HODGKIN LYMPHOMA — STAGE III AND IV

Expert Panel on Radiation Oncology—Hodgkin Lymphoma: John P. Plastaras, MD, PhD; Ranjana Advani, MD; Leslie K. Ballas, MD; Bouthaina S. Dabaja, MD; Sughosh Dhakal, MD; Christopher R. Flowers, MD, MS; Chul S. Ha, MD; Bradford S. Hoppe, MD, MPH; David B. Mansur, MD; Nancy P. Mendenhall, MD; Monika L. Metzger, MD; Kenneth B. Roberts, MD; Ronald Shapiro, MD; Sonali M. Smith, MD; Stephanie A. Terezakis, MD; Karen M. Winkfield, MD, PhD; Anas Younes, MD; Louis S. Constine, MD.

Summary of Literature Review

Introduction/Background

This review for Hodgkin lymphoma (HL) addresses the treatment of patients who are newly diagnosed with clinical stage III or IV HL. There is only low-level evidence on the treatment of stage III or IV nodular lymphocyte-predominant HL, so this review is limited to advanced-stage classical Hodgkin lymphoma (cHL). It is the result of a comprehensive review of the literature and expert opinion. A combination of doxorubicin, bleomycin, vinblastine, and dacarbazine (ABVD) has been the most widely used chemotherapy regimen for HL. Alternatives to ABVD have been developed for patients with locally extensive or advanced-stage disease, including BEACOPP (a combination of bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone) and its variants and the Stanford V regimen. Dose-escalated BEACOPP and its variants have been shown to be superior to standard-dose chemotherapy, although at the expense of increased toxicity, including infertility and leukemia risks. Stanford V, when given with radiation therapy (RT) as specified by the original protocol (involved-field RT [IFRT] to initial sites ≥5 cm and/or macroscopic splenic disease), yields results comparable to those of ABVD. Most randomized trials showed no significant benefit with the addition of consolidative RT after a complete response (CR) to chemotherapy based on computed tomography (CT) scans. There are data to suggest, however, that there may be a role for RT in patients with initially bulky disease or in patients with a partial response (PR) to chemotherapy. Given the strong prognostic value of early positron emission tomography (PET) findings showing the disease’s response to chemotherapy in patients with advanced-stage HL, investigators are exploring response-adapted therapy, including changing the chemotherapy regimen or omitting RT based on early PET-detected response.

Diagnosis of Advanced-Stage Hodgkin Lymphoma

The diagnosis of HL should be confirmed pathologically by a biopsy specimen of sufficient size to enable architectural interpretation and immunohistochemical analysis. The most recent iteration of the World Health Organization classification combines nodular sclerosis, mixed cellularity, lymphocyte-rich, and lymphocyte-depleted HL into one entity termed cHL [1]. Staging for HL most commonly employs the 1989 Cotswold revision of the Ann Arbor staging classification. Stage III is defined as involvement of lymph node regions or structures on both sides of the diaphragm (with or without splenic involvement), whereas stage IV requires involvement of extranodal sites [2]. Similarly, the Cancer Staging Manual (7th edition) by the American Joint Committee on Cancer (AJCC) defines stage III as involvement of lymph nodes on both sides of the diaphragm, which also may be accompanied by extralymphatic extension or by involvement of the spleen. Stage IV disease is defined as diffuse or disseminated involvement of 1 or more extralymphatic organs, or isolated extralymphatic organ involvement in conjunction with disease in distant sites, or any involvement of the liver or bone marrow, or

1Principal Author, University of Pennsylvania Health System, Philadelphia, Pennsylvania. 2Stanford Cancer Center, Stanford, California, American Society of Clinical Oncology. 3University of Southern California Keck School of Medicine, Los Angeles, California. 4University of Texas MD Anderson Cancer Center, Houston, Texas. 5University of Rochester Medical Center, Rochester, New York. 6Emory University, Atlanta, Georgia, American Society of Clinical Oncology. 7University of Texas Health Science Center at San Antonio, San Antonio, Texas. 8University of Florida Proton Therapy Institute, Jacksonville, Florida. 9University Hospitals Seidman Cancer Center Case Western Reserve University School of Medicine, Cleveland, Ohio. 10University of Florida, Gainesville, Florida. 11St. Jude Children’s Research Hospital, Memphis, Tennessee, American Society of Clinical Oncology. 12Yale University School of Medicine, New Haven, Connecticut. 13Richard L. Roudebush VA Medical Center, Indiana University School of Medicine, Indianapolis, Indiana. 14University of Chicago, Chicago, Illinois, American Society of Hematology. 15Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins Hospital, Baltimore, Maryland. 16Massachusetts General Hospital, Boston, Massachusetts. 17Memorial Sloan Kettering Cancer Center, New York, New York, American Society of Clinical Oncology. 18Panel Chair, University of Rochester Medical Center, Rochester, New York.

The American College of Radiology seeks and encourages collaboration with other organizations on the development of the ACR Appropriateness Criteria through society representation on expert panels. Participation by representatives from collaborating societies on the expert panel does not necessarily imply individual or society endorsement of the final document.

