PROGNOSTIC SIGNIFICANCE OF ACTUAL DOSE INTENSITY IN DIFFUSE LARGE CELL LYMPHOMA: RESULTS OF A TREE-STRUCTURED SURVIVAL ANALYSIS

BY

LARRY W. KWAK
JERRY HALPERN
RICHARD A. OLSHEN
SANDRA J. HORNING

TECHNICAL REPORT NO. 127
OCTOBER 1988

PREPARED UNDER THE AUSPICES OF
PUBLIC HEALTH SERVICE GRANT 5 R01-GM21215-13

DIVISION OF BIOSTATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
PROGNOSTIC SIGNIFICANCE OF ACTUAL DOSE INTENSITY
IN DIFFUSE LARGE CELL LYMPHOMA:
RESULTS OF A TREE-STRUCTURED SURVIVAL ANALYSIS

BY
LARRY W. KWAK
JERRY HALPERN
RICHARD A. OLSHEN
SANDRA J. HORNING

TECHNICAL REPORT NO. 127
OCTOBER 1988

PREPARED UNDER THE AUSPICES
OF
PUBLIC HEALTH SERVICE GRANT 5 R01-GM21215-13

DIVISION OF BIOSTATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
PROGNOSTIC SIGNIFICANCE OF ACTUAL DOSE INTENSITY IN DIFFUSE LARGE CELL LYMPHOMA: RESULTS OF A TREE-STRUCTURED SURVIVAL ANALYSIS

Larry W. Kwak, M.D., Ph.D.\textsuperscript{1}  
Jerry Halpern, Ph.D.\textsuperscript{2}  
Richard A. Olshen, Ph.D.\textsuperscript{3}  
Sandra J. Horning, M.D.\textsuperscript{1}

\textsuperscript{1}Division of Oncology, Department of Medicine and the \textsuperscript{2}Division of Biostatistics, Department of Health Research and Policy, Stanford University Medical Center, Stanford, CA 94305

\textsuperscript{3}Department of Mathematics and Laboratory for Mathematics and Statistics, University of California, San Diego, La Jolla, CA 92093

Supported in part by NIH grants CA 34233 and CA 41628, the John Simon Guggenheim Memorial Foundation, and NSF Grant DMS 85-05609. Larry W. Kwak was a Fellow of the American Cancer Society and Richard A. Olshen was a Research Scholar in Cancer of the American Cancer Society.

Presented in part at the 24th annual meeting of the American Society of Clinical Oncology, May 22-24, 1988, New Orleans, LA.

Address reprint requests to Sandra J. Horning, MD, Division of Oncology, Room M211, Stanford University Medical Center, Stanford, CA 94305.

Running Head: Dose Intensity in Diffuse Large Cell Lymphoma
ABSTRACT

While diffuse large cell lymphoma (DLCL) is considered to be highly curable with current therapy, treatment failures are observed even with intensive combination chemotherapy regimens. In order to study the prognostic significance of actual dose intensity of chemotherapy in DLCL, we retrospectively analyzed 115 previously untreated patients treated at Stanford between 1975 and 1986 with CHOP (cyclophosphamide, Adriamycin, vincristine, prednisone), (M)BACOD (methotrexate, bleomycin, Adriamycin cyclophosphamide, vincristine, dexamethasone) or MACOP-B (methotrexate, Adriamycin, cyclophosphamide, vincristine, prednisone, bleomycin). The actual relative dose intensity (RDI), the amount of drug actually administered to each patient during the first 12 weeks of therapy, was calculated as standardized to CHOP and analyzed in combination with clinical factors prognostic for survival by univariate analysis. Multivariate recursive partitioning (tree structured) survival analysis identified the actual RDI of Adriamycin > 75% as the single most important predictor of survival. A model incorporating the actual RDI of Adriamycin and performance status, in combination with serum lactate dehydrogenase and extranodal disease, defined three overall prognostic groups of patients, with respective three-year survival rates of 89%, 63%, and 18%. The three prognostic groups remained distinct, even when restricted to complete responders. This model was also predictive of survival when dose intensity was analyzed relative to the optimum dose defined for each of the three regimens and when applied to a subgroup of patients age 50 years or younger. There were no statistically significant differences among the treatment regimens. We conclude that actual RDI is an important prognostic factor for survival in DLCL and that analysis of RDI early in the course of treatment may allow modification of therapy.

