Risk, Cure and Complications in Advanced Hodgkin Disease

Sandra J. Horning

Stanford University, Stanford, CA

Current therapy for Hodgkin disease is aimed at high cure rates and optimal survivorship. Although intensified chemotherapy with the escalated BEACOPP regimen resulted in higher rates of cure and survival compared with COPP/ABVD in the high-profile HD9 randomized controlled trial (RCT), this regimen has not been universally adopted by patients and physicians due to the attendant increased risks of early and late complications. Appropriately, questions emerge as to whether the results of this trial should be interpreted as establishing the superiority of BEACOPP over the current standard ABVD therapy, and whether clinical or biologic prognostic factors can better select patients for more intensive treatment. The widespread availability and high predictive accuracy of functional imaging with PET scans has led to promising, preliminary studies assessing early response to therapy with this diagnostic tool. In this review, the characteristics and outcomes of patients treated with ABVD in RCT will be made and compared with COPP/ABVD in HD9; clinical and biologic prognostic factors will be discussed, including PET imaging; and newer strategies targeted at minimizing treatment complications while maximizing cure rates will be discussed. Although enthusiasm for PET imaging is great, the challenges for using this diagnostic tool for risk-adapted therapies are substantial. Importantly, physicians and patients should be aware of these challenges, support the RCT that seek to address them, and carefully weigh risks and benefits for individual patients.

Introduction

The treatment of Hodgkin disease (HD) at any stage of presentation is highly successful, and survival statistics demonstrate significant gains from 1996-2002 compared with 1984-1986. Over the last three decades, management has evolved from a heavy reliance on extended field irradiation supplemented with alkylating agent-based chemotherapy to anthracycline-based chemotherapy with selective consolidation with involved field irradiation, often at a lower dose. In the past decade, superior outcomes in early-stage disease have been reported in individual clinical trials with combined modality therapy versus chemotherapy alone in early-stage disease and with more intensive chemotherapy in advanced-stage disease, but at the cost of an increased risk for early and late morbidities. These experiences have resulted in a tension between the desire to cure as many patients as possible with primary therapy and the desire to render survivors with as few early and late adverse effects as possible. Additional advances within the past decade have included the study and use of positron emission tomography (PET) scans as more accurate predictors of disease status in HD and characterization of HD on the cellular and molecular levels. There is, naturally, a great desire to bring this new knowledge to bear in the clinical management and to specifically address the noted tension.

From this background, three questions emerge:

1. Is it appropriate to extrapolate the results of COPP/ABVD in the German Hodgkin Lymphoma Study Group (GHSG) HD9 trial to conclude that BEACOPP (cyclophosphamide, adriamycin, etoposide, vincristine, bleomycin, procarbazine, prednisone) is superior to ABVD (adriamycin, bleomycin, vincristine, dacarbazine)?
2. Can we risk-adapt, that is, distinguish clinical or biologic prognostic features to identify patients who will benefit or will be harmed from more intensive treatment?
3. Can newer therapeutic strategies result in both higher survival rates and better survivorship?

ABVD in Randomized Controlled Trials

The ABVD combination, introduced in the 1970s by Bonadonna and colleagues, became the standard of care in HD based on randomized clinical trials conducted by the U.S. cooperative groups. The more recent Intergroup study reported by Duggan et al had no upper age limit, and older age had a major impact on efficacy and toxicity. In this study, 36% of patients were older than 40 years (Table 1). Only stage III-IV patients were eligible but patients relapsing after prior radiation therapy were included. Based on the standards of the time, treatment was delivered with dose modifications for peripheral blood counts, and growth factors were not to be used to maintain optimal dose and schedule. The failure-free survival (FFS) at 5 years was 63% and overall survival (OS) was 81% in the ABVD arm (n = 412) (Table 2). FFS was inferior for patients who were older, had stage IV disease (45% of the population) and had received prior radiation. No radiotherapy was planned in this study. Of note, 75% of the deaths occurred among those older than 55 years. Retrospective application of the Inter-
national Prognostic Score (IPS), which was less favorably distributed than expected (Table 1), did not correlate with FFS. Although efficacy was similar to the MOPP (mechloroethamine, vincristine, procarbazine, prednisone, doxorubicin, bleomycin, vinblastine)/ABV combination, this randomized controlled trial (RCT) was closed prematurely due to an increased risk of death and second malignancy with MOPP/ABV.11

