP group: 96% of 282 pts treated vs. 67% of 39 pts treated HR 0.37 (95% CI, 0.25-0.56), p=0.0001 Age ≧35 yrs vs. <35 yrs: ABVD-AVD Anti-müllerian hormone recovery: 37% (standard deviation [SD] 10), median 5.7 pmol/L (IQR 2.1–9.7) before treatment vs. 1.4 pmol/L (0.2–4.2) at 2 yrs vs. 127% (SD12), median 10.0 pmol/L (IQR

Main study; Author(s); Comparison; Objectives	Companion publications; Author(s); Objectives; Design; Notes	Companion publication related question	Summary results of the companion publication
			7.5–21.1) before treatment vs. 13.4 pmol/L (7.1–22.3) at 2 yrs, p<0.0001. Recovery time ABVD-AVD vs. BEACOPP: 208.9 ds vs. 529.9 ds, p<0.0001 Proportion of pts aged ≧35 yrs showing recovery: 64 (79%) at 1 yr, and 41 (50%) at 2 yrs after treatment. Proportion of pts <35 yrs: 176 (79%) at 1 yr, and 211 (95%) at 2 yrs after treatment; HR 0.49 (95% CI, 0.37–0.65), p<0.0001.
ECHELON-1 Connors, 2018 [9]; Straus, 2020 [86]; Radford, 2020 [87]; Straus, 2021 [88] Comparison: ABVD vs. A-AVD Objectives: to compare first-line therapy with A-AVD vs. ABVD in pts with stage III or IV classic HL	Ramchandren, 2019 [94] Objectives: To present safety and efficacy outcomes for pts from the main trial who were treated in North America (85 cities across the United States and Canada) Design: subgroup analysis of the original study	Question 1 ABVD and modified ABVD	Subgroup analysis of North American pts: N= 497 (37% of the entire population); ABVD n=247, A-AVD n=250 mPFS by independent review facility at a median follow=up of 24.7 mos: 57 vs. 38 events (progression, death, confirmed non-response), HR 0.596 (95% CI, 0.365-0.899), p= 0.012. AE Any AE grade ≥3 67% vs. 81% Drug-related Grade ≥3 AE 56% vs. 77%
	Evens, 2022 [61] Objectives: Pre-specified subgroup analysis of pts ≥60 yrs of age included in the original study. Pre-specified analyses and post-hoc analyses for the extended follow-up of the efficacy and safety of A-AVD vs. ABVD. Design: subgroup analysis of the original study. Results at 60.9 mos follow-up.	Question 2	N=186 pts with untreated advanced-stage HL from the ECHELON-1 study. ABVD: n=102, median age 66 yrs (range 60-83 yrs); A-AVD: n=84, median age 68 (range 60-82 yrs). Stage III (35%) and IV (64%). Comparison ABVD vs. A-AVD: 5-year PFS per investigator assessment: 61.6% (95% CI, 50.9–70.7) vs. 67.1% (95% CI, 55.1–76.5), HR=0.820 (95% CI, 0.494–1.362, p=0.443 AE 181 pts for this analysis: n=98 vs. 83 Pts received a median of 6 cycles of treatment. ABVD: 71% of pts required a dose modification of bleomycin. Dose reduction: 9%, dose held: 4%, dose delayed: 49%, bleomycin discontinued 28% A-AVD: 80% of pts required a dose reduction/modification of brentuximab vedotin. Dose reduction: 31%, dose held: 5%, dose delayed: 61%, Brentuximab vedotin discontinued 20%. Treatment-related mortality ABVD vs. A-AVD: 5.1% vs. 4.4% Grade ≥3 AE: 80% vs. 88% Any grade pulmonary AE: 13% vs. 2% (this AE was associated with 3 of the 5 treatment-related deaths in the ABVD arm). Any grade febrile neutropenia on study: 17% vs. 37% Grade ≥3 neutropenia: 59% vs. 70%

Main study; Author(s); Comparison; Objectives	Companion publications; Author(s); Objectives; Design; Notes	Companion publication related question	Summary results of the companion publication
			Grade ≥3 peripheral neuropathy: 3% vs. 18%
	Grigg, 2021 abs at 62.9 mos follow-up [95] Objectives: To describe efficacy and safety outcomes for the AYA pts enrolled in ECHELON-1 Design: Exploratory analysis of the original study	Question 1 ABVD and modified ABVD	Pts in the AYA subgroup of the ECHELON-1 trial: N=771, n=375 ABVD, vs. n=396 A-AVD; age range 15-39 yrs. Results at a median follow-up of 60.7 mos 5-yr PFS: 79.4% vs. 86.3%, HR 0.64 (95% CI, 0.45-0.92, p=0.013. AE Peripheral neuropathy: 40% vs. 64% p values nr SPM: 1.4% vs. 1.8% p values nr
	Hutchings, 2022 abs [62] Objectives: To report the pre-specified OS analysis from ECHELON-1, and long-term safety data, at 73 mos follow-up (cut-off June 1, 2021) Design: subgroup analysis of the original trial.	Question 1 ABVD and modified ABVD Question 2	OS events: A-AVD: 39 vs. ABVD: 64 [HR] 0.590; 95% CI 0.396-0.879; p=0.009). Subgroups: stage III (HR: 0.863; 95% CI 0.452-1.648) stage IV disease (HR: 0.478; 95% CI 0.286-0.799), PET2-negative pts (HR: 0.583; 95% CI 0.338-0.856) PET2-positive pts (HR: 0.163; 95% CI 0.037-0.717), pts aged <60 years (HR: 0.509; 95% CI 0.291-0.890), pts aged >=60 years (HR: 0.829; 95% CI 0.469-1.466), Different geographic locations: Europe (HR: 0.783; 95% CI 0.467-1.315) and North America (HR: 0.327; 95% CI 0.153- 0.699). PFS estimate: A-AVD 82.3% (79.1-85.0) vs. ABVD 74.5% (70.8-77.7), (HR: 0.678; 95% CI: 0.532-0.863). Safety: A-AVD vs. ABVD no difference. In both A-AVD and ABVD groups, Treatment-related peripheral neuropathy (PN): 86% (379/443) vs. 87% (249/286) PN cases resolving: (72% s 79%) or improving (14% s 8%) . SPM: A-AVD: 23 cases vs. ABVD 32 in the arm were reported. Pregnancies and live births A-AVD vs. ABVD: 49 vs. 28 and 42 vs. 19.

Main study; Author(s); Comparison; Objectives	Companion publications; Author(s); Objectives; Design; Notes	Companion publication related question	Summary results of the companion publication
HD15 Engert, 2012 [58] 8×eBEACOPP vs. 6×eBEACOPP vs. 8×BEACOPP14 (HD13: 2×ABVD vs. ABV vs. AV; **HD14** 4× ABVD +30Gy RT vs. 2×eBEACOPP +2SABVD+30Gy RT) Objectives: to compare 2 reduced-intensity chemotherapy regimens followed by PET-guided RT, and to assess the influence of erythropoietin on quality of life.	Behringer, 2013 [96] Objectives: To analyze gonadal function of survivors of HL Design: Cohort study presents separate results for HD15 (Reported here) NOTE:	Question 1 High dose vs. low dose chemotherapy.	N=649 pts from study HD15 [58], Included also data from HD13, and HD14 Only results for pts with advanced stage (HD15): *Female survivors (n=232):* After therapy AMH levels=0 µg/L in all age groups. FSH: <30 yrs old: 11.1 U/L; ≥30 yrs old: 29.7 U/L, no difference between arms (p>0.15) Menstrual cycle after therapy: <30 yrs old: 82% resumed cycle ≥30 yrs old: 45% resumed cycle *Male survivors (n=417)* Serum levels of inhibin B and FSH: no difference between arms. Levels compatible with proven fertility: 0.5%; levels corresponding to oligospermia: 88.8%
	Kreissl, 2020 [97] Objectives: To describe a complete range of HRQoL (as measured with the EORTC QLQ-C30) domains from diagnosis up to 5 yrs of survivorship and analyze the influence of pt, lymphoma, and treatment characteristics. Design: Pooled analysis: only results for advanced stage pts are reported here (HD15). Data collected from January 2003 to December 2009. Pts completed the HRQoL questionnaires: immediately after diagnosis, after 2-4 cycles of chemotherapy, after end of first-line treatment and at predefined foloow-up examinations up to 5 yrs.		Only HD15 data are reported here: 5 functioning scales: physical functioning [PF], role functioning [RF], cognitive functioning [CF], emotional functioning [EF], and social functioning [SF], 3 symptom scales: fatigue [FA], pain [PA], and nausea and vomiting [NV], and 6 single-item symptoms: dyspnea [DY], appetite loss [AP], constipation [CO], diarrhea [DI], sleeplessness [SL], and financial [FI]. Pts responded to the questions in a graded format ranging from 1 (not at all) to 4 (very much). Raw data were linearly transformed to 14 HRQoL scores ranging from 0 to 100 each. High scores on the functioning scales represent a better level of functioning; high scores on the symptom scales indicate a higher level of symptom; an absolute change of >10 points is considered a clinically relevant difference for all domains analyzed. Baseline: In advanced-stage patients scores were already low at baseline: mean scores of impairments in HD15: Emotional function (EF), 33.5; Social function (SF), 37.0; Fatigue (FA), 40.1; Sleep problems/sleeplessness (SL), 33.9; and Financial problems (FI, 27.5); Dyspnea (DY), 27.9; Appetite loss (AP), 27.0; Pain (PA), 20.4; Physical functioning (PF), 17.5. During Chemotherapy:

Main study; Author(s); Comparison; Objectives	Companion publications; Author(s); Objectives; Design; Notes	Companion publication related question	Summary results of the companion publication
			HRQoL worsened in most domains during treatment, with fatigue (FAd), role functioning (RFd), and social functioning (SFd) domains (d) being most severely affected: FAd of 51.9; RFd of 54.8, and SFd of 51.1, in HD15. During Follow-up: 1 yr 1: "FI" was the most affected domain of HRQoL, FI 36.8 ≥ 2 yrs: HRQoL was stable. High and persistent deficits in HRQoL exist in survivors independently from the chemotherapy received.
HD9 Diehl, 1998 [64]; Diehl, 2003 [54]; Engert, 2009 [81] Comparison: COPP/ABVD vs. bBEACOPP vs. eBEACOPP Objectives: To evaluate if moderate dose escalation and/or acceleration of standard polychemotherapy is beneficial in advanced-stage HL	Ballova, 2005 [63] Objectives: to compare the bBEACOPP regimen with standard COPP-ABVD in older pts with HL Design: RCT – includes 75 pts aged 66-75 yrs, 68 of which were assessable. 26 treated with COPP/ABVD and 42 with bBEACOPP	Question 2	COPP/ABVD vs. bBEACOPP No statistically significant difference between arms for OS and FFTF At 5 yrs: OS rate: 50% all pts FFTF rate: 55% v 74%, p=0.13 At 6.6 yrs: 8% vs. 21% died of acute toxicity (p value nr) SPM: 12% vs. 14% (p values nr)
	Engel, 2000 [103] Objectives: To test the practicability including schedule adherence, acute hematological toxicity and need for supportive treatment	Question 1 High vs. low dose BEACOPP	Treatment-related death (from acute toxicity): bBEACOPP vs. eBEACOPP 6 (1.4%) vs. 4 (0.96%)
	Anonymous 2005 [120] Erratum	NA	On page 2387, in Table 1, the dose of cyclophosphamide in the regimen of increased-dose BEACOPP should have read 1250 mg per square meter, rather than 1200 mg per square meter, as printed

Main study; Author(s); Comparison; Objectives	Companion publications; Author(s); Objectives; Design; Notes	Companion publication related question	Summary results of the companion publication
HD21 Borchmann, 2022 abs [22] Comparison: 2×eBEACOPP vs. 2× BrECADD then PET; PET negative: 2×eBEACOPP vs. 2×BrECADD PET positive: 4×eBEACOPP vs. 4×BrECADD Objecitves: Part 1 (Tolerability): To test if using a brentuximab based therapy can reduce treatment-related morbidity. Part 2 efficacy: to test the efficacy of eBEACOPP vs. BrECADD	BrECAPP vs. BrECADD Eichenauer, 2017 [98], Damaschin, 2021 [99] Damaschin, 2022 [100] 3 yrs follow-up of the study comparing BrECADD vs. BrECAPP at 34 mos follow-up Feasibility Phase II trial for HD21 Objectives: To test a modified BEACOPP regimen with the introduction of brentuximab vedotin to reduce side effects of treatment BrECADD vs. BrECAPP. This is the report of the final analysis. Design: RCT	Question 1 Modified BEACOPP vs. BEACOPP	N=104, 101 pts in efficacy analysis, 102 pts included in safety analysis n=52 BrECAPP vs. n=52 BrECADD [98] OS at 18 mos [98]: nr OS at 34 mos (3-yr estimate) [99]: 100% vs. 95.4% (95% CI, 89.2-100) Complete response* 42 of 49 pts assessed (86%, [95% CI, 73-94]) vs. 46 of 52 pts assessed (88%, [95% CI, 77-96]) Complete remission* 46 of 49 pts assessed (94%, [95% CI, 83–99) vs. 46 of 52 pts assessed (88%, [95% CI, 77–96]) PFS rates estimates at 18 mos [98]: 95% (95% CI, 85-100) vs. 89% (95% CI, 77-100) PFS rates at 34 mos (3-yr estimate) [100]: 90.2% (95% CI, 80.9-99.5) vs. 89.7% (95% CI, 81.0-98.3) AE Acute grade ≥3 AE: 94% vs. 88% Hematological grade ≥3 AE: 92% vs. 87% Infection: 8% vs. 15% Peripheral neuropathy (all grades): 32% vs. 35%
QUESTION 3			
HD0801 FIL Ricardi, 2021 [75] Comparison Objectives: To examine: a) whether an early PET response-adapted strategy (high-dose chemotherapy + autologous bone marrow transplant is safe and effective (phase II non-randomized [102]), and	Rigacci, 2020 [101] Objectives: To analyze the clinical and biological characteristics and the outcomes of PET-2-negative pts who had a positive end-of-treatment PET, as well as PET-2-positive pts who were offered salvage treatment Design: Retrospective cohort	Question 1 PET-adapted strategies	Of 39 PET+ pts, 38 were treated with salvage therapy. 27 achieved a complete remission. At 27 mos for these pts: 3-yr PFS (ITT analysis): 54% (95% CI, 33%-71%) 3-yr OS: 77% (95% CI, 55%-89%) For the entire population at 27 mos follow up: 3-yr PFS: 80% (95% CI, 76%-83%) 3-yr OS: 97% (95% CI, 94%-98%)

Main study; Author(s); Comparison; Objectives	Companion publications; Author(s); Objectives; Design; Notes	Companion publication related question	Summary results of the companion publication
b) whether PET-2-negative pts could benefit from RT consolidation for areas of bulky disease, provided they maintained negativity for the entire ABVD treatment (randomized comparison [75];) Question 3	Zinzani, 2016 [102] Objectives: To examine whether an early PET response-adapted strategy (high-dose chemotherapy + autologous bone marrow transplant) is safe and effective Design: phase II prospective cohort study (non-randomized)	Question 1 PET driven strategies	103 pts of 512 evaluable were PET-2 positive. 81 received salvage treatment. 2-yr PFS: PET positive: 76% PET negative: 81%

NOTES:
a The Authors used data from HD13 (early stage HL), HD14 (intermediate stage HL), and HD15 (advanced stage HL), and reported results for advanced stage separately.

A-AVD=brentuximab vedotin, doxorubicin, vinblastine, and dacarbazine; Abs=abstract; ABV=doxorubicin, vinblastine, and dacarbazine; ABVD=doxorubicin, bleomycin, vinblastine, and dacarbazine; AE=Adverse events; AMH=anti-Müllerian hormone; aOR=adjusted odds ratio; AP=appetite loss; aSCT=autologous stem-cell transplant; AV=doxorubicin and vinblastine; AVD=doxorubicin, vinblastine, and dacarbazine; AYA=adolescents and young adults; BEACOPP=bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone; BrECADD=brentuximab vedotin, etoposide, cyclophosphamide, doxorubicin, dacarbazine, dexamethasone; BrECAPP=brentuximab vedotin, etoposide, cyclophosphamide, doxorubicin, procarbazine, prednisone; CF=cognitive functioning; CO=constipation; COPP/ABVD=cyclophosphamide, vincristine, procarbazine, and prednisone alternating with doxorubicin, bleomycin, vinblastine, and dacarbazine; DI=diarrhea; DS=Deauville five-point scale; DY=dyspnea; eBEACOPP=escalated BEACOPP; EF=emotional functioning; EORTIC QLQ-C30=European Organization for Research and Treatment of Cancer Quality of Life Questionnaire; FA=fatigue; FI=financial; FSH=follicle-stimulating hormone; G-CSF=granulocyte-stimulating factor; HD=Hodgkin's disease; HL=Hodgkin Lymphoma; HR=Hazard ratio; IQR=interquartile range; ITT=Intention to treat; MDS=myelodysplastic syndromes; mos=months; mPFS=modified progression-free survival; nr=not reported; NV=nausea and vomiting; ON=osteonecrosis; OR=odds ratio; OS=overall survival; PA=pain; PET=positron emission tomography; PET-2=pts status after 2 cycles of PET; PF=physical functioning; PFS=Progression-free survival; PN=peripheral neuropathy; pts=patients; QOL=Quality of life; R-eBEACOPP=eBEACOPP with rituximab; RF=role functioning; sCD30=soluble protein called tumour necrosis factor; SD=standard deviation; SF=social functioning; SL=sleeplessness; SPM=Secondary primary malignancies; SPMN=secondary primary malignant neoplasia; TARC=Thymus and activation-regulated chemokine; TTR-F=time-to-recovery from fatigue; TTR-W=time-to-return to work; yr(s)=year(s)