Key Words:  Dose Intensity
            Diffuse Large Cell Lymphoma
            Tree Structured Survival Analysis
INTRODUCTION

Diffuse large cell lymphoma (DLCL) is the most common histologic subtype among the diffuse aggressive lymphomas, accounting for approximately one-third of all non-Hodgkin's lymphomas (1). Once almost uniformly fatal, as many as 50 to 70% of advanced stage patients and 80% of limited stage patients remain alive on the plateau phase of survival curves following treatment with modern combination chemotherapy regimens (2-7). While a number of prognostic variables have been identified by several groups of investigators, reliable identification of patients at high risk for treatment failure is made difficult by the heterogeneity of DLCL (8-15).

Recently, there has been widespread interest in the concept of dose intensity, the amount of chemotherapeutic drug delivered per unit time, as a significant factor affecting outcome. The correlation between dose intensity of combination chemotherapy and treatment outcome has been demonstrated for a number of tumors, including breast cancer, ovarian cancer, Hodgkin's disease, small cell lung cancer, and multiple myeloma (16-21). DeVita has suggested that dose intensity is also correlated with survival in DLCL and that combinations of six or more drugs are superior to CHOP (cyclophosphamide, Adriamycin, vincristine, and prednisone) (22). However, dose intensity has been a poorly controlled variable in almost all studies of DLCL, and it is not clear that the more intensive multidrug regimens are superior to CHOP, delivered at full doses every three weeks, to comparable patients. For these reasons, we undertook a survival analysis of 115 patients with DLCL treated at Stanford between 1975 and 1986 with three combination chemotherapy regimens: CHOP, (M)BACOD (methotrexate, bleomycin, Adriamycin, cyclophosphamide, vincristine, dexamethasone), and MACOP-B (methotrexate, Adriamycin, cyclophosphamide, vincristine, prednisone, bleomycin). Using the multivariate statistical approach of recursive partitioning, or tree-structured, survival analysis, we analyzed drug dose intensity, based on the amount of drug actually delivered during the first 12 weeks of therapy, in combination with pretreatment clinical prognostic factors. In the process of this analysis, a model was constructed which produces three distinct prognostic groups with markedly different overall survival.

PATIENTS AND METHODS

Patient Characteristics

One hundred fifteen previously untreated patients with a histologic diagnosis of DLCL, intermediate grade or immunoblastic, as defined by the Working Formulation, who were seen and treated at Stanford between 1975 and 1986 comprised the study group (1). Patients were staged according to the Ann Arbor criteria (23). Initial staging procedures in all patients included complete physical examination, chest radiograph, routine blood counts and chemistry tests, bone marrow
biopsy, bipedal lymphangiography and/or abdominal and pelvic computed tomography. Additional tests were performed as clinically indicated.

Patients in this study received CHOP, (M)BACOD, or MACOP-B treatment in standard doses in three week cycles for CHOP and (M)BACOD and weekly for MACOP-B (2,4,24). The high dose methotrexate in (M)BACOD was given as 1 G/m² on day 14 of each cycle with leucovorin rescue. Selection of treatment regimen was individualized for each patient, but in most cases reflected the regimen with the best results reported at that time. Thus, the standard therapy for advanced stage patients with DLCL at Stanford was CHOP prior to 1981, (M)BACOD between 1981 and 1985, and MACOP-B beginning in 1985. CHOP continued to be used after 1981 for limited stage patients, especially for those with non-bulky or asymptomatic disease. Patients who received radiation therapy were included in the analysis only if it was given after the completion of chemotherapy.

Sixty-eight patients were treated with CHOP, 31 with (M)BACOD, and 16 with MACOP-B. The clinical characteristics of these patients are shown in Table 1. There were several differences in age and sex ratio among the groups, whereas the distribution of ECOG performance scores was essentially constant across all three groups. Performance status was based on the Eastern Cooperative Oncology Group scale, in which 0 is associated with no symptoms, 1 with symptoms but continued ability to ambulate, 2 with bedridden status less than 50% of the day, 3 with bedridden status greater than 50% of the day, and 4 with chronic bedridden status and a requirement for assistance for daily maintenance. Overall, approximately one-third of patients presented with B symptoms. The (M)BACOD group contained the highest proportion of advanced stage patients (74%). Overall, the majority of patients were Ann Arbor stage III or IV, and the (M)BACOD and MACOP-B groups contained the highest proportion of patients with elevated pretreatment lactate dehydrogenase (LDH), multiple extranodal sites of involvement, and bulk of disease, defined as the largest dimension of a single tumor mass. The size of the largest mass was determined by review of both the medical record and relevant radiographic studies and included both nodal and extranodal sites. Actual dimensions of masses ≥3cm in diameter were used in the analysis. Sixty-four per cent of patients presented with extranodal sites of disease of which the gastrointestinal tract was the most commonly involved extranodal site (19%). Other sites of extranodal disease included pleura (13%), bone marrow (12%), bone (11%), lung (10%), testis (9%), skin (7%), pericardium (6%), kidney (6%), and liver (3%).