Gobbi and colleagues tested ABVD in a three-arm Italian study involving 122 ABVD patients with stage IIB-IV disease.11,12 In this study, the age cap was 65 years (although the upper age boundary of treated patients was 68 years) and the median age was 31 years, with 23% patients older than 45 years. Radiation therapy was delivered to patients with bulky disease or to a maximum of two sites of persistent radiographic disease for partial response (PR) or uncertain complete response (CRu) patients. Six cycles of chemotherapy were given and were supported by growth factor for severe neutropenia. The IPS distribution was as expected despite the fact that 34% of patients had stage IIB disease. Stage IV disease was reported in 24%. At 5 years, the freedom from progression was 85% and the OS was 90%. Only 3% of patients experienced disease progression. In this study, an unusual definition of FFS was used, in which PR was scored as treatment failure. However, using that definition, the FFS was 78% at 5 years.

The ABVD regimen was tested by Johnson and colleagues as the standard arm in randomized comparisons with multidrug combinations (ChlVPP/PABlOE, ChlVPP/EVA).13 Adults of all ages were eligible, such that the upper age boundary was 77 years and the median age was 35 years. Patients with advanced disease and those with early-stage disease requiring systemic chemotherapy (based on risk factors) were eligible. Overall, 19% patients had stage IV disease. Patients received 6 to 8 cycles of chemotherapy, supported by growth factors to maintain optimal dosing, and radiation was delivered to sites of bulky or residual disease. Of note, relative dose intensity was more than 80%, and 44% used growth factors. At 3 years, the estimated event-free survival was 75% for the 406 ABVD-treated patients, and the OS was 90%. Outcomes were inferior for patients older than 45 years with the multidrug combinations (in agreement with the U.S. Intergroup results) compared to ABVD, whereas no significant differences were seen in the overall study according to IPS. In this study, ABVD emerged as the preferred regimen based on superior tolerance.

With current anti-emetics and vascular access, ABVD is considered a well-tolerated regimen. In the U.S. Intergroup trial, pulmonary toxicity as defined by protocol occurred in 30% of ABVD patients, with an increased incidence in those who were older and had received prior radiation. A small number of fatal events related to bleomycin pulmonary toxicity has been reported in multiple studies; the incidence of this in the Intergroup study was 2%. Fatal respiratory toxicity in the UK study was reported as 1.5%. Remarkably, no cases of acute leukemia were recorded among any of the patients treated with ABVD alone in the Intergroup study, the UK study or the Italian study, consistent with prior reports that ABVD does not increase the risk of subsequent leukemia. Early data suggested that fertility among both men and women is maintained with ABVD,
although this has generally been an extrapolation from semen analyses and reports of menstrual function, alone or combined with hormone measurements. Recently, Hodgson et al reported that 70% of women (n = 36) treated with ABVD were able to conceive within 12 months after 4 to 6 treatment cycles compared with 75% of controls.14