Reprint requests to: publications@acr.org
nodular involvement of the lung(s). A bone marrow biopsy has been commonly used in patients with advanced-stage disease, but the additional information of bone marrow sampling is becoming controversial in patients staged with PET/CT [3]. Most recently, The Lugano classification modifies the Ann Arbor system by formally eliminating the need for routine bone marrow biopsies in HL, creating a separate category termed “II bulky” (with acknowledgement that this may be treated as limited or advanced disease based on histology and prognostic factors), removing the designation “E” for extranodal sites in advanced-stage disease, and removing the designation “X” for bulky disease. However, the definition of bulky remains as commonly defined by the Ann Arbor staging system as a mediastinal mass that has a width greater than one-third the maximum intrathoracic diameter on an upright posteroanterior chest radiograph or a mass >10 cm on a CT scan [4]. It is important to note that clinical trials have defined bulky disease differently, with the smallest dimensions used being 5 cm or the presence of focal splenic nodules (macroscopic splenic disease) [5-7]. Due to similarity in outcomes, patients with stage IIIB/IIBX HL have been included in many trials with stage III and IV patients.

Prognostic Factors

The International Prognostic Project [8] identified risk factors that, when added together, give a score (International Prognostic Score [IPS]) that indicates the likelihood of freedom from disease progression. This study was devised using patients mainly from the 1980s treated with chemotherapy with or without radiation. Seven adverse prognostic factors had similar independent prognostic effects. The 7 factors are male sex, age 45 years or older, stage IV disease, albumin <4 g/dL, a hemoglobin level <10.5 g/dL, a leukocyte count of at least 15,000 per cubic millimeter, and a lymphocyte count of <600 per cubic millimeter or <8% of the white cell count or both. Five-year freedom-from-progression rates based on the number of factors ranged from 84% for those with zero factors to 42% for those with 5 or more factors [8]. A modern analysis of patients treated between 1980 and 2010 demonstrated that the IPS is still prognostic but the range of outcomes is narrowed, with 5-year freedom-from-progression rates ranging from 88% to 62% [9]. Alternate prognostic systems may offer simpler and more predictive models using modern-era patients [10]. The IPS itself does not include all potentially important adverse prognostic features on presentation. For example, it does not include bulky disease. Other prognostic factors that may be important include B symptoms, histologic subtype, number of nodal sites of disease, number of organs involved, bone marrow involvement, erythrocyte sedimentation rate, platelet count, serum alkaline phosphatase, serum lactate dehydrogenase, and comorbidities [11-16]. With the increasing use of PET for the staging and restaging for lymphoma patients, PET-detected early response to chemotherapy has been identified as a powerful prognostic tool in chHL. In a study by Gallamini et al [17] that included 260 patients with stage II-IVB disease, the 2-year progression-free survival (PFS) rates for patients with positive versus negative PET results after 2 cycles of chemotherapy (PET-2) were 12.8% and 95.0%, respectively (P<0.0001). In multivariable regression analysis that included PET-2 status and IPS as a continuous variable, IPS lost its prognostic value, and only PET-2 status had significant independent prognostic value for PFS (hazard ratio [HR], 38.3; P<0.0001). Although this study arguably catalyzed a generation of response-adapted trials, it is important to note that PET interpretation in this study preceded modern consensus criteria and subsequent studies have not shown as wide a separation of PET-positive and PET-negative groups. Nevertheless, trials have been designed investigating the role of response-adapted therapy based on early PET findings (discussed below).

Assessment of Response

Response assessment in chHL can be challenging. It is common to note a residual mass on a CT scan, especially in a patient initially presenting with bulky disease. It is important to note that the definition of response has varied from trial to trial, but the 2007 Cheson response criteria have standardized response definitions in the PET era for all malignant lymphomas [18]. PET imaging forms the basis of the response criteria, eliminating the complete response/unconfirmed (CRu) category. For a CR, residual masses of any size are permitted if the PET scan is negative. Although the CRu category has been removed from response criteria, only moderate reproducibility of PET scan interpretation among nuclear medicine experts suggests that uncertainty regarding response has not been eliminated but rather has shifted from one modality (CT) to another (PET) [19]. A numeric scale of metabolic responses using the Deauville criteria has gained traction as a tool in some research studies and other published guidelines [4,20,21]. A 2014 update of the Cheson criteria incorporates the Deauville criteria (also referred to as the “5-point scale”) and states that PET/CT should be the standard of care for remission assessment in fluorine-18-2-fluoro-2-deoxy-D-glucose (FDG)–avid lymphomas, including HL. Although the size of anatomic residual disease may carry prognostic significance (for example, in the German Hodgkin Study Group [GHSG], ACR Appropriateness Criteria® 2 Hodgkin Lymphoma — Stage III and IV.
HD12 [22], residual disease was associated with more relapse), PET-based response assessment at the end of chemotherapy is an undeniably valuable tool with a high negative predictive value [23] (see Variant 1). Interim PET assessment during a course of chemotherapy is an intriguing tool to develop “response”-adapted rather than just “risk”-adapted treatment strategies [17]. In another study of 304 patients with HL undergoing ABVD chemotherapy, PET scans performed after 2 cycles were prognostic for sustained remission. In the advanced-stage patients, the 9-year PFS was 88.6% in the patients with negative interim PET scans versus 28.7% in those with positive PET scans [24]. The best way to use interim PET scans remains a subject of intense research, both regarding the correct PET parameters to use [25] and how to intensify/deintensify therapy as discussed below (ongoing GHSG HD18).