All patients underwent restaging with baseline staging procedures during treatment and at the completion of treatment. Bone marrow biopsies were, for the most part, repeated only in patients having positive biopsies on presentation. Complete response was defined as the disappearance of all clinical evidence of disease and either the normalization or stabilization of all laboratory and radiographic abnormalities. To be classified as a complete response, the response had to be maintained for at least 30 days following the completion of therapy. Determination of the number of cycles to complete remission was based on restaging studies described above. A value of zero cycles was
assigned to patients who had undergone surgical resection of all gross disease prior to the initiation of chemotherapy (11 patients).

**Dose Intensity**

Actual chemotherapy doses were available for all 115 patients. The method of Hryniuk and Bush (16) was used to calculate the dose intensity (DI) of each drug actually administered to the patient. For the purposes of this analysis, we considered the amount of each drug, normalized to body surface area (mg/m²), administered during the first 12 weeks of therapy. Twelve weeks was selected, because (1) the MACOP-B regimen is of 12 weeks duration; (2) if found to be important, DI could then be used to make treatment decisions, including changes in therapy, relatively early in the course of treatment; and (3) a previous study has shown that response during the first three to four cycles of chemotherapy is prognostically significant (25). Although this method may introduce bias in patients with early disease progression, the primary regimen was changed in only three patients prior to the 12th week of treatment.

DI was expressed as a decimal fraction of the dose prescribed in a standard regimen over the same time frame (relative dose intensity). CHOP was selected as the standard regimen because the group treated with CHOP represented the majority of our patients and also because CHOP remains the standard of comparison for the second and third generation combination chemotherapy regimens. Average relative dose intensity (RDI) for the combination of drugs was calculated by taking the arithmetic mean of RDI of the four individual drugs for each patient. The projected relative dose intensities for each drug and the four-drug average for each of the regimens standardized to CHOP are shown in Table 2. These are based on 100% of the calculated dose for that particular regimen without any delays in treatment. From these calculations it is apparent that the regimen with the highest dose intensity of cyclophosphamide is CHOP, while MACOP-B has the highest dose intensity of Adriamycin. Several assumptions were made in calculating actual RDI: (1) a maximum value of 1.00 was allowed for the RDI of prednisone/dexamethasone when calculating average RDI, since no clear dose-response relationship exists for steroids given in therapeutic doses, (2) one mg of dexamethasone was equivalent to 6.7 mg of prednisone, based on relative glucocorticoid activity, (3) the capping off of vincristine doses at 2.0 mg for CHOP meant that the RDI of this drug for patients treated with (M)BACOD varied with body surface area above 1.43 m², and (4) for two patients who had RDI of a single drug > 1.00 (other than prednisone), a value of 1.00 was assigned when calculating average RDI, and (5) three patients treated with (M)BACOD received higher doses of methotrexate, between 1 and 3 G/m². A maximum value of 1.50 was assigned to the RDI of this drug when calculating average RDI (for each regimen as its own standard).
Statistical Methods

Survival curves were calculated from the date of initiation of treatment according to the actuarial method of Kaplan and Meier (26). The generalized Wilcoxon test of Gehan was used to assess the significance of differences between patient groups (27). The ability of each variable alone to predict survival was tested by a univariate Cox analysis (28). Multivariate Cox analyses with up to eight variables were also done. For a particular model, each variable was treated separately and up to five variables were tested in combination for their additional prognostic significance beyond that of the other variables in the model (28,29).