BEACOPP in Randomized Controlled Trials
Diehl and colleagues introduced standard and escalated BEACOPP based upon a mathematical model indicating the potential for an increased cure rate with dose intensification.2 BEACOPP combines etoposide with traditional cytotoxic agents, cyclophosphamide and procarbazine, as well as doxorubicin. Growth factors are integral to the delivery of escalated BEACOPP, which features major intensification of cyclophosphamide and etoposide. The HD9 study, in which standard and escalated BEACOPP were compared with COPP/ABVD, has been presented, published, and updated on several occasions.2,15 Eligible patients were 15 to 65 years old with stage IIB-IIIA disease with risk factors or with IIB-IV disease. Patients received 8 cycles of chemotherapy followed by 36 Gy radiotherapy to sites of bulky or residual disease. Table 3 tabulates data from the HD9 study comparing outcomes of standard and escalated BEACOPP with COPP/ABVD. Given the caveats of comparison due to early closure of the COPP/ABVD arm, escalated BEACOPP yields statistically superior FFS and OS compared with COPP/ABVD at 5 years, and a recent update indicates that its superiority has only increased by 10 years.15 The survival benefit is remarkable, especially considering that ABVD did not demonstrate superior survival compared to MOPP, albeit in a smaller trial.16 The IPS correlated with treatment efficacy in each treatment arm in this study. Based on these outstanding efficacy data, escalated BEACOPP became the new standard treatment in Germany for patients 65 years old or less.

The HD9 elderly study was simultaneously conducted in 75 patients 65 to 75 years old and compared COPP/ABVD with standard BEACOPP.17 Escalated BEACOPP was considered to be unsuitable for patients older than 65 years. In contrast to the results in younger patients, there was no difference in the treatment arms and the results were unsatisfactory in both arms, with 46% FFS and 50% OS. Toxicity was unacceptable with standard BEACOPP, which was associated with 21% acute fatal toxicity.

As anticipated, the toxicity of escalated BEACOPP was considerable, including nearly universal grade 3-4 neutropenia, 70% thrombocytopenia, 66% grade 3-4 anemia, and 22% grade 3-4 infections. These complications were more apparent in the final four treatment cycles. At 10 years, secondary leukemia was reported in 14 patients treated with escalated BEACOPP (3.2%).18 Fertility was compromised in men and women as anticipated.19 A desire to ameliorate these adverse effects while maintaining the therapeutic benefit led to subsequent GHSG studies as outlined below.

Extrapolating the COPP/ABVD Results to ABVD
Table 1 makes apparent the variation in the patient populations studied with ABVD and the COPP/ABVD arm of the HD9 study. The Intergroup trial stands out as having more stage IV patients, a higher proportion of IPS 4-7 patients, older-age patients, and the inclusion of patients who had failed prior radiotherapy. The United Kingdom study also has older patients but fewer stage IV patients. Despite these differences, the British, Italian and German studies are relatively consistent with regard to IPS scores. In Table 2, the efficacy and treatment delivery of ABVD and COPP/ABVD in these studies is described. The FFS data are inconsistent across the ABVD studies, and the reported dose intensity of drug delivery varied. Analyses of risk factors was inconsistent with regard to IPS score; advanced age and stage IV disease were recognized as prognostic factors in the U.S. Intergroup study.

The explanation for the lower FFS in the Intergroup study is subject to speculation, but might relate to the intensity of delivered therapy as well as a less favorable patient population. Neutropenia is common with the ABVD regimen and, historically, dose reductions were based on total white blood count on days 1 and 15. Several authors have demonstrated that dose intensity can be maintained with the addition of growth factors.19-21 Recently, Evens et al have suggested that full doses of ABVD can be given safely despite neutropenia and without growth factors. In a single institution experience, they were able to deliver ABVD on a median 28.2-day schedule and maintain 99% dose intensity.15 Further, their results in advanced-stage patients (n = 40) were excellent: 87% FFS at 5 years. These results are of interest based on theoretical concerns that G-CSF exacerbates severe lung injury in preclinical models.

Table 3. DH9 study results according to treatment arm and International Prognostic Score (IPS).