Table 4-7A. Analyses of the German Hodgkin Study Group prospective individual patient data repository

Study, Objectives	Research Question	Dataset used	Population(s) in the original studies, characteristics	Comparison(s) in the original studies	Results of the pooled analysis
Borchmann, 2019 [104] Objectives: To analyze osteonecrosis complications in pts treated in the German Hodgkin Study Group RCTs Design: descriptive analysis of pts with osteonecrosis from the original studies	Question 1	Includes data from HD12, HD15, and HD18 (It also included data from HD10, HD11, HD13, and HD14 and compares the treatment for early vs. advanced HL) Borchmann, 2011 [59], Engert, 2012 [58], and Borchmann, 2018 [16]	5719 pts with advanced stage treated in the German Hodgkin Lymphoma Study Group RCTs HD10-15 and HD18 Pts with advanced HL Age (median) 33 yrs (16-60): Gender (male): 57%	HD12: A) 8×eBEACOPP+RT vs. B) 8×eBEACOPP no RT vs. C) 4×eBEACOPP +4×bBEACOPP + RT vs. D) 4×eBEACOPP +4×bBEACOPP no RT HD15: n=728 8×eBEACOPP vs. 6×eBEACOPP vs. 8×$BEACOPP_{14}$ HD18: PET-2 positive pts: 8×eBEACOPP (Standard) + 6×eBEACOPP (New Standard)[c] vs. 8×R-eBEACOPP PET-2 negative pts: per protocol analysis 8×eBEACOPP (Standard) + (New Standard)[c] vs. 4×eBEACOPP	Cumulative incidence: (advanced stage): 1.0% [95% CI: 0.71–1.21] Localization: Both advanced and early stage: most pts 72.7% (104/143) had two areas affected, usually bilateral osteonecrosis of the femoral head Intervention: surgery 56.1% (32/57) Outcome of osteonecrosis in HL: severity of symptoms: grade 2: 45.5% vs. grade 3: 54.5% (30 vs. 36 pts) Timing of osteonecrosis: (also early stage) Within 3 yrs after initiation of chemotherapy (83.3%, 55/66 pts) with a peak at 18 mos. Advanced stage pts had earlier onset than early stage pts: 22.3 mos vs. 47.4 mos Risk factors (only advanced age): cumulative corticosteroid dose (OR 1.20 [95% CI: 1.03-1.40], p=0.017), and age (OR 0.76 [95%CI0.58-0.99], p=0.044) were significant prognostic factors
Bröckelmann, 2017 [105] Objectives: To comprehensively analyze HL late relapse (>5 yrs) after risk-adapted first-line treatment	Question 1	HD9 Diehl, 2003 [54] and HD12 [16]	Pts with newly diagnosed HL in unfavourable stage IIB or IIIA or stage IIIB or IV	8×COPP/ABVD no G-CSF vs 8×bBEACOPP no G-CSF vs 8×eBEACOPP G-CSF	15-yr estimates: Cumulative incidence of relapse for advanced stage: 3.9% (95% CI: 2.6-5.2) Cumulative incidence of relapse: early favourable vs. early unfavourable vs. advanced: 5.3% vs. 3.9% vs. 3.9%, p=0.01; In multivariate analysis: Advanced vs. early favourable HR 0.7 (95% CI: 0.4-0.98), p=0.04 Male pts (including also early stage) were more at risk than women: HR 1.6 (95% CI:1.1-2.3), p=0.01 PFS for advanced vs. early favourable stage: HR 2.6 (95% CI: 1.8-3.9), p<0.001 Including pts with early stage: OS of pts in continuing long-term remission vs. those who experienced a late relapse: at 11.2 yrs: 95.2% (95.2-96.4%) vs. 86.1% (80.1-92.1); HR 2.5 (95%CI: 1.7-3.5)
Haverkamp, 2015 [106]	Question 1	HD12 and HD15	3,309 Pts with advanced stage HL	HD15: 8×eBEACOPP vs 6×eBEACOPP vs.	Bleomycin was discontinued in 17.6% and vincristine in 32.6%. A total of 157 patients (4.7%) received ≤4 cycles of bleomycin, and 218 (6.6%) received ≤3 cycles of vincristine; these were

Study, Objectives	Research Question	Dataset used	Population(s) in the original studies, characteristics	Comparison(s) in the original studies	Results of the pooled analysis
Objectives: to analyze the impact of bleomycin and vincristine dose reductions in case of treatment-related adverse events				8×BEACOPP$_{14}$ HD12: A) 8×eBEACOPP+RT in responding pts with initial bulk or residual tumour vs. B) 8×eBEACOPP no RT vs. C) 4×eBEACOPP +4×bBEACOPP + RT vs. D) 4×eBEACOPP +4×bBEACOPP no RT	compared with patients receiving >4 cycles of bleomycin or >3 cycles of vincristine. After a median follow-up of 59 and 67 mos for PFS and OS, respectively, there was no significant difference in PFS or OS in patients receiving ≤ or >4 cycles of bleomycin (5-yr PFS difference, 1.7% (95% CI, -4.2%-7.6%); 5-yr OS difference, 1.5% (95% CI, -2.6%-5.5%). No significant difference in pts receiving ≤ or >3 cycles of vincristine (5-yr PFS difference, -1.3% (95% CI, -5.6%-3.1%); 5-yr OS difference, -0.1% (95% CI, -3.1%-2.9%)
Eichenauer, 2014 [107] Objectives: To provide data on incidence, outcome, and risk factors for the development of treatment related AML-MDS Design: Retrospective analysis	Question 1	HD9, HD12, HD15, PROFE, and BEACOPP-14 pilot trials	Pts with advanced stage HL	HD15: 8×eBEACOPP vs 6×eBEACOPP vs. 8×BEACOPP$_{14}$ HD12: A) 8×eBEACOPP+RT in responding pts with initial bulk or residual tumour vs. B) 8×eBEACOPP no RT vs. C) 4×eBEACOPP +4×bBEACOPP + RT vs. D) 4×eBEACOPP +4×bBEACOPP no RT HD9 8×COPP/ABVD no G-CSF (enrollment for this group was stopped early for benefit at the 1st interim analysis, 2 yrs with 125 pts in the COPP/ABVD, 131 pts in the bBEACOPP arm and 65 pts in the eBEACOPP arm) vs 8×bBEACOPP no G-CSF 8×eBEACOPP G-CSF	Cumulative incidence of treatment-related AML-MDS Difference between treatment groups Differences in OS after treatment-related AML-MDS diagnosis between pts undergoing aSCT and those who did not (Multivariate analysis) 11,952 pts treated for newly diagnosed HL within German Hodgkin Study Group trials between 1993 and 2009 were considered. At a median follow-up of 72 months: treatment-related (t)-AML/MDS diagnosis: 106/11,952 pts (0.9%). Median time from HL treatment to t-AML/MDS: 31 mos. t-AML/MDS pts median age vs. age in the whole group: 43 vs. 34 yrs, p<0.0001. Risk of developing AML/MDS pts who received ≥4 eBEACOPP cycles vs. pts who received <4 eBEACOPP cycles of BEACOPP or no BEACOPP chemotherapy: 1.7% vs. 0.7% vs. 0.3%, p<0.0001. Median OS for all t-AML/MDS pts: 7.2 mos. However, t-AML/ MDS pts who had allogeneic stem cell transplantation: median OS not reached after a median follow-up of 41 mos (p<0.001).
Wongso, 2013 [108] Objectives: To analyze clinical course and risk factors associated with	Question 1	HD9, HD12, and HD15	Pts with advanced stage HL	HD15: 8×eBEACOPP vs 6×eBEACOPP vs. 8×BEACOPP$_{14}$	Incidence of TRM Causes of TRM Risk factors for TRM TRM score

Study, Objectives	Research Question	Dataset used	Population(s) in the original studies, characteristics	Comparison(s) in the original studies	Results of the pooled analysis
treatment-related mortality (TRM) Design: Retrospective analysis				HD12: A) 8×eBEACOPP+RT in responding pts with initial bulk or residual tumour vs. B) 8×eBEACOPP no RT vs. C) 4×eBEACOPP +4×bBEACOPP + RT vs. D) 4×eBEACOPP +4×bBEACOPP no RT HD9 8×COPP/ABVD no G-CSF (enrollment for this group was stopped early for benefit at the 1st interim analysis, 2 yrs with 125 pts in the COPP/ABVD, 131 pts in the bBEACOPP arm and 65 pts in the eBEACOPP arm) vs 8×bBEACOPP no G-CSF 8×eBEACOPP G-CSF	Among a total of 3402 pts, TRM of 1.9% (64 of 3402) was mainly related to neutropenic infections (n=56; 87.5%). Twenty of 64 (31.3%) events occurred during the first course of eBEACOPP. Higher risk of TRM was seen in pts age ≥40 yrs with poor performance status (PS) and in pts age ≥50 yrs. PS and age were then used to construct a new risk score; those with a score ≥2 had TRM of 7.1%, whereas patients who scored 0 or 1 had TRM of 0.9%.