A new and important aspect of our work is a recursive partitioning approach to survival analysis. The method is based on extensions by Gordon and Olshen of previous work on classification and regression trees (30-34). In this computer-intensive approach to survival analysis, the group of 115 patients was first divided into subgroups that were as alike as possible in survival time. The division was effected by an optimally chosen yes-no question concerning an individual variable. To begin, a particular variable was considered, and for each observed value, v, of the variable, every patient was tentatively put into one of two subgroups depending on whether that patient's value exceeded v or not. For each subgroup pair, a logrank statistic was computed for the difference (35). The particular variable, its value v, and the subgroup pair for which the difference was greatest was saved for further use. Every quantitative variable was treated in this way (qualitative variables like marrow involvement are treated in a similar manner). At this point, for each variable, a candidate optimal yes-no question was found. The next step was to pick the best among these. We made that choice by using a (fourth power Wasserstein) criterion which compared the horizontal distance between the survival functions in each subgroup pair (30).

By successively repeating this process on each chosen subgroup, a large binary tree was generated. A smaller subtree of the large tree was selected by cross-validation, and this small tree is reported (36). The technique has a built in method for handling missing data, of which there is very little in our study group (37). Hemoglobin was excluded for three patients in whom gastrointestinal bleeding was documented at presentation, and the LDH value for a single patient was not available; otherwise, there were no missing data for the remainder of the 115 patients.

A Kaplan-Meier survival curve was made for each distinct group identified by recursive partitioning, and pairs of these curves were compared with Gehan's generalized Wilcoxon test.

RESULTS

Response and Survival Data

Figure 1a shows the Kaplan-Meier survival curve for the total group of 115 patients. The actuarial five-year survival is 58% with a median follow-up of survivors of three years (range, 8 months - 7.5 years). The survival of the patients is stratified by regimen in Figure 1b. As shown,
when the three regimens were compared for survival, no statistically significant differences were observed. Complete response (CR) rates were high for all three groups, although the (M)BACOD regimen was associated with a somewhat lower CR rate (68%) compared to the other two regimens (CHOP 87% and MACOP-B 94%). Differences in patient characteristics among the three groups and the fact that the regimens, for the most part, were not used concurrently may account for the lack of observable differences in survival. For instance, the (M)BACOD group contained a higher proportion of patients with adverse prognostic factors (Table 1), which may, in addition, explain its lower CR rate, and there is a relatively small number and shorter follow-up of the MACOP-B patients.

**Univariate Analysis of Clinical Features for Survival (Table 3)**

Eleven pre-treatment clinical factors were evaluated individually for the entire group of 115 patients as prognostic indicators of survival (Table 3). ECOG performance status, number of extranodal sites of disease, and Ann Arbor stage, all analyzed as continuous or ordered variables, were significant (p < 0.05), with poor performance status, greater number of extranodal sites, and more advanced stage each associated with an adverse effect on survival. Bone marrow involvement and the presence of B symptoms were also associated with a significantly adverse effect on survival. Hemoglobin and LDH, both analyzed as continuous variables, were of borderline statistical significance, as was gender, in favor of females. Single factors not predictive for survival included age and bulk, analyzed as continuous variables, histology (intermediate grade vs. immunoblastic), and specific nodal or extranodal sites of involvement other than bone marrow. The small number of patients involved in each specific site of involvement may have precluded the full analysis of its influence on survival.

The effect of consolidative radiation therapy on survival was also evaluated. Thirty patients had been selected to receive radiation therapy to sites of initial bulky disease following the achievement of CR and completion of chemotherapy. When they were compared to patients having an initial mass \( \geq 5 \text{ cm} \) who had not received consolidative radiation therapy in CR (n=23) a significant difference in survival was observed in favor of radiotherapy (Table 3) suggesting a beneficial role for radiation therapy as an adjunct to chemotherapy in this setting.

The time interval required to achieve complete remission was also examined as a variable for survival. Among the 96 patients (83%) who achieved a complete remission, the number of cycles of chemotherapy required to achieve CR, analyzed as a continuous variable, was not of prognostic significance (Table 3). These results should be interpreted with caution, however, as not all patients were restaged at the same time points in their therapy, and some patients may have achieved complete remission at time points somewhat earlier than indicated by formal restaging.
Relative Dose Intensity

Shown in Table 2 are the median actual RDI and the corresponding projected RDI, of each drug and of the four-drug average for each of the regimens, standardized to CHOP, based on doses of chemotherapy received by each of the 115 patients during the first 12 weeks of therapy as described above. With the exception of six patients, there were no arbitrary dose reductions for advanced age or poor performance status. The most frequent reason for dose attenuation or treatment delay was leukopenia (8%), or an episode of fever with absolute neutropenia associated with a prior cycle of treatment (12%). Liver dysfunction necessitated dose attenuation of Adriamycin in one patient, and prednisone was removed from the regimen of one patient because of severe mental status changes and in another patient because of severe muscle weakness. Neurotoxicity associated with vincristine resulted in dose attenuations or deletions of this drug from the regimen of 11% of patients. However, patients in all three treatment groups received close to full projected doses as shown by the small differences between projected and actual RDI for each of the drugs within each regimen.