**Treatment Options Overview**

**Chemotherapy Regimens**

Combination chemotherapy has been the mainstay of treatment since the 1970s after the National Cancer Institute reported results with the mechlorethamine, vincristine, procarbazine, and prednisone (MOPP) regimen [26]. Since that seminal report, several approaches have been attempted, including 1) substitution of some of the chemotherapeutic agents while maintaining a similar regimen; 2) use of entirely new regimens; 3) use of non–cross-resistant agents in combination with MOPP, such as the MOPP/doxorubicin, bleomycin, and vinblastine (ABV) hybrid; 4) use of a dose-dense regimen with or without radiation, such as dose-escalated BEACOPP or Stanford V; and 5) use of a time-condensed regimen such as BEACOPP-14, where baseline-dose BEACOPP is given every 2 weeks rather than every 3 weeks with granulocyte colony-stimulating factor support (see Variant 2 and Variant 3).

Unlike with most other advanced-stage malignancies, chemotherapy can cure the majority of cHL patients. Following the seminal paper by Devita et al [26], several randomized trials showed that ABVD or ABVD-containing regimens were significantly superior to MOPP [11,14,27]. The MOPP/ABV hybrid was subsequently found to be more toxic and to offer no significant improvement in outcomes compared to ABVD [28]. As a result, ABVD became the most widely used regimen in the United States and offers cure in approximately 85%–88% of patients, with a favorable short-term and long-term toxicity profile.

The dose-escalated BEACOPP regimen, developed by the GHSG, yielded significantly better survival compared with conventional-dose regimens in the GHSG HD9 study [29]. In this 3-arm trial, patients with stage IIB-IV HL were randomized to 8 cycles of cyclophosphamide, vincristine, procarbazine, and prednisone alternating with ABVD (COPP-ABVD), baseline-dose BEACOPP, or dose-escalated BEACOPP. In the update of the study with 1196 evaluable patients and a median follow-up of 111 months, the 10-year freedom-from-relapse rate was significantly higher in the dose-escalated BEACOPP arm than in the baseline BEACOPP and COPP-ABVD arms (82%, 70%, and 64%, respectively; P<0.0001) [30]. The corresponding 10-year overall survival (OS) rates were 86%, 80%, and 75%, respectively (P=0.0005). However, in the standard arm of this trial, COPP-ABVD was used rather than standard ABVD. Also, patients on the dose-escalated BEACOPP arm had a significantly higher 10-year cumulative incidence of acute myelogenous leukemia/myelodysplasia (3.2%, 2.2%, and 0.4%, respectively; P=0.03). In subsequent trials, the GHSG has employed dose-escalated BEACOPP as the standard comparator arm.

To reduce treatment-related toxicity, the GHSG explored the efficacy of modified BEACOPP as well as the value of RT in advanced-stage HL. The GHSG HD12 trial was a 4-arm study of 1670 patients with bulky stage IIB and stages III-IV disease comparing 8 cycles of dose-escalated BEACOPP versus 4 cycles of escalated and 4 cycles of baseline BEACOPP [22]. There was no significant difference between the chemotherapy arms with regard to either 5-year freedom-from-treatment failure (86.4% versus 84.8%) or 5-year OS rate (92% versus 90.3%). There was concern that the less intensive chemotherapy arm resulted in more disease progression (1.1% versus 3.3%, P=0.006) without any improvement in treatment-related deaths (2.4% versus 3.4%). The value of RT in this study is discussed below.

The GHSG HD15 trial compared 3 chemotherapy regimens (8 cycles versus 6 cycles of dose-escalated BEACOPP versus 8 cycles of BEACOPP-14) [31] and was the first large-scale prospective trial to incorporate a PET-adapted approach. The BEACOPP-14 regimen, a time-intensified variant of standard-dose BEACOPP administered every 2 weeks rather than 3 weeks with growth factor support for 8 cycles, showed no acute leukemia or myelodysplasia based on a median follow-up of 34 months [32]. Importantly, 30 Gy was given only to subjects with 2.5-cm residual masses that were PET positive. This reduced the overall percentage of patients
receiving additional RT from 71% in their prior HD9 trial to 11% in the HD15 trial. There was high adherence to the planned RT, but patients with a positive end-of-treatment PET retained an inferior PFS (86.2% versus 92.6%) compared to patients with a negative end-of-treatment PET, irrespective of the residual tumor size. However, it is important to note that despite central PET review, this study preceded the standardization of PET interpretation that is now considered standard. There was improved survival at 5 years in subjects treated with only 6 cycles of dose-escalated BEACOPP compared to 8 cycles (91.9% versus 95.3%). This difference was mostly due to treatment-related deaths and second malignancies. Of note, BEACOPP-14 had an intermediate 5-year OS between the other 2 regimens tested (94.8%) and was statistically noninferior to 8 cycles of dose-escalated BEACOPP [23].