ABV=doxorubicin, vinblastine, and dacarbazine; ABVD=doxorubicin, bleomycin, vinblastine, and dacarbazine; AML/MSD=acute myeloid leukemia and myelodysplastic syndromes; aSCT=autologous stem-cell transplant; AV=doxorubicin and vinblastine; bBEACOPP=baseline BEACOPP; BEACOPP=bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, prednisone; CI=confidence interval; COPP/ABVD=cyclophosphamide, vincristine, procarbazine, and prednisone alternating with doxorubicin, bleomycin, vinblastine, and dacarbazine; eBEACOPP=escalated BEACOPP; G-CSF=granulocyte-colony stimulating factor; HL=Hodgkin lymphoma; HR=hazard ratio; mos=months; OR=odds ratio; OS=overall survival; PET=positron emission tomography; PET-2=pts status after 2 cycles of PET; PFS=progression-free survival; PS=performance status; pts=patients; RT=radiotherapy; t-AML/MDS=treatment-related AML/MDS; TRM=treatment-related mortality; yrs=years;

Table 5-7A. Companion studies of the included trials reporting outcomes that were not considered critical or important.

Original study Objectives	Companion study Objective, Design
GITIL/FIL HD0607 Gallamini, 2020 [60] (final analysis); Gallamini, 2018 [48] (Long-term results) Objectives: 1) To test a risk-adapted strategy on PET positive pts after 2 ABVD cycles 2) To investigate the role of consolidation RT in PET-2 pts treated with ABVD up-front who presented with a baseline mass ≥5 cm	Gallamini, 2021 abs [109] Objectives: To report on the predictive value of Total Metabolic Tumour Volume on ABVD outcomes in PET-2-negative pts Design: Subgroup analysis
ECHELON-1 Connors, 2018 [9]; Straus, 2020 [86]; Radford, 2020 [87]; Straus, 2021 [88] ABVD vs. A-AVD Objectives: to compare first-line therapy with A-AVD vs. ABVD in pts with stage III or IV classic HL	Straus, 2020 [110] Objectives: To test the effect of G-CSF primary prophylaxis vs. no G-CSF Design: Exploratory analysis of the main study
	Radford, 2019 [111] Objectives: To evaluate mPFS according to baseline sCD30 and TARC Design: Exploratory ad-hoc analysis
	Suri, 2019 [112] Objectives: To develop population pharmacokinetic and exposure-response models to examine sources of variability in exposure and safety/efficacy end points Design: Subgroup analysis of the original study
HD15 Engert, 2012 [58] Objectives: to compare 2 reduced-intensity chemotherapy regimens followed by PET-guided RT, and to assess the influence of erythropoietin on quality of life. Comparisons: 8×eBEACOPP vs. 6×eBEACOPP vs. 8×BEACOPP14	Engert, 2010 [113] Objectives: To determine if epoetin alfa reduces anemia-related fatigue, improves health-related pt-reported outcomes, and has an impact on FFTF, and OS Design: RCT double blind study
	Kobe, 2008 [114] Objectives: To evaluate the negative predictive value of PET after 6 to 8 cycles of BEACOPP Design: cohort study (subgroup analysis)
	Kreissl, 2016 [115]

Original study Objectives	Companion study Objective, Design
	Objectives: To describe fatigue from diagnosis to 5 yrs post treatment; to estimate the effect of disease, patient, and treatment variables on fatigue; and to identify pts subgroups with different longitudinal fatigue courses. Design: cohort
	Borchmann, 2019 [116] Objectives: To study incidence, type, timing, and risk factors for thrombotic events in pts with early, intermediate and advanced HL Design: Pooled analysis[a]: only results for advanced stage pts are reported here. The Authors used data from HD13 (early-stage HL), HD14 (intermediate-stage HL), and HD15 (advanced-stage HL), and reported results for advanced stage separately. Incidence of thrombosis: 155 events (7.3%) in the advanced-stage group, higher compared to early (0.7%) and intermediate (1.3%) groups (p<0.001) Incidence in the subgroup of pts of 50-60 yrs of age: 10% (27/269) Timing: most thrombotic events occurred during chemotherapy: (indirect evidence): venous events: 138/175, 78.9%; arterial events: 9/15 (60%) Risk factors for thrombosis (indirect evidence) - logistic regression analysis: Relative frequency: 8×BEACOPP14: 9.4% (67/710), vs. 8×eBEACOPP: 5.7% (40/705), p=0.0079 6×eBEACOPP: 5.1% (36/711) Age: OR (per yr) 1/41.02, p=0.01 Smoking: OR=1.61, p=0.02 Subgroup of female using oral contraception (n=248) vs. no contraception: 6.8% (10/147) vs. 3.9% (4/101), OR=1.77, (CI 0.54 to 5.81)
HD12 Borchmann, 2011 [59] Comparison: 8×eBEACOPP + RT 8×eBEACOPP 4×eBEACOPP+4bBEACOPP+RT 4×eBEACOPP+4bBEACOPP Objectives: To test a regimen that could decrease the toxicity associated with 8 cycles of eBEACOPP + RT while maintaining efficacy, specifically: 1) Chemotherapy: to test the effect of replacing the last 4×eBEACOPP with 4 cycles of baseline BEACOPP; 2) Radiotherapy: to evaluate the impact of using RT for consolidation in pts responding to chemotherapy who had initial bulky disease (≥5 cm or residual disease 1.5 cm	Eich, 2007 [117] Objectives: To report on a subset of 1084 pts analyzed by the RT quality control program Design: subgroup analysis of an RCT at the time of the 5th interim analysis

Original study Objectives	Companion study Objective, Design
AHL 2011 Casasnovas, 2019 [17] Casasnovas, 2022 [84] (for updated late treatment-related AE such as SPM, and fertility) 6×eBEACOPP vs. PET-driven treatment: All pts received 2×eBEACOPP, then PET2. Standard group: 2×eBEACOPP (irrespective of PET results) then PET4. If PET-4-negative: 2×eBEACOPP; if PET-4 positive: salvage treatment. PET-driven treatment: PET2-positive pts: 2×eBEACOPP PET2-negative pts: 2×ABVD Then PET4: Consolidation If PET-4-negative: 2×ABVD If PET-4 positive: salvage treatment	Chevreux, 2022 [118] Objectives: to examine the impact of social disparities on the disease features at diagnosis and analyze how the sociodemographic patient features could impact the HL outcome of patients with advanced-stage HL enrolled in the AHL2011 Design: cohort
HD18 Borchmann, 2018 [16] Objectives: PET-2 positive pts: To show superiority of R-eBEACOPP vs. eBEACOPP for PFS. At the 3-yr interim analysis futility was concluded for this question and these pts were no longer randomized, but they all received further 4 cycles of eBEACOPP. Pet-2 negative pts: To show non-inferiority of eBEACOPP with reduced number of cycles compared to standard eBEACOPP for PFS.	Ferdinandus, 2024 [119] Objectives: to examine the distribution of residual metabolic disease in PET-2 and the prognostic relevance of multiple involved regions Design: cohort, retrospective

A-AVD=Brentuximab vedotin plus doxorubicin, vinblastine, and dacarbazine; abs=abstracts; ABVD=doxorubicin, bleomycin, vinblastine, and dacarbazine; AE=Adverse events; bBEACOPP=baseline BEACOPP; BEACOPP=bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone; eBEACOPP=escalated BEACOPP; FFTF=Freedom from treatment failure; G-CSF=granulocyte-colony stimulating factor; HL=Hodgkin lymphoma; mPFS=modified PFS; OR=odds ratio; OS=Overall Survival; PET=positron emission tomography; PET-2=pts status after 2 cycles of PET; PFS=Progression-free survival; pts=patients; R-eBEACOPP=eBEACOPP with rituximab; RT=radiotherapy; sCD30=soluble protein called tumour necrosis factor; SPM=Secondary primary malignancies; TARC=thymus and activation regulated chemokine; yrs=years
a This document reported on early and advanced HL

Table 6-7A. Question 2: Non-randomized single-arm trials: General characteristics and summary results of included trials.