<table>
<thead>
<tr>
<th>Patient Population</th>
<th>Standard BEACOPP</th>
<th>Escalated BEACOPP</th>
<th>COPP/ABVD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N = 469</td>
<td>N = 466</td>
<td>N = 260</td>
</tr>
<tr>
<td></td>
<td>5 y 10 y</td>
<td>5 y 10 y</td>
<td>5 y 10 y</td>
</tr>
<tr>
<td>FFS, %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPS 0.1</td>
<td>81 79</td>
<td>92 91*</td>
<td>79 78*</td>
</tr>
<tr>
<td>IPS 2.3</td>
<td>72 71</td>
<td>87 83†</td>
<td>67 59†</td>
</tr>
<tr>
<td>IPS 4-7</td>
<td>74 56</td>
<td>82 71‡</td>
<td>59 54‡</td>
</tr>
<tr>
<td>OS, %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPS 0.1</td>
<td>93 85</td>
<td>95 94</td>
<td>92 88</td>
</tr>
<tr>
<td>IPS 2.3</td>
<td>86 84</td>
<td>90 87§</td>
<td>84 73§</td>
</tr>
<tr>
<td>IPS 4-7</td>
<td>81 63</td>
<td>82 70</td>
<td>67 61</td>
</tr>
</tbody>
</table>

* P = .015
† P = .0001
‡ P = .020
§ P = .0027
Abbreviations: BEACOPP, cyclophosphamide, adriamycin, etoposide, vincristine, bleomycin, procarbazine, prednisone; COPP/ABVD, cyclophosphamide, vincristine, prednisone, procarbazine/adriamycin, bleomycin, vinblastine, dacarbazine; FFS, failure-free survival; OS, overall survival.
and retrospective data relating older age and G-CSF use with bleomycin-associated lung toxicity.

Notably, COPP/ABVD cycles in the HD9 study were delivered over a median 46.3 days compared with the planned 30 days. As described by Diehl et al, this compares with a median cycle length of 24.3 days for standard BEACOPP and 24.7 days for escalated BEACOPP.

Taken together, these RCT results suggest that it is not appropriate to extrapolate the COPP/ABVD results from the HD9 study to suggest that they are a fair assessment of modern ABVD treatment. The marked variation in the patient populations (particularly with regard to age and stage) and the differences in treatment delivery lead to the conclusion that a RCT comparing ABVD with BEACOPP, as is in progress in the European Organization for Research on Treatment of Cancer (EORTC), is necessary.

Clinical and Biologic Prognostic Factors

The IPS reported by Hasenclever et al. resulted from a large, multinational effort coordinated by the GHSG. The IPS was variably prognostic in the studies described above, but it nicely separated outcomes in the HD9 study and was also found to be prognostic by other investigators. In Table 3, the HD9 results are reported by IPS by efficacy measures at 5 and 10 years. Remarkably, the favorable impact of escalated BEACOPP only increases over time, whereas the benefit of standard BEACOPP appears to wane. The most strikingly apparent finding is that it is necessary to treat patients with all IPS scores to achieve the benefit from escalated BEACOPP. The benefits are not restricted to high-risk disease. Accordingly, selection of high-risk patients is not an effective strategy to discern the subset of patients who benefit from dose intensification.

The other conclusion that can be drawn from Table 3 is the relatively unfavorable outcome for IPS 4-7 patients. At 10 years, 30% of patients treated with escalated BEACOPP are estimated to be deceased. More effective strategies are needed for this subgroup. In addition, as described above, older age is a consistently adverse prognostic factor. Differences in both treatment delivery and underlying biology may contribute to poor outcomes.

Clearly, BEACOPP is not indicated in patients older than 65 years, where it appears to be harmful. In contrast, treatment with ABVD in optimal doses appears to yield better results than alternative strategies.

Tissue-based and serologic biomarkers have been extensively explored as measures of prognosis in Hodgkin lymphoma (HL). Variations in morphology, classical phenotypic markers (CD15, CD20), proteins associated with apoptosis (bcl-2), T-cell composition in the microenvironment, and other tissue markers have been reported to be associated with prognosis. However, none of these has gained traction due to inconsistent results, small populations, variable therapies, and the greater strength and reliability of clinical prognostic factors. Other investigators proffered measures of tumor burden, such as β2 microglobulin and soluble CD30. The significance of Epstein-Barr virus expression by in situ hybridization has emerged as interesting marker that may be neutral or favorable in younger patients but is unfavorable in older patients, where it may serve as an indicator of immunodeficiency. Assessment of cytokine expression in the serum (IL-10) and cytokine polymorphisms has also correlated with clinical outcomes. Although these are important and promising avenues of investigation, at this time biomarkers do not appreciably impact therapeutic decision making in HD.