Variations of the BEACOPP regimen have also been investigated by groups other than the GHSG. Federico et al [33] in Italy initially reported that 4 escalated and 2 standard doses of BEACOPP resulted in superior PFS compared with ABVD for patients with an IPS ≥ 4 (5-year rates, 81% and 68%, respectively; \( P = 0.038 \)). However, BEACOPP did not significantly improve 5-year OS rates. In the long-term follow-up analysis the PFS benefit of BEACOPP was lost and 10-year OS rates were identical at 84% [34]. The Gruppo Italiano Terapie Innovative nei Linfomi (GITIL) and Intergruppo Italiano Linfomi (IIL) cooperative groups compared 6–8 cycles of ABVD versus 4 cycles of dose-escalated followed by 4 cycles of baseline BEACOPP as first-line therapy, with preplanned high-dose therapy as salvage [35]. Salvage therapy with reinduction/high-dose autologous transplant was given to patients with <80% response or with relapse. RT was given to any previously bulky or residual masses. The BEACOPP variant was associated with a significantly higher 7-year rate of freedom from first progression compared to ABVD (85% versus 73%, \( P = 0.004 \)); however, significant differences were not observed for freedom from second progression (88% versus 82%, \( P = 0.15 \)) or OS (89% versus 84%, \( P = 0.39 \)). The authors argued that when proper salvage treatment is used, long-term disease outcomes are near equivalent with ABVD, and thus overtreatment with BEACOPP can be avoided.

The European Organization for Research and Treatment of Cancer (EORTC) protocol 20012 compared 4 cycles of escalated-dose BEACOPP followed by 4 cycles of standard-dose BEACOPP to 8 cycles of ABVD in high-risk stage III/IV patients (IPS ≥3). No RT was used in either arm. Interim results showed superior 4-year PFS with BEACOPP (72.8% versus 83.4%, \( P = 0.005 \)) but no significant difference in 4-year OS rates (86.7 versus 90.3, \( P = 0.208 \)) [36]. We await long-term follow-up to determine if the EORTC 20012 results will continue to mimic the GITIL/IIL study.

Stanford V is a 12-week, 7-drug regimen that is administered on a weekly basis. Promising phase II results have been reported, with a low risk of pulmonary toxicity, infertility, leukemia, or myelodysplasia [5,7,37]. Patients with initial disease measuring ≥5 cm and/or macroscopic splenic disease receive 36 Gy of IFRT starting 2 weeks after chemotherapy completion. Compared to 6 cycles of ABVD, the cumulative chemotherapy doses are significantly less with Stanford V for adriamycin (300 versus 150 mg/m²) and bleomycin (120 versus 30 IU/m²). The importance of RT as part of Stanford V was highlighted by a multi-institutional randomized trial from Italy [38] comparing modified Stanford V, mechlorethamine, vincristine, procarbazine, prednisone, epidoxorubicin, bleomycin, vinblastine, lomustine, doxorubicin, and vindesine (MEC hybrid) with ABVD. In this study, RT was limited to patients with initially bulky disease (>6 cm) or with PR to chemotherapy and those with only up to 2 sites of disease. The 5-year relapse-free survival rates of patients on the modified Stanford V arm were significantly lower than those of patients treated with MEC hybrid or ABVD (73%, 92%, and 88%; \( P < 0.01 \)). The inferior outcome with the modified Stanford V was attributed to the fact that RT was not given according to the Stanford V guidelines, such that only 76% of patients with a CR to chemotherapy received RT and sometimes delayed beyond the 2 weeks prescribed by the original regimen. In addition, the results were difficult to interpret because the response evaluations that determined if patients would continue on a study arm were performed at different times: after 8 and 12 weeks of modified Stanford V, 16 weeks of ABVD, and 24 weeks of MEC hybrid. A phase III study from the United Kingdom National Cancer Research Institute Lymphoma Group [39] suggests that Stanford V is as effective as ABVD when RT is given as specified by the Stanford group.