Study (Author, date, Name, location)	Design Objectives Follow-up	Patients N. Age Stage Gender	Patients' co-morbidity, polypharmacy	Intervention(s)	Outcomes/ results
ABVD and ABVD-like					
Yildiz, 2021 [65] Turkey	Retrospective cohort (chart review from 2004 to 2020) To demonstrate the treatment results of patients with advanced age: OS To demonstrate clinicopathological features and chemotherapy tolerance. Follow-up: 45.2 mos	N=51 (includes early and advanced stage) Age (median, range): 66 yrs (60-76 yrs) Stage: I: 6 (11.8%) II: 13 (25.5%)[a] III: 13 (25.5%) IV: 19 (37.3%) Gender: Male 56.9%	Co-morbidities: present in 78.4% (including diabetes mellitus, hypertension, coronary artery disease, chronic obstructive pulmonary disease, dementia, Parkinson's disease, rheumatological diseases requiring systemic treatment, and secondary malignancies)	First-line treatment: ABVD combination (n=45, 88.2%) 1 to 8 cycles AVD (n=3, 5.9%) Bendamustine/brentuximab (n=1, 2%) ABD (n=1, 2%) GCVP (n=1, 2%) Consolidation RT (n=12, 23.6%)	OS median: 123.8 mos (Std. Error: 56.145; 95% CI: 13.78-233.87) OS at 5 yrs (only advanced stage pts): 54.4% Adverse events: Treatment-related death: 1 (2%) Refused treatment: 1 (2%) Dose reduction due to treatment-related toxicity: 15 (29.4%)

Study (Author, date, Name, location)	Design Objectives Follow-up	Patients N. Age Stage Gender	Patients' co-morbidity, polypharmacy	Intervention(s)	Outcomes/ results
Wrobel, 2019 [66] PLRG-R9 study Poland	Retrospective analysis To provide data on treatment outcomes and complications in elderly HL pts Follow-up (median): 36 (1–215) mos	N=350 of these 149 > 60 yrs of age Age: median 70 yrs Stage: IIB: 18% III: 53% IV: 28% Gender: Male 50%	Comorbidities: 98 (66%) CVD: 74 (50%)	ABVD/ABVD like +/- RT 86% CHOP/PVAG: 7% BEACOPP: 3% Palliative: 4%	Results only for the advanced cohort *(Table 2 in the paper)* TRD: 7 (6%) *Table 2* All pts: 3 yrs EFS: 0.48 (95%CI: 0.37–0.59), 3 yrs OS: 0.76 (95%CI: 0.66–0.85) ABVD treated pts: 3 yrs EFS: 0.54 (95%CI: 0.43–0.66), 3 yrs OS: 0.80 (95%CI: 0.71–0.90) In the multivariate analysis, significant predictors for inferior OS were CVD (most important) (HR=2.76; 95%CI: 1.57–4.87; p=0.00044) whereas the HR for age >70 was 1.73 (95%CI: 0.94–3.19; p=0.079); for EFS: poor performance status (ECOG <2) (HR=1.68; 95%CI: 1.05–1.59, p=0.014), age >70 years (HR=1.42; 95%CI:1.0–2.49, p=0.05), the presence of CVD (HR=1.43; 95%CI: 0.96–2.11, p=0.078;) and the presence of extranodal disease (HR=1.68; 95%CI:1.04–2.71, p=0.033).
Cokgezer, 2022 [74][b]	Retrospective chart review Objectives: to evaluate response, toxicity, and survival in pts ≥50 yrs of age Follow-up: (only pts ≥60 yrs old) median: 47 mos (1–180) mos	N=101 of which 49 of age ≥60 yrs old Age (only pts ≥60 yrs old): median 66 yrs (range 60–85) yrs Stage (only pts ≥60 yrs old): Advanced: 51% Gender (only pts ≥60 yrs old): male, 49.5%	Comorbidity in pts ≥60 yrs old: 83.7%	ABVD: 79% (n=39) of pts ≥60 yrs old AVD and mini-CHOP (n=49) (400 mg/m^2 cyclophosphamide, 25 mg/m^2 doxorubicin, 1.4 mg/m^2 vincristine, and 40 mg/m^2 prednisolone) were given to 3 patients (3%) among cases who were ≥60 years old	Response CR: 92.7% In multivariate analysis: Age ≥60 was a predictor of OS HR: 3.711 (95% CI, 1.698-8.112), when the reference was age 50-59, p=0.001

Study (Author, date, Name, location)	Design Objectives Follow-up	Patients N. Age Stage Gender	Patients' co-morbidity, polypharmacy	Intervention(s)	Outcomes/ results
Brentuximab vedotin alone or in combination					
Evens, 2018 [67] US	Phase 2 single arm To improve the curability of older pts with newly diagnosed HL. To examine several geriatric assessment measures and to correlate these measures and other clinical factors with patient outcome. Follow-up (median): 23 mos (range, 2-49 mos)	N=48 Age: median 69 (range 60 to 88) Stage: IIB:19 % III: 37% IV: 44% Gender: Male 63%	The median CIRS-G comorbidity score at baseline was 7 (range, 1 to 20) with 25 (52%) of 48 of pts having at least one grade 3 or 4 deficit, and 15 (31%) of 48 patients had a CIRS-G score ≥10. Of pts with CIRS-G score ≥10, 5 (33%) of 15 had loss of IADLs, and of pts with loss of IADLs at baseline, 4 (67%) of 6 had a CIRS-G score ≥10.	Sequential BV given before and after AVD (i.e., 2 lead-in doses of BV, six cycles of AVD, and four consolidation doses of BV)	2-yr PFS 84% (95% CI, 69% - 92%) 2-yr EFS 80% (95% CI, 65% - 89%) 2-yr OS 93% (95% CI, 80% - 98%); AE: 2 pts came off study because of toxicity and 3 were lost to follow-up.
Gibb, 2021 [68] BREVITY trial UK	Multicentre Phase 2 response-adapted design To test BV monotherapy as front-line therapy Follow-up: 36 mos	35 pts with advanced stage HL unfit for standard therapy Age: 77 (72-82) Stage: II: 7 (20%) III: 12 (34%) IV: 16 (46%) Gender: Male 63%	LVEF reduced with associated comorbidities or cardiac risk factors: 11% Impaired respiratory function: 3% ECOG PS 1, 2 or 3 and aged >60 years: 28% LVEF and ECOG PS: 26% LVEF, impaired respiratory function and ECOG PS: 3% Impaired respiratory function and ECOG PS:14% LVEF and impaired respiratory function:3% Impaired cardiac function, LVEF and ECOG PS:3% Impaired cardiac function, impaired respiratory function and ECOG PS:3% Impaired cardiac function and ECOG PS: 6%	BV monotherapy administered at an initial dose of 1.8 mg/kg every 3 wks as a 30-min outpatient IV infusion for 4 cycles	PFS median time: 7.3 mos (95% CI 5.2–9.0) At 12 and 24 mos, 13.7% (95% CI 4.3–28.4) and 6.9% (95% CI 1.2–19.6) of pts were progression free OS median: 19.5 months (95% CI 12.6–not reached) AE 18 patients had at least one related AE of Grade ≥3 and 14 serious AEs (SAEs) TRD: 3 (9%)
Friedberg, 2024 [69][c] SGN35-015 (NCT01716806) BREVITY study*	Phase 2 open label not comparative Objectives: 1) To provide long-term follow-up data on the Brentuximab vedotin +dacarbazine (BV-	N= Part B: 22 pts Pard D: 21 pts Age Part B median 69 yrs (range 62-88) Part D: median 77 yrs (range 60-88)	Part B: Limitations to performing physical functioning tasks: 73% At least 1 comorbidity that interfered with quality of life or ≥3 comorbidities: 50%.	BV alone or in combination with dacarbazine or with bendamustine or with nivolumab; there were 6 parts of the study: part A (BV), part B (BV-DTIC),	Part B – OR rate: 95% (95% CI, 77.2-99.9) CR 64% Part BOS median: not reached at 63.6 mos (95% CI 53.4 mos-not estimable)