Early as well as end of therapy response has long been recognized to predict outcome in HD. However, the common occurrence of residual radiographic abnormalities hampered response assessment. PET using [18F] fluorodeoxyglucose (FDG) has emerged as a powerful functional imaging tool for restaging and response assessment. Studies in advanced HD, in particular, indicate that PET is highly predictive at the conclusion of chemotherapy. PET is also highly predictive of outcome in advanced HL after 2 cycles of chemotherapy and is more powerful than conventional clinical prognostic parameters. Recently, a series of 260 consecutive patients enrolled in prospective studies of PET imaging in Denmark and Italy were reported together. Patients with stage IIA with risk factors and stages IIB-IV without age restriction were enrolled. All but 11 patients received ABVD. PET scans were obtained at baseline and after 2 cycles of therapy with no change in therapy planned based on the results. The population was quite favorable, with only 17% stage IV patients, and the IPS distribution was 41% 0-1, 50% 2-3, and 9% 4-7. At 2 years, progression-free survival was more than 90% for IPS 0, more than 80% for IPS 1 and 2, and fell to approximately 60% for IPS 3 and above.

PET scans in the Gallamini et al study were visually interpreted by consensus with clinical input and considered to be negative if minimal residual uptake was present (defined as equal to or only slightly higher than the mediastinal blood pool or standardized uptake value [SUV] of 2.0-3.5). PET-2 scans were positive in 50 (19%) patients, 13% of IPS 0-2 patients, and 38% of IPS 3-7 patients. PET-2 scans were strikingly predictive of status at 2 years, with FFS of 95% for negative versus 12.8% for positive patients (P < .001). Outcomes were equally predictive among low and high IPS patients and PET-2, stage IV and age older than 45 years were most predictive of outcome in multivariate analyses. PET-2 was by far the strongest predictive factor.

Taken together, these data indicate that age is a consistent clinical prognostic factor and that elderly patients may be harmed by more intensive strategies. The IPS defines patients generally at higher risk for treatment failure, but patients who benefit from more intensive treatment are found in all IPS strata. Patients with IPS 4-7 continue to pose a challenge, as 30% fail escalated BEACOPP by 10 years. Finally, early PET imaging, as reported in selected centers, is the most prognostic determinant of treatment success or
failure. These early PET data are consistent with older data indicating that the rate of response by conventional measures is prognostic, preclinical models in curative disease, and clinical experience with tumor markers in another highly curable malignancy, testicular cancer.

New Therapeutic Strategies to Attain High Cure Rates and Optimal Survivorship

Modifications of the BEACOPP regimen were tested in subsequent GHSG studies. The HD15 study compared 8 cycles of escalated BEACOPP with 4 cycles of escalated and 4 cycles of standard BEACOPP. In presented but unpublished analyses, these treatment arms appeared to be equally effective and confirmed the original escalated BEACOPP results. Further, outcomes after a second randomization to radiotherapy or observation were not different. Another modification of BEACOPP is a 14-day schedule supported by growth factors. As cumulative doses are less with these modifications, it is anticipated that the risk of secondary leukemia would be less, and that is what has been observed to date. Likewise, the degree of acute hematologic toxicity is lessened. Effects on male and female fertility remain concerns, however. Efficacy with these modifications has been excellent, maintaining FFS above 85%. However, it will be important to carefully assess these newer strategies as the mature data become available for full analysis.