The Eastern Cooperative Oncology Group (ECOG) E2496 Intergroup phase III trial was a pivotal trial comparing failure-free survival between 6–8 cycles of ABVD with the Stanford V regimen in 794 eligible patients [40]. RT in this trial was administered in the ABVD arm only to patients with bulky mediastinal disease (>1/3 mediastinal mass ratio or 10 cm), whereas on Stanford V, RT was administered to nodal sites ≥5 cm and macroscopic splenic disease if present. With a median follow-up of 6.4 years, there was no significant difference in the 5-year failure-free survival between ABVD and Stanford V (74% versus 71%, \( P = 0.32 \)). Similarly, OS was identical at 88% for
both arms. Importantly, this study preceded the era of routine PET imaging. Nevertheless, the authors conclude equivalency between the 2 arms and propose that, given the increased lymphopenia and increased neuropathy associated with Stanford V, ABVD for 6–8 cycles plus RT to bulky mediastinal disease remains the standard of care. Currently, ABVD is rarely given for a full 8 cycles, and 6 cycles is preferred.

The debate continues regarding the optimal choice of systemic therapy for advanced-stage HL [4,30]. Escalated BEACOPP has unarguably better PFS compared with ABVD; however, when long-term end points like OS are considered, the difference narrows. To date, there still has not been a study directly comparing 6 cycles of escalated BEACOPP (best arm from the GHSG HD15 study) and optimized ABVD with a preplanned salvage strategy (from the GITIL/IIL study). Many physicians and patients struggle with the balance between aggressively curing the disease with frontline therapy, while accepting the considerable toxicities, and more gentle initial therapy, while accepting that salvage treatment will be not only toxic but less effective. Stanford V is also an option, but extensive use of RT is part of the package that needs to be accepted up front, something that is not always palatable to young patients. New approaches such as the addition of brentuximab to chemotherapy in treatment-naïve patients is being studied [41]. The increasing utility of response-adapted therapy using functional imaging should help balance the risk-benefit profile for future patients.

The Role of Radiation Therapy

The role of RT in advanced-stage HL is controversial because most early reports used chemotherapeutic regimens such as MOPP or MOPP hybrids, relatively large RT fields (mantle fields or subtotal nodal irradiation), and limited response-assessment tools. Ferme et al [42] addressed the question of whether subtotal nodal irradiation could substitute for additional cycles of chemotherapy. In this trial, patients were randomized to 6 cycles of MOPP/ABV or ABVPP (doxorubicin, bleomycin, vinblastine, procarbazine, and prednisone). If patients achieved a CR or a PR of at least 75% after 6 cycles, they were randomized to either consolidation with 2 additional cycles of chemotherapy or subtotal nodal irradiation to 36 Gy. No differences in 10-year disease-free survival (73% versus 78%, \( P=0.07 \)) or OS rate (84% versus 79%, \( P=0.29 \)) were detected. Without a clear benefit and very large subtotal nodal irradiation RT fields, the RT arm did not become the new reference arm for the EORTC.

To test the impact of RT after effective chemotherapy, the EORTC randomized 421 advanced-stage HL patients in CR after MOPP-ABV chemotherapy to observation or IFRT (16–24 Gy) [43]. This study failed to demonstrate a benefit to consolidative RT if a CR was obtained after chemotherapy. The authors separately reported results in 227 patients who failed to achieve a CR and received consolidative IFRT [44]. The 8-year event-free survival and OS rates were statistically no different than those for patients who achieved a CR to chemotherapy, which had suggested a potential role for RT after a PR to chemotherapy.

Johnson et al [45] analyzed the outcomes of nonrandomized consolidative IFRT given after chemotherapy in the initial treatment of advanced HL. At least 30 Gy of IFRT was delivered to residual masses or sites of originally bulky disease. With a median follow-up of 6.9 years, significantly higher PFS rates (71% versus 86%, \( P<0.0001 \)) and OS rates (87% versus 93%, \( P=0.014 \)) were found in patients who received RT. As in most other studies, PET was not used to assess response.

The GHSG HD12 trial randomized adult patients to 30 Gy of RT to initial sites of disease >5 cm and residual disease versus observation after BEACOPP (8 cycles of escalated dose or 4 cycles of escalated dose plus 4 cycles of standard dose) [22]. The use of RT in this study was determined by CT-based responses. At 5 years, there was a benefit to RT in terms of freedom-from-treatment failure (90.4% versus 87%, \( P=0.08 \)) but no OS advantage (HR, 1.09; 95% confidence interval [CI], 0.74 to 1.60). In subgroup analyses, they found that omission of RT in patients with CT-based residual disease significantly worsened freedom-from-treatment failure by 5.8% at 5 years, whereas the outcomes for patients in CR with initial bulky disease were not significantly different (1.1% difference). Of note, 22.9% of patients who were not assigned to undergo RT received it either due to prespecified reasons or against panel recommendations. This may have resulted in an underestimate of the impact of RT in patients with residual disease. The authors concluded that RT should not be routinely omitted in patients with CT-based residual disease, but RT can be omitted if they are in CR, even with initially bulky disease.