Study (Author, date, Name, location)	Design Objectives Follow-up	Patients N. Age Stage Gender	Patients' co-morbidity, polypharmacy	Intervention(s)	Outcomes/ results
	DTIC) combination, and 2) To report on a subsequent multicenter phase 2 trial that evaluated the safety and antitumor activity of combination therapy with BV-nivolumab in previously untreated older patients with cHL. Follow-up: Part B: median 63.6 mos. Part D: median 51.6 mos	Stage Part B: 73% stage III, IV Part D: 71% stage III, IV Gender: Part B, male: 73% Part D: male 71%	Part D: Limitations to performing physical functioning tasks: 43% At least 1 comorbidity that interfered with quality of life or ≥3 comorbidities: 38%.	part C (BV-bendamustine), part D (BV-nivolumab), part E (BV), and part F (BV). This paper reports only parts B and D. Part B: BV (1.8 mg/kg) and DTIC (375 mg/m^2) Part D: BV (1.8 mg/kg) and nivolumab (3 mg/kg) on day 1 of each 3-week cycle for up to 16 cycles.	PFS median: 47.2 mos (95% 10.8 mos-not estimable) AE (during treatment): peripheral sensory neuropathy: 48% Grade ≥3 treatment-related AE: 76% Part D: OR rate: 86% (95% CI, 63.7-97) CR 67% OS median: not reached (95% CI not estimable -not estimable) PFS median: not reached (95% CI 9.4 mos-not estimable) AE (during treatment): peripheral sensory neuropathy: 77% Grade ≥3 treatment-related AE: 45%
Cheson, 2020 [70] ACCRU trial USA	Single arm, phase 2, multicentre Objectives: to evaluate the efficacy and safety in older pts and younger pts unsuitable for ABVD Follow-up: 21.2 mos (range 15.6-29.9 mos) *Note:* the study was closed at the first interim analysis after 25 pts were enrolled because only 64% of the first 25 pts had achieved partial or complete response. According to the design, with an OR rate of < 65% the treatment would be considered ineffective.	N=46 – pts ≥60 yrs old: 96% Age (median): 71.5 yrs (range 64-77 yrs) Stage: II: 33% IV: 20% IV: 46% Male: 54%	Not reported	BV and nivolumab	OR rate: 61% (95% CI, 45-75) Stage III and IV only: 60% OS median: not reached PFS: median 18.3 mos (95% CI, 12.7-not reached) AE: TRD: 2% Peripheral neuropathy: 48%

Study (Author, date, Name, location)	Design Objectives Follow-up	Patients N. Age Stage Gender	Patients' co-morbidity, polypharmacy	Intervention(s)	Outcomes/ results
Other treatments					
VEPEMB					
Proctor, 2012 SHIELD study [71] UK, Germany (registration arm)	(1) a prospective phase 2 trial of evaluate VEPEMB and (2) a prospective registration study of pts not treated with VEPEMB and designated as frail or not frail and treated at the discretion of the clinician. Follow-up (median): 36 mos	N=175 Age (median) 73 yrs (range 61-85 yrs) Stage Early (1A/2A): n= 31 Advanced: n=144 Gender (male): VEPEMB 53 (51.5%) ABVD 15 (42.9) CLVPP 13 (68.4) Other 9 (500)	No information on the advanced stage only group. Pts were divided between frail and not frail according to a rating scale. Frail pts (n=13) were excluded from the VEPEMB study. No one in this category reached CR on any form of therapy, and they all died. 67 of 103 non-frail VEPEMB pts achieved CR. In this group 34 of 103 died.	VEPEMB (n=103, of which 72 with advanced stage) ABVD (n=35) CLVPP (n=19) Other (n=18)	VEPEMB arm (excluded pts designated "frail") OS rate at 3 mos: 81% PFS rate at 3 mos: 74% TRD (overall includes early stage): 7% AE (all pts): > Grade 2 haemato-toxicity (includes neutropenic sepsis): 100% (all required G-CSF) Pulmonary fibrosis 1 (1%) Dose reductions: 67% of pts CLVPP group (n=19) No response or progression: 9 (47%) In multivariate analysis in the VEPEMB cohort for PFS identifies CR as a significant factor ($p<0.001$) along with age linked to failure of comorbidity assessment ($p<0.001$). *page 6010*
Levis, 2004 [72] Italy This is the VEPEMB that initially introduced	Phase 2 single arm To devise a less toxic treatment for older pts with HL that does not increase the relapse rate. Follow-up: 12 mos	N=105 Age (mean) 71 (range 66-83) Stage IA-IIA: 48 (45%) IIB-IV: 57 (54%) *Note: this study presented the results for advanced stage patients separately.* Gender: male 53%	Included both advanced and early stage: Pts with co-morbidities: 39 (37%)	6×VEPEMB+RT to bulky/not responding areas.	Only advanced stage: CR rate: 58% 5-yr FFS: 34% 5-yr actuarial RFS rate: 66% 5-yr actuarial OS rate: 32% 5-yr actuarial DSS rate:37% Grade 3-4 neutropenia: 100% (required G-CSF) Chemotherapy plan interruption/modification: 26% Hospitalization during treatment: 21% Prognostic factors in multivariate analysis: stage, systemic symptoms and comorbidity affected OS, DSS and FFS
PVAB					
Ghesquieres, 2024[b] [73] France LYSA study	Prospective Phase 2 multicentre non randomized	N=89 Age (median) 68 yrs (range 61-88 yrs) Stage:	Co-morbidity was measured with the Cumulative Illness Rating Scale for Geriatrics (CIRS- G). Total score, median 3 (0-12)	PVAB (i.e., 6 cycles of prednisone [40 mg/m^2, days 1-5], vinblastine [6 mg/m^2, day 1], doxorubicin [40 mg/m^2, day 1], and	Response: CMR: 77.5% (95% CI 67-86) PR: 65% Stable disease: 1%

Study (Author, date, Name, location)	Design Objectives Follow-up	Patients N. Age Stage Gender	Patients' co-morbidity, polypharmacy	Intervention(s)	Outcomes/ results
	Objectives: To investigate a new first-line therapy for older pts with advanced-stage HL. Follow-up: 42 mos (95% CI 40-47)	IIB: 3% III: 34% IV: 63% Gender: male 65%	No grade 3-4 comorbidity: 91% ≥1 grade 3-4 comorbidity: 9%	bendamustine [120 mg/m^2, day 1])	PFS rate at 4 yrs: 50% (95% CI 39-61) Toxicity: TRD: 4.5% Death from SPM:6 (25%) Pts who experienced at least 1 grade ≥ Adverse event: 94% Pts who experienced at least 1 severe AE: 31.5%

* Historical preface: Previously to this, Forero-Torres et al. 2015 [128] and Gibb et al. (BREVITY trial) [68,129] (we excluded these) studied the use of brentuximab alone, and found a high relapse rate. Therefore, here the combination with dacarbazine and bendamustine was tested and published in a 2016 abstract Yasenchaket al. trial [130]. This study is the full publication of that abstract.

a This study [65] did not specify whether stage II included patients with risk factors (stage IIB) or not.
b This study was identified as an abstract, initially published in 2021 [131], and it was initially excluded as an abs publication of an ongoing trial. We contacted the Authors and found the full text publication of the final results and added it to this table.
c Historical preface to this publication: Previously to this, Forero-Torres et al. 2015 [128] and Gibb et al. (BREVITY trial) [68,129] (we excluded these) studied the use of brentuximab alone, and found a high relapse rate. Therefore, here the combination with dacarbazine and bendamustine was tested in a 2016 Yasenchak trial (old abstract not included).

ABD=Doxorubicin, bleomycin, dacarbazine; ABVD=doxorubicin, bleomycin, vinblastine, and dacarbazine; AE=adverse events; AVD=doxorubicin, vinblastine, and dacarbazine; BEACOPP=bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone; BV=brentuximab vedotin; BV-DTIC=brentuximab vedotin + dacarbazine; CHOP=Cyclophosphamide, doxorubicin, vincristine and prednisolone; CI=confidence interval; CIRS-G=Cumulative Illness Rating Scale-Geriatric; CLVPP=chlorambucil, vinblastine, procarbazine hydrochloride, and prednisolone; CMR=complete metabolic response; CR=complete remission; CVD=cardiovascular disorders; DSS=disease-specific Survival; ECOG=Eastern Cooperative Oncology Group; EFS=event-free survival; FFS=failure-free survival; G-CSF=granulocyte-colony stimulating factor; GCVP=gemcitabine, cyclophosphamide, vincristine, dexamethasone; HL=Hodgkin lymphoma; HR=hazard ratio; IADLs=instrumental activities of daily living; IV=intravenous; LVEF=left ventricular ejection fraction; mos=months; OR rate=objective response rate; OS=overall survival; PFS=progression-free survival; PS=Performance Status; pts=patients; PVAB=prednisone, vinblastine, doxorubicin and bendamustine; PVAG=prednisone, vinblastine, doxorubicin, and gemcitabine; RFS=relapse-free survival; RT=radiotherapy; SAEs=serious AEs; SPM=secondary primary malignancies; Std=standard; TRD=treatment-related deaths; VEPEMB=vinblastine, cyclophosphamide, procarbazine, etoposide, mitoxantrone and bleomycin; wks=weeks; yrs=years.