When evaluated separately as continuous variables for the entire group of 115 patients, average RDI, as well as the RDI for cyclophosphamide and Adriamycin were significant (p < 0.05) for survival (Table 2). The RDI for vincristine and prednisone were not significant for survival in this analysis. Similarly, the RDI for methotrexate and bleomycin evaluated for the 47 patients treated with (M)BACOD or MACOP-B had no significant effect on survival (p > 0.05, data not shown).

Tree Structured Survival Analysis (Multivariate Recursive Partitioning)

The covariates selected for this analysis were: (1) the pre-treatment clinical factors that were prognostically significant in the univariate analysis (ECOG performance status, number of extranodal sites, stage, marrow involvement, B symptoms, and LDH), (2) age, sex, and bulk, which were not statistically significant in the univariate analysis but were included because other retrospective analyses have found these factors to be prognostically important, (3) average RDI and the RDı for cyclophosphamide and Adriamycin, and (4) treatment regimen.

Tree-structured survival analysis identified the RDI of Adriamycin > 75% as the first split point (Figure 2). This divided the patients into groups of 92 and 23 patients; superior survival was associated with the group who received higher doses of Adriamycin (Figure 3a). The significance of actual dose intensity as a determinant of survival was further supported by the identification of average RDI (all four drugs) and RDI of cyclophosphamide as the surrogate covariates. The term surrogate covariate refers to the covariate which gave rise to the next best split (37).

The less favorable group (those with RDI of Adriamycin ≤ 75%) of patients was best further subdivided by pre-treatment LDH resulting in two terminal subgroups (Figure 2, subgroups I and II). Patients with RDI of Adriamycin ≤ 75% and pretreatment LDH > 1.4 x normal formed a particularly poor prognostic group (Figure 3b). The surrogate covariate for this node was stage.
Figure 2 shows that patients who received > 75% RDI of Adriamycin were best further subdivided by ECOG performance status. Patients with poor initial performance scores (ECOG 3 or 4) despite > 75% RDI of Adriamycin (subgroup III), had a very poor prognosis as demonstrated in Figure 3c. The surrogate covariates for this node were number of extranodal sites and age.

The favorable group of patients with RDI Adriamycin > 75% and good performance status was then further subdivided by the number of extranodal sites of disease. As shown in Figure 2, this generated the most favorable subgroup (IV) of all patients, and a second subgroup (V) with an intermediate prognosis. These results are shown in Figure 3d. The surrogate covariates for this node were age and the RDI of Adriamycin and cyclophosphamide.

The resulting five subgroups of patients at each terminal node divide naturally into three overall prognostic groups. The survival curves of these three groups are shown in Figure 4. The favorable group (three-year survival 89%) consists of those patients with RDI of Adriamycin > 75%, ECOG performance status 0-2, and no extranodal disease (subgroup IV). The poor prognostic group (three-year survival 18%) consists of those patients with RDI of Adriamycin ≤ 75% together with significantly elevated LDH (subgroup II), or poor performance status, regardless of dose intensity (subgroup III). The intermediate group (three-year survival 63%) consists of the remaining patients: those with RDI of Adriamycin ≤ 75% without significantly elevated LDH (subgroup I), and those with RDI Adriamycin > 75% and good performance status, but at least one extranodal site of disease (subgroup V). The complete response rates for the three groups were 97% for the favorable, 84% for the intermediate, and 60% for the poor prognostic groups. It is important to note that 60% of the patients in the poor prognostic group achieved CR, as it suggests that the effect of the variables we have identified on survival is not solely related to the ability to achieve a CR. This point is further illustrated by the survival curves in Figure 5, which show that the three prognostic group comparisons maintain statistical significance when this model is applied only to the 96 patients who achieved a CR.