As part of the GHSG HD15 trial for advanced-stage HL [23], RT was only given to masses >2.5 cm that were PET positive. In this study that used a backbone of 3 different BEACOPP-based regimens, the 12-month negative predictive value of PET/CT was 94.1%. The restrictions on RT in this study, combined with the efficacy of the intense chemotherapy, resulted in having only 11% of all subjects receiving RT, much lower than in the non-PET approach used in HD9 (70% of subjects received RT). The 4-year PFS of PET-positive subjects (86.2%) was
significantly worse than PET-negative subjects (92.6%), despite the protocol-prescribed use of RT in the PET-positive subjects (95% CI for difference, 0.9–12.0; log-rank $P=0.022$). These PET-positive subjects did fare better than other PET-positive subjects reported in other series where treatment was not adjusted according to the PET findings [17], suggesting that the role of RT in patients with PET-residual disease after chemotherapy deserves further study.

To summarize, most of the modern data agree that consolidative RT does not confer a benefit when patients are in CR after chemotherapy, even if initially bulky [22,23,42,43]. In contrast, the practice of adding RT to residual masses is supported by data that are either randomized [22] or from single arms based on good outcomes compared to CR patients when RT was part of the strategy [44,45]. The response-adapted use of PET/CT to guide RT has been demonstrated in the GHSG HD15, where the negative predictive value of PET/CT was excellent even when a residual mass was present. In particular, the patients with PET-negative residual masses >2.5 cm treated with BEACOPP alone on this trial fared just as well as patients in anatomic CR [23]. There currently are not good data on the predictive value of PET in advanced-stage patients after multiple cycles of ABVD, and it is unclear whether results from BEACOPP-based trials can be extrapolated.

**Response-Adapted Strategies**

Interim PET response-adapted strategies are being tested in the next generation of advanced-stage HL studies. An intergroup study in the United States led by the Southwest Oncology Group (S0816) is a phase II trial randomizing patients after 2 cycles of ABVD; patients with a negative interim PET proceed to 4 additional cycles of ABVD, whereas patients with a positive interim PET have their treatment changed to escalated BEACOPP [46]. A second intergroup study, still ongoing, is testing a similar approach in bulky cHL using interim PET assessment as the determinant of 2 chemotherapy approaches; there is no RT in this trial [47]. Both intergroup studies have central PET assessment using the now-standard Deauville criteria. The ongoing GHSG HD18 trial was also designed to evaluate the role of intensifying or reducing systemic therapy based on early PET results and also evaluates the potential role of the anti-CD20 monoclonal antibody, rituximab, in cHL. After 2 cycles of dose-escalated BEACOPP, PET-negative patients are randomized to either 6 more cycles of dose-escalated BEACOPP or only 2 more cycles. PET-positive patients are randomized to 6 more cycles of dose-escalated BEACOPP with or without rituximab. An interim futility analysis of this study has shown no benefit of adding rituximab [48]. The GITIL is currently conducting a phase II trial involving PET-adapted chemotherapy in patients with previously untreated stage IIB-IVB HL (GITIL-HD0607). On this trial, patients at first receive 2 cycles of ABVD chemotherapy. Individuals with PET-2–positive disease, as determined by a PET Reviewing Committee, are randomized to undergo 4 cycles of escalated-dose BEACOPP or rituximab-BEACOPP chemotherapy. In contrast, patients with PET-2–negative disease receive 4 additional cycles of ABVD. PET-2–negative patients with initially bulky disease are then randomized to IFRT to 30 Gy after 6 cycles of ABVD or no further treatment. The Fondazione Italiana Linfomi has opened the HD0801, which also applies a response-adapted strategy after initial 2 cycles of ABVD. We await the results of these studies to best understand how to employ the power of interim PET responses.

**Radiation Treatment Planning With Involved-Site Radiation Therapy Paradigm**

Recently, a new set of field designs, involved-site RT (ISRT), has been developed and endorsed by the steering committee of the International Lymphoma Radiation Oncology Group [49]. Led by experienced radiation oncologists specializing in lymphoma who initially organized the standardization of IFRT fields in the 2-D era a decade ago, the ISRT fields are a “modernized” version of IFRT. These new field designs were developed to take into consideration modern technology, including the use of staging PET/CT scans, 3-D and 4-D treatment planning with CT scanners, conformal treatment techniques, and the use of image guidance to replace the antiquated IFRT that was based on 2-D treatment planning and bony anatomy. These fields are expected to be somewhat smaller than the traditional IFRT but larger than involved-node RT (INRT) for patients who do not have adequate imaging necessary for INRT treatment planning. In the context of advanced-stage HL, the gross tumor volume is the residual mass after chemotherapy. Margins are added for uncertainties and motion depending on the location in the body, such as the thorax and/or upper abdomen. Volume-based planning techniques used to avoid critical structures are similar to limited-stage HL but may take into account the additional cumulative chemotherapy typically used in advanced-stage HL.