Table 7-7A. Excluded abstracts and abstracts of ongoing trials published in 2020-2024.

Date	Study	Design / PUBLICATION TYPE	Intervention	comparison	Outcomes	# pts: Intervention/ (comparison)	Age yrs	Comments
2022	Fiad, 2022 [132] GATLA LH-05	Retrospective analysis including only pts >60 yrs ABS	ABVD PET adapted	=	PFS, OS	59 Pts	≥ 60	Not fully published as of March 21, 2024 This is a subgroup analysis of the GATLA LH-05 study (which was published in 2019 https://pubmed.ncbi.nlm.nih.gov/30864146/ THIS WAS A PET TRIAL)
2021	Makita, 2021 [133] Horizon Study	Multicentre retrospective ABS	ABVD	=	OS, PFS, EFS	171	71	It was excluded as ABS of retrospective trial Not fully published as of March 21, 2024. Authors contacted and full publication in preparation
2020	Phillips T #5223	A retrospective analysis with propensity score matching (medical and pharmacy claims data in the US Symphony Health Solutions database) ABS	A-AVD, ABVD	=	Pts characteristics, De-escalation, use of PET	340 in A-AVD 651 in ABVD	≥60	Not fully published as of Dec 5, 2023
2020	De Colella HALO trial (NCT02467946) #16839	Prospective international multicenter open-label phase I/II study ABS	Brentuximab vedotin and bendamustine	=	Feasibility and efficacy of BV BE	59	62-79	Not fully published as of Dec 5, 2023 Similar to Yasenchak, this study, of which we identified 4 publications [134-137] tested the combination of brentuximab and bendamustine
2020	Aussedat G. LYSA #5858	Prospective Phase 2 ABS	PVAG	=	CR PR, PFS, OS	49	>60	Not fully published as of Dec 5, 2023 Includes also early unfavourable
Abstracts of interim analyses								
2022	Greil R. GHSG-HD21 #16753	Single arm of older pts parallel to the HD21 RCT ABS of interim	BrECADD	eBEACOPP PET-driven	Toxicity, efficacy	85 pts	older	Not fully published as of Dec 5, 2023 ABS of ongoing trial Additional non comparative cohort to the HD21 trial.
2021	Flinn I. SGN35-027, trial in progress (EudraCT 2020-004027-17) #5632	Open-label, multiple part, multicenter, phase 2 clinical trial. Part A will evaluate the efficacy and safety of A-AVD when administered with growth factor prophylaxis in pts with Stage III/IV cHL. Part B will evaluate the	Brentuximab vedotin + nivolumab, doxorubicin, dacarbazine	=	Efficacy	Part A, 50 (Part B, Part C, 150) only part A has advanced stage pts	≥ 60	Not fully published as of Dec 5, 2023 EXCLUDE abs of interim

Date	Study	Design / PUBLICATION TYPE	Intervention	comparison	Outcomes	# pts: Intervention/ (comparison)	Age yrs	Comments
		combination of AN+AD in pts with Stage I or II cHL with bulky mediastinal disease or Stage III or IV cHL. Part C will evaluate the efficacy and tolerability of AN+AD in pts with Stage I or II cHL without bulky disease. ABS of INTERIM						
2020	Herrera AF. #5542	Phase 3 RCT SWOG S1826 ABS of interim	N-AVD	A-AVD	OS, PFS	98	≥60	Not fully published as of Dec 5, 2023 ABS of interim ABS with final results published on Dec 12 at the ASH meeting
2023	Herrera, 2023 [23]	Phase 3 RCT SWOG S1826 2nd planned interim analysis ABS of interim	N-AVD	A-AVD	PFS*, OS, EFS, at 2 yrs; Metabolic complete response rate, incidence of AE at 10 yrs QOL ancillary studies PFS: @ 1 yr 86% vs. 94%, HR 0.48, 99% CI 0.27–0.87, one-sided p=0.0005) TRD: 7 (0.14%) vs. 4 (0.082%) Grade ≥3 hematologic AE: 30.5% vs. 48.4% Pneumonitis (any grade): 3.2% vs. 2.0% Peripheral neuropathy (any grade, ≥3 grade): Sensory: 54.2%,7.8% vs. 28.1%, 1.2%	N=976 pts with advanced stage HL Stage: III, IV Gender: 56% male	Age: median 27 (range 12 to 83); 24% were <18	These are the results of the 2nd planned interim analysis at 12.1 mos of follow-up. According to NCT https://classic.clinicaltrials.gov/ct2/show/NCT03907488 primary completion is June 30 2023, and estimated study completion date is 2025-04-01

Date	Study	Design / PUBLICATION TYPE	Intervention	comparison	Outcomes	# pts: Intervention/ (comparison)	Age yrs	Comments
					Motor: 6.8% vs. 4%			
2023	Borchmann [24] NCT02661503	Multicentre, phase 3, open-label, non-inferiority RCT Part 2 efficacy ABS of interim	BrECADD	2×eBEACOPP vs. 2× BrECADD then PET; PET negative: 2×eBEACOPP vs. 2×BrECADD PET positive: 4×eBEACOPP vs. 4×BrECADD	PFS at 5 yrs	150; at this interim analysis 100 PFS events	≤60	36 mos interim analysis
2022	Borchmann, 2022 abs [22] NCT02661503	Multicentre, phase 3, open-label, non-inferiority RCT Part 1 tolerability ABS of interim	BrECADD	2×eBEACOPP vs. 2×BrECADD then PET; PET negative: 2×eBEACOPP vs. 2×BrECADD PET positive: 4×eBEACOPP vs. 4×BrECADD	Treatment-related morbidity	1500 pts	18 to 60 yrs old	This study has final results, but it is not fully published in a peer-reviewed format

A-AVD=Brentuximab vedotin plus doxorubicin, vinblastine, and dacarbazine; ABS=abstract; ABVD=doxorubicin, bleomycin, vinblastine, and dacarbazine; AE=Adverse events; AN-AD=Brentuximab vedotin + nivolumab and Brentuximab vedotin + dacarbazine; AVD=doxorubicin, vinblastine, and dacarbazine; BE=bendamustine; BEACOPP=bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, prednisone; BrECADD=brentuximab vedotin, etoposide, cyclophosphamide, doxorubicin, dacarbazine, dexamethasone; BV=Brentuximab vedotin; cHL=classical Hodgkin lymphoma; CI=confidence interval; CR=complete remission; eBEACOPP=escalated BEACOPP; EFS=event-free Survival; HR=hazard ratio; N-AVD=nivolumab-AVD; OS=overall survival; PET=positron emission tomography; PFS=progression-free survival; pts=patients; PR=partial remission; PVAG=prednisone, vinblastine, doxorubicin, and gemcitabine; QOL=quality of life; RCT=randomized controlled trial; TRD=treatment-related death; US=United States.

Appendix 8: Amendments to the project plan

First amendment

On September 20, 2022, the Working Group noted that no RCT studies with a sample of ≥100 patients were located for Question 2. Therefore, the search was widened for the older population to studies that were either RCT or comparative observational with a sample of a minimum of 30 patients.

Second amendment

On July 14, 2023, after considering the possibility of conducting a network meta-analysis to determine which is the best treatment for patients included in Question 1 (patients younger than 60 years of age), the Working Group acknowledged that two pivotal studies, the SWOG S1826, and the HD21, may be practice changing, and worth to include in this guideline. Unfortunately, at this date (October 16, 2023) we only have abstract publications of interim analyses of the SWOG S1826 trial [125,138], and Part 1 (tolerability) for the HD21 trial [22]. The estimated completion date for these trials is 2024/2025.

In absence of, peer-reviewed, full publications of the SWOG S1826 (NCT03907488) and HD21 we cannot make recommendations that include the use of modified BEACOPP and modified ABVD treatments because the results are incomplete, particularly for methods details and adverse events, and after the publication of these studies, changes may occur that invalidate the resulting recommendations. Therefore, the decision was made to complete the quality assessment of the included studies for Question 1 and Question 3. Then, this part of the guideline will be paused, and it will be presented as a summary of the evidence for the time being. When the full publications of the SWOG and HD studies will appear, we will conduct an update search, and integrate these studies and any other new evidence, and complete the recommendations for the younger population. This could be a timely update of this guideline after publication, or we can make this change before publication, depending on when the newly published studies appear.