One of the aims of this study was to identify a subgroup of younger patients with poor prognosis who may be candidates for very aggressive or investigational therapy, such as bone marrow transplantation. To this end, we applied the model to the subgroup of patients age ≤ 50 years (n = 58). The analysis for this subset of patients was not as extensive, in part because of the smaller number of patients, and the majority of the patients fell into a favorable prognostic group; that is, high RDI of Adriamycin and good performance status (Figure 6a). However, using the same covariates with slightly different split points, it was possible to identify a small subgroup of patients at high risk for failure (Figure 6b). The characteristics of this poor prognostic group (actuarial two year survival approximately 20%) were: (1) RDI of Adriamycin ≤ 85% and LDH > 1.1 times normal or (2) ECOG performance status 3 or 4, despite a high RDI of Adriamycin.

An alternative method of analyzing dose intensity would be to consider the amount of drug delivered during the first 12 weeks of therapy relative to the optimum dose defined for each regimen.
Such an approach would allow more general application to other regimens used in the treatment of DLCL. While it is understood that 75% of the optimum Adriamycin dose in one regimen may not be equivalent to that in another regimen, considering each regimen as its own standard also would reflect indirectly the impact of additional drugs in the regimen not included in CHOP (e.g. bleomycin, methotrexate). When this approach was used in calculating dose intensity, the RDI for Adriamycin, cyclophosphamide, and the average RDI (four or six drugs) were again highly significant for survival in univariate analysis ($P < 0.01$, data not shown). Moreover, when these values of RDI were incorporated into the multivariate model, again three overall prognostic groups were generated (Figure 7).

The same covariates were also included in multivariate Cox regression analyses for survival. Only a single patient had to be excluded from these analyses because of missing data (pre-treatment LDH was not available). Among the pre-treatment clinical factors that were prognostically significant in the univariate analysis (Table 3) or identified in the tree-structured analysis, only ECOG performance status consistently retained additional significance beyond that of various combinations of factors ($p < 0.05$). Of interest, gender, in favor of females, also was found to be significant in combination with a number of pre-treatment clinical factors ($p = 0.02$). The individual RDI of Adriamycin and cyclophosphamide, as well as the average RDI, also retained significance independently ($p = 0.02$, 0.04, and 0.07 respectively), and in combination ($p = 0.048$ by the likelihood ratio test), when analyzed together with ECOG performance status. Again, treatment regimen provided no significant additional prognostic information, when analyzed with RDI and ECOG performance status.

Discussion

There has been increasing interest in the concept of dose intensity with respect to cancer chemotherapy. Retrospective studies of chemotherapeutic regimens for a variety of cancers have reaffirmed the importance of dose in achieving a maximum therapeutic effect in responsive tumors and have highlighted the implications of this concept for the design and analysis of future clinical trials (16-21,38). In our analysis of 115 patients with DLCL, we have shown that actual RDI, as measured during the first 12 weeks of therapy, is an important prognostic factor for survival in DLCL. Using a tree-structured analysis, the characteristic that generated two patient groups most different in prognosis was the actual RDI of Adriamycin. Univariate analysis and multivariate Cox regression analysis also confirmed the impact of actual RDI on survival in DLCL.

Dose intensity has been shown to be important in the treatment of lymphomas. Carde and co-workers reported that dose intensity during the first three cycles of MOPP (nitrogen mustard, vincristine, prednisone, procarbazine) chemotherapy was significantly related to outcome in Hodgkin's
disease (19). Using a hypothetical nine-drug regimen as the standard of comparison, DeVita and colleagues at the National Cancer Institute (NCI) used dose intensity analysis to calculate projected dose intensities for a variety of treatment programs used in the management of diffuse aggressive lymphomas (22). They found a strong correlation of average RDI (based on nine drugs) to long-term survival, but no relationship of outcome to separately calculated two-drug average RDI for Adriamycin and cyclophosphamide. While our analysis confirms a correlation between average RDI and outcome, there are several important differences between the NCI analysis and our study which may explain the discrepancy with regard to the individual RDI of Adriamycin and cyclophosphamide. First, the availability of specific data on actual doses of drugs delivered to each of the 115 patients in our study allowed precise calculation and subsequent analysis of actual dose intensity. It is important to emphasize the difference between actual and projected dose intensity. Actual dose intensity may vary significantly from projected dose intensity for individual patients because of the nature of dose adjustments based on drug toxicity or other factors. Clearly, actual dose intensity is the datum of greater importance and would be expected to correlate much more closely with outcome than dose intensity calculated from an intended protocol (39). Second, initial dose reductions up to 50% were allowed for patients over age 65 treated with CHOP in the Southwest Oncology Group (SWOG) studies cited in the study by DeVita et al. The effect of these dose reductions on actual dose intensity could not have been reflected in calculations of projected dose intensity, and this factor may, at least in part, explain the lack of correlation of the two-drug average RDI with outcome. Third, two assumptions made when using the DI method to calculate average RDI of two drugs, (1) that both drugs in a multidrug regimen are equally effective and (2) that the effects of the individual drugs are additive, are avoided in a multivariate analysis in which the RDI of Adriamycin and cyclophosphamide are treated as separate variables. Finally, calculations of dose intensity for CHOP in the NCI study were based on four-week cycles; at Stanford CHOP is routinely administered every three weeks as originally described by McKelvey et al. (24).