The appeal of the PET-2 data described above is the opportunity to adjust treatment early in the therapeutic course based on the hypothesis that dose intensification will effect a higher cure rate. Certainly, the incremental efficacy of dose intensification was established in the HD9 study and also serves as the basis for treating recurring HD with myeloablative chemotherapy and stem cell transplantation. Dose adjustment could be approached by starting with dose intense therapy (i.e., escalated BEACOPP) and then de-escalating based on PET-2 or, alternately, escalating to dose-intense treatment based on PET-2. Dann et al studied the former approach in 108 patients with HD dichotomized as intermediate risk (n = 69, IPS 0-2) or high risk (n = 39, IPS 3-7). Intermediate-risk patients received standard BEACOPP and high-risk patients received escalated BEACOPP with a plan to crossover based on the results of a PET or Gallium scan after 2 cycles. Of 11 intermediate-risk patients with positive studies after standard BEACOPP, 10 crossed over, of whom 2 progressed. Six of the 58 scan-negative patients remaining on standard BEACOPP had progressive disease, yielding an 84% progression-free survival for the cohort of 69 patients. Among the 39 high-risk patients, just 7 continued with escalated BEACOPP and, for the entire high-risk cohort, there were 4 treatment failures and one early discontinuation. Although the sample size is small, this interesting study suggests that risk-adapted treatment is feasible and effective.

Nonetheless, there are several caveats to risk-adapted approaches that must be considered (Table 4). First the ability to reproducibly interpret PET scans with the 92% accuracy reported by Gallamini et al must be confirmed. In contrast to prior reports from these authors, the interpretations were made by consensus, incorporated clinical information, and in an unreported number, over-ruled prior interpretations. It was not stated that the reviewers were fully blinded. In particular, the concept of minimal residual uptake, interpreted by blood pool and with an SUV of 2-3.5, has not been adopted in other risk-adapted studies, such as those in non-Hodgkin lymphoma. Another issue is one of strategy. De-escalation, as in the Dann et al study, does not protect patients from the higher risks of intensive treatment, whereas escalation may not fully recapture the benefit of more intensive treatment. The ability to successfully treat PET-positive patients with a second-line or cross-over treatment has not been established. Whether or not it is essential to treat patients with HD with aggressive therapy from the time of diagnosis, as promoted by the GHSG, is an unanswered question. Further, it is likely that patients with PET-2–positive studies after ABVD are also patients at greatest risk for failing escalated BEACOPP (most

Table 4. Challenges for therapy risk-adapted by early positron emission tomography (PET).

| PET scan acquisition and interpretation | • Consensus versus standard reading |
| • Use of other clinical information |
| • Adjudication of conflicts |
| • Minimal residual uptake |
| • Visual versus SUV versus rate of change |
| • Technical: equipment, recording |
| • Inexperienced readers |
| • Factoring likelihood of false positives and negatives |
| Escalation of therapy | • Not known to be effective |
| • Early PET-positive may also predict failure of intensive therapy |
| • Not an option for elderly |
| De-escalation of therapy | • Does not eliminate all risks (acute toxicity, sterility) |
| • Not known to be effective |
| Study design and conduct | • Assumptions made on clinical risk |
| • Assumptions made on efficacy of cross-over design |
| • Necessity for blinding |
| • Equipose for patients and physicians based on published and presented results |
notably IPS 4-7), as has been established prior to dose-intense therapy and autotransplantation. Of course, risk-adapted approaches do not address the most challenging population, the elderly. Finally, risk-adaptation based on early PET may not reveal cases that relapse later after treatment, as seen in the approximately 10% of PET-2-negative patients who progressed in the Dann et al study. 32

These caveats should not deter investigators from meeting the challenges to rigorously study risk-adaptation in clinical trials so that results can disseminate to practicing physicians and their patients. However, during the conduct of such studies, it will be essential to have a control arm for which the results of early functional imaging will not trigger a clinical decision. It will also be important that the physician community fully understands the issues and can communicate them to their patients in order for these important RCT to be conducted expeditiously.

Correspondence
Sandra Horning, MD, Stanford University, 875 Blake Wilbur Drive, Suite CC-2338, Stanford, CA 94305-5821; phone (650) 725-6456; fax (650) 725-8222; sandra.horning@stanford.edu

References