The decision was made to continue work on Question 2 (patients older than 60 years) and complete recommendations for that population.

Third amendment

On October 13, 2023, the Working Group examined the available evidence for Question 2, and it became apparent that no comparative evidence has been found for some of the interventions of interest for Question 2 (i.e., PVAG, CHOP, CHLVPP, brentuximab alone, and other agents). In consultation with the PEBC leadership the decision was made to search for single arm Phase II trials for these agents. These trials will be eligible for inclusion if they have a minimum sample size of 30 patients (of higher depending on how many relevant articles will be included) and if they include a minimum of 80% of patients of age ≥60 years, and a minimum of 80% of patients with advanced-stage HL. Quality of these studies will not be assessed because they are considered at high risk of bias.

Fourth amendment

On December 15, 2023, the Working Group acknowledged that two pivotal studies are still ongoing, and not yet published in full text. An early abstract publication of the SWOG S1826 [23,24] study indicates that treatment with nivolumab may be advantageous for older patients as well as for younger ones in terms of survival outcomes and reduced adverse effects.

Therefore, the decision was made to limit this document to an evidence summary and to resume work and make recommendations once this study will be fully published.

Appendix 9: Glossary

Acronym	Definition
4bB	4× baseline BEACOPP
4eB	4× escalated BEACOPP
A-AVD	Brentuximab vedotin plus doxorubicin, vinblastine, and dacarbazine
ABD	Doxorubicin, bleomycin, dacarbazine
abs	Abstracts
ABV	Doxorubicin, vinblastine, and dacarbazine
ABVD	Doxorubicin, bleomycin, vinblastine, and dacarbazine
ABVD DD-DI	Dose-dense/dose-intense ABVD
AE	Adverse Events
AHRQ	Agency for healthcare research and quality
AMH	anti-Müllerian hormone
AML	Adult acute myeloid leukemia
AML-MSD	Acute myeloid leukemia and myelodysplastic syndromes
AMSTAR	Assessing the methodology quality of Systematic Reviews
AN+AD	Brentuximab vedotin + nivolumab and Brentuximab vedouin + dacarbazine
aOR	Adjusted odds ratio
AP	Appetite loss
ASCR	Autologous stem cell rescue
ASCT	Autologous stem cell transplantation
AUCinf	Area Under the Plasma Concentration-time Curve From Time 0 to Infinity for Brentuximab Vedotin
AV	Doxorubicin and vinblastine
AVD	Doxorubicin, vinblastine, and dacarbazine
AYA	Adolescents and young adults
B_{14}	$BEACOPP_{14}$
bB	Baseline BEACOPP
bBEACOPP	Baseline BEACOPP
b/c	Because
BE	Bendamustine

BEACOPP	Bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone
$BEACOPP_{14}$	BEACOPP for 14 days
BEAM	Carmustine, etoposide, cytarabine, melphalan
BrECADD	Brentuximab vedotin, etoposide, cyclophosphamide, doxorubicin, dacarbazine, dexamethasone
BrECAPP	Brentuximab vedotin, etoposide, cyclophosphamide, doxorubicin, procarbazine, prednisone
BV	Brentuximab vedotin
BV-DTIC	Brentuximab vedotin + dacarbazine
CADTH	Canadian agency for drugs and technologies in health
CEC	Carboplatin, etoposide, and cyclophosphamide
CF	Cognitive functioning
cHL	Classical Hodgkin lymphoma
CHLVPP	Chlorambucil, vinblastine, procarbazine, prednisolone
CHOP	Cyclophosphamide, doxorubicin, vincristine and prednisolone
CI	Confidence interval
CIRS-G	Cumulative illness rating scale-geriatric
CLVPP	Chlorambucil, vinblastine, procarbazine hydrochloride, and prednisolone
Cmax	Maximum serum concentration
CMR	Complete metabolic response
CO	Constipation
COPP	Cyclophosphamide, vincristine, procarbazine, and prednisone
COPPA	COPP/ABVD
COPP/ABVD	Cyclophosphamide, vincristine, procarbazine, and prednisone alternating with doxorubicin, bleomycin, vinblastine, and dacarbazine
COPPEBVCAD-CEC	Cyclophosphamide, lomustine, vindesine, melphalan, prednisone, epirubicin, vincristine, procarbazine, vinblastine, and bleomycin
CR	Complete remission
cRT	Consolidation radiotherapy
CR(u)	Complete remission (unconfirmed)
CT	Computed tomography
CVD	Cardiovascular disorders
DFI	Disease-free interval
DFS	Disease-free survival

DOCR	Duration of complete remission
DOR	Duration of response
d(s)	Days
DS	Deauville five-point scale
DSS	Disease-specific survival
DY	Dyspnea
eB	Escalated BEACOPP (in figures)
eBEACOPP	Escalated BEACOPP (in text and tables)
EBVD	Epirubicin, bleomycin, vinblastine, dacarbazine
ECOG	Eastern cooperative oncology group
EF	Emotional functioning
EFS	Event-free survival
EMBASE	Excerpta medica database
EORTC QLQ-C30	European organization for the research and treatment of cancer quality of life questionnaire C-30
EPISTEMONIKOS	Health evidence database
FA	Fatigue
Feb	Febrile
FFFP	Freedom from first progression
FFP	Freedom from progression
FFS	Failure-free survival
FFSP	Freedom from second progression
FFTF	Freedom from treatment failure
FI	Financial
FSH	Follicle-stimulating hormone
G-CSF	Granulocyte colony-stimulating factor
GCVP	Gemcitabine, cyclophosphamide, vincristine, dexamethasone
Gy	Gray
HDT	High-dose therapy
HL	Hodgkin lymphoma
HR	Hazard ratio
IADLs	Instrumental activities of daily living

IF-RT	Involved field radiotherapy
IHP	International harmonization project
IPD	Individual participant data
IPI	International prognostic index
IPS	International prognostic score
IQR	Interquartile range
ITT	Intention-to-treat
IV	Intravenous
LH	Luteinizing hormone
LVEF	Left ventricular ejection fraction
MCR	Medical research council
MDS	Myelodysplastic syndromes
MEDLINE	Medical literature analysis and retrieval system online
MOPP-ABV	Dacarbazine and mechlorethamine, vincristine, procarbazine, prednisone/doxorubicin, bleomycin, vinblastine
mos	Months
mPFS	modified PFS
NA	Not applicable
N-AVD	Nivolumab-AVD
NFT	No further treatment
NHL	Non-Hodgkin lymphoma
NIHR HTA	National institute for health and care research, health technology assessment
nr	Not reported
NS	Not significant
NV	Nausea and vomiting
ON	Osteonecrosis
OR	Odds ratio
OR Rate	Objective response rate
ORR	Overall response rate
Obs	Observation
OS	Overall survival
PA	Pain

PET	Positron emission tomography
PET-2	pts status after 2 cycles of PET
PF	Physical functioning
PFS	Progression-free survival
PN	Peripheral Neuropathy
POI	Premature ovarian insufficiency
PR	Partial remission
PRISMA-IPD	Preferred reporting items for systematic reviews and meta-analyses for individual patient data systematic reviews
PROSPERO	International prospective register of systematic reviews
PS	Performance status
pts	Patients
PVAB	Prednisone, vinblastine, doxorubicin and bendamustine
PVAG	Prednisone, vinblastine, doxorubicin, and gemcitabine
QOL	Quality of life
RCT	Randomized Controlled Trial
R-eBEACOPP	eBEACOPP with rituximab
RF	Role functioning
RFS	Relapse-free survival
Ritux	Rituximab
RR	Relative risk
RT	Radiotherapy
SAEs	serious adverse events
sCD30	Soluble protein called tumour necrosis factor
SD	Standard deviation
SEER	Surveillance, epidemiology and end results
SF	Social functioning
SL	Sleeplessness
SOF	Summary of findings
SPM	Second primary malignancies
SPMN	Second primary malignant neoplasms
SQ	Study question

Stanford V	Doxorubicin, vinblastine, mechlorethamine, vincristine, bleomycin, etoposide, prednisone
Std	Standard
t-AML/MDS	Treatment-related AML/MDS
TARC	Thymus and activation-regulated chemokine
TRD	Treatment-related deaths
TRM	Treatment-related mortality
TTP	Thrombotic thrombocytopenic purpura
TTR-F	Time-to-recovery from fatigue
TTR-W	Time-to-return to work
UK	United Kingdom
US	United States
VEPEB	Vinblastine, etoposide, bleomycin, epirubicin, and prednisone
VEPEMB	Vinblastine, cyclophosphamide, prednisolone, procarbazine, etoposide, mitoxantrone, and bleomycin
vs	Versus
wks	Weeks
yrs	Years