In conclusion, the use of RT in advanced-stage cHL depends on the context of the chemotherapy used and the final response assessment; interim response assessment may also have a future role in determining treatment. RT is not recommended for patients with nonbulky (<5 cm) stage III-IV HL who achieve a CR with full-course
chemotherapy [42,43]. In the context of dose-escalated BEACOPP, RT is only required in PET-avid residual masses >2.5 cm [23]. In the context of ABVD chemotherapy as used in the GITIL/IIL, there are clinical data that support RT of initial bulky lymphoma to 25.2 Gy and areas of residual disease to 30.6 Gy [35]. For initially bulky sites [45,50] or if there are residual PET-positive site(s) after chemotherapy [44], RT is an option after 6 cycles of ABVD chemotherapy (see Variant 4 and Variant 5). Radiation doses of 30–36 Gy are usually appropriate after ABVD or BEACOPP [45,51]. After chemotherapy with the Stanford V regimen, IFRT to 36 Gy should be given to all sites that initially are ≥5 cm and to macroscopic splenic disease [6] (see Variant 1 and Variant 4).

**Summary of Recommendations**

- ABVD × 6 cycles every 4 weeks is the most commonly accepted chemotherapy regimen for advanced-stage HL in North America.
- Other chemotherapy regimens, including Stanford V (which incorporates IFRT to 36 Gy to initial sites of disease measuring ≥5 cm and macroscopic splenic disease) and dose-escalated BEACOPP × 6 cycles, can be considered. RT depends on the type of chemotherapy used and the final response assessment.
- Functional imaging is likely to have an important impact on future standard approaches in cHL.

**Summary of Evidence**

Of the 51 references cited in the ACR Appropriateness Criteria® Hodgkin Lymphoma — Stage III and IV document, 37 are categorized as therapeutic references including 23 well designed studies, 7 good quality studies, and 1 quality study that may have design limitations. Additionally, 14 references are categorized as diagnostic references including 2 well designed studies and 2 quality studies that may have design limitations. There are 16 references that may not be useful as primary evidence.

The 51 references cited in the ACR Appropriateness Criteria® Hodgkin Lymphoma — Stage III and IV document were published from 1970-2010. While there are references that report on studies with design limitations, 32 well designed or good quality studies provide good evidence.

**Supporting Documents**

For additional information on the Appropriateness Criteria methodology and other supporting documents go to www.acr.org/ac.

**References**


The ACR Committee on Appropriateness Criteria and its expert panels have developed criteria for determining appropriate imaging examinations for diagnosis and treatment of specified medical condition(s). These criteria are intended to guide radiologists, radiation oncologists and referring physicians in making decisions regarding radiologic imaging and treatment. Generally, the complexity and severity of a patient’s clinical condition should dictate the selection of appropriate imaging procedures or treatments. Only those examinations generally used for evaluation of the patient’s condition are ranked. Other imaging studies necessary to evaluate other co-existent diseases or other medical consequences of this condition are not considered in this document. The availability of equipment or personnel may influence the selection of appropriate imaging procedures or treatments. Imaging techniques classified as investigational by the FDA have not been considered in developing these criteria; however, study of new equipment and applications should be encouraged. The ultimate decision regarding the appropriateness of any specific radiologic examination or treatment must be made by the referring physician and radiologist in light of all the circumstances presented in an individual examination.
**Clinical Condition:** Hodgkin Lymphoma — Stage III and IV

**Variant 1:** 25-year-old woman with cHL: clinical stage (CS) IIIA, with mediastinal, splenic, and para-aortic involvement, all sites nonbulky (<5 cm), and IPS <3; completed 6 cycles of ABVD, followed by restaging PET/CT showing a CR (Deauville 2) and residual mediastinal mass of 2 cm.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>No RT</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>20–29 Gy ISRT</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>30–36 Gy ISRT</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

**Rating Scale:** 1,2,3 Usually not appropriate; 4,5,6 May be appropriate; 7,8,9 Usually appropriate

**Variant 2:** 25-year-old woman with cHL: CS IIIA, with mediastinal, splenic, and para-aortic involvement, all sites nonbulky (<5 cm), and IPS <3.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABVD × 6 cycles (every 4 weeks)</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Stanford V with 36 Gy to macroscopic splenic disease</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Dose-escalated BEACOPP × 4 cycles and baseline BEACOPP × 4 cycles (every 3 weeks)</td>
<td>5</td>
<td>This option may be appropriate but there was disagreement among panel members on the appropriateness rating as defined by the panel’s median rating. This arm was tested in HD12 against dose-escalated BEACOPP × 8 with no better results (and maybe worse PFS).</td>
</tr>
<tr>
<td>Dose-escalated BEACOPP × 6 cycles (every 3 weeks)</td>
<td>7</td>
<td>See HD15 results (Engert 2012) [23].</td>
</tr>
<tr>
<td>Dose-escalated BEACOPP × 8 cycles (every 3 weeks)</td>
<td>3</td>
<td>See HD15 results (Engert 2012) [23].</td>
</tr>
<tr>
<td>BEACOPP-14 (standard) × 8 cycles (every 2 weeks)</td>
<td>5</td>
<td>This option may be appropriate but there was disagreement among panel members on the appropriateness rating as defined by the panel’s median rating. See HD15 results (Engert 2012) [23].</td>
</tr>
</tbody>
</table>

**Rating Scale:** 1,2,3 Usually not appropriate; 4,5,6 May be appropriate; 7,8,9 Usually appropriate
Clinical Condition: Hodgkin Lymphoma — Stage III and IV

Variant 3: 52-year-old man with cHL: CS IVB, with bulky (>10 cm) mediastinal, splenic, para-aortic, and bone marrow involvement, albumin <4 g/dL, hemoglobin <10.5 g/dL, and IPS 5.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type of Chemotherapy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABVD × 6 cycles (every 4 weeks)</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Stanford V with 36 Gy to sites ≥5 cm and macroscopic splenic disease</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Dose-escalated BEACOPP × 4 cycles and baseline BEACOPP × 4 cycles (every 3 weeks)</td>
<td>5</td>
<td>This option may be appropriate but there was disagreement among panel members on the appropriateness rating as defined by the panel’s median rating. This arm was tested in HD12 against dose-escalated BEACOPP × 8 with no better results (and maybe worse PFS).</td>
</tr>
<tr>
<td>Dose-escalated BEACOPP × 6 cycles (every 3 weeks)</td>
<td>7</td>
<td>See HD15 results (Engert 2012) [23].</td>
</tr>
<tr>
<td>Dose-escalated BEACOPP × 8 cycles (every 3 weeks)</td>
<td>5</td>
<td>This option may be appropriate but there was disagreement among panel members on the appropriateness rating as defined by the panel’s median rating. See HD15 results (Engert 2012) [23].</td>
</tr>
<tr>
<td>BEACOPP-14 (standard) × 8 cycles (every 2 weeks)</td>
<td>5</td>
<td>This option may be appropriate but there was disagreement among panel members on the appropriateness rating as defined by the panel’s median rating. See HD15 results (Engert 2012) [23].</td>
</tr>
</tbody>
</table>

**Rating Scale:** 1,2,3 Usually not appropriate; 4,5,6 May be appropriate; 7,8,9 Usually appropriate

Variant 4: 52-year-old man with cHL: CS IVB, with bulky (>10 cm) mediastinal, splenic, para-aortic (<5 cm), and bone marrow involvement, albumin <4 g/dL, and hemoglobin <10.5 g/dL; 50% reduction in mediastinal mass with residual FDG avidity limited to the mediastinum after 2 cycles of ABVD (Deauville 4); restaging PET/CT after 4 cycles of ABVD showed a complete metabolic response (Deauville 2); completed 6 cycles of ABVD. Final PET/CT assessment shows a 4-cm residual mediastinal mass, Deauville 2.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Additional Therapy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No further therapy</td>
<td>5</td>
<td>There are limited data on no additional therapy.</td>
</tr>
<tr>
<td>Radiation therapy</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>2 additional cycles of ABVD</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>2 additional cycles of ABVD followed by RT</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td><strong>Radiation Therapy (if given)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20–29 Gy to all initial sites followed by a boost to the mediastinum to a total dose of 30–36 Gy</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>20–29 Gy to mediastinum only</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>30–36 Gy to mediastinum only</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

**Rating Scale:** 1,2,3 Usually not appropriate; 4,5,6 May be appropriate; 7,8,9 Usually appropriate
Clinical Condition: Hodgkin Lymphoma — Stage III and IV

Variant 5: 52-year-old man with cHL: CS IVB, with bulky (>10 cm) mediastinal, splenic, para-aortic (<5 cm), and bone marrow involvement, albumin <4 g/dL, and hemoglobin <10.5 g/dL; 50% reduction in mediastinal mass with residual FDG avidity limited to the mediastinum after 2 cycles of ABVD (Deauville 4); restaging PET/CT after 4 cycles of ABVD showed a partial metabolic response (Deauville 4). After 6 cycles of ABVD, the residual mediastinal mass is 4 cm and final PET/CT assessment still Deauville 4 only in mediastinum. All other sites are <2.5 cm.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additional Therapy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No further therapy</td>
<td>2</td>
<td>This option needs a biopsy.</td>
</tr>
<tr>
<td>Radiation therapy</td>
<td>6</td>
<td>This option is used only if biopsy is not possible or if biopsy is negative.</td>
</tr>
<tr>
<td>2 additional cycles of ABVD</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>2 additional cycles of ABVD followed by RT</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Rebiopsy, if positive for HL, salvage chemotherapy and autologous transplant</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

| Radiation Therapy (if given)           |        |                                                    |
| 20–29 Gy to all initial sites followed by a boost to the mediastinum | 5      | This option may be appropriate but there was disagreement among panel members on the appropriateness rating as defined by the panel’s median rating. |
| 20–29 Gy to mediastinum only           | 4      |                                                    |
| 30–36 Gy to mediastinum only           | 7      |                                                    |
| 30–36 Gy to mediastinum only with boost to PET-avid area | 8      | Boost would be >36 Gy.                           |
| >36 Gy to mediastinum only             | 6      | This option is used if PET is still positive after transplant. |

Rating Scale: 1,2,3 Usually not appropriate; 4,5,6 May be appropriate; 7,8,9 Usually appropriate