Nevertheless, several limitations, outlined in recent reviews, remain with all studies utilizing the dose intensity approach (39-41). Among these is the assumption that drug scheduling is relatively unimportant. In addition, it remains to be seen to what extent dose intensity correlates with outcome independent of total dose. For example, among 121 patients with DLCL treated with at least eight cycles of m- or M-BACOD chemotherapy, no relationship was found between survival and the percentage of the total prescribed dose of each drug actually administered (13).

One of our goals in evaluating this series of patients with DLCL was the identification of prognostic groups that would have implications for the design of randomized therapeutic trials and for determination of optimal therapy for individual patients. The actual RDI of Adriamycin and performance status, in combination with pre-treatment LDH and the number of extranodal sites of disease, defined three groups of patients with markedly different survival. The identification of these
prognostic groups underscores the predictive power of the novel statistical approach of recursive partitioning, or tree-structured, survival analysis. The tree-growing paradigm of examining nearly every allowable split on each variable and selecting the best of these splits at each point, has often been applied to classification data, such as the identification of clinical factors predictive for myocardial infarction (31,34), and in this study has been extended to the context of censored survival data. The analysis presented here illustrates the prospective use of this approach and, once the prognostic groups have been identified, the ease with which a given patient may be properly assigned to one of them. This technique represents a significant advance in the technology of survival analysis. Other advantages of the tree-structured recursive partitioning approach are the very few assumptions that are necessary to use it (and hence its broad applicability), its automatic identification of interactions; that is, synergistic effects among the variables included for analysis; and the natural way that it handles missing data (30,37,42).

The three risk groups generated in the model were prognostic when limited to the population of complete responders. The relatively high CR rate of 60% in the poor prognostic group, suggested that the remissions among the three risk groups were not equally durable. These results are in contrast to the observation that the improvement in overall survival observed between two generations of chemotherapy regimens relates solely to the ability of the regimens with greater dose intensity to induce higher CR rates (22). RDI and the clinical prognostic factors we have identified may not only exert their influence on the achievement of a CR, but may also be used to predict groups at higher risk for relapse following CR. We have also shown in Figure 7 that the model remains valid when the dose intensity analysis is standardized to regimens other than CHOP. In fact, when each regimen was considered as its own standard, a range of values for RDI Adriamycin was found over which three significantly different risk groups were generated. The precise limits of the split point on RDI Adriamycin for general application to patients with DLCL must be determined by comparison with other retrospective and future prospective studies of actual RDI of this drug in other regimens. Thus, in Figure 7 the split on RDI of Adriamycin > 75% refers to 75% of the optimal dose of Adriamycin in the particular regimen used.

The favorable prognostic group (three-year survival 89%) appears to be effectively treated with existing chemotherapy regimens. Significant improvements over existing survival rates for this group of patients may be difficult to demonstrate and achieve without substantial additional toxicity. Conversely, the poor prognostic group (three-year survival 18%) requires alternative therapy. Several reports have suggested a role for high dose therapy and autologous bone marrow transplantation in selected patients with relapsed non-Hodgkin's lymphoma (43,44). Younger patients at high risk for treatment failure may be candidates for primary bone marrow transplantation following cytoreductive chemotherapy to complete remission, or minimal disease status, if they can be confidently identified early on. This approach as reported by Gulati et al appears very promising (45). The observation that
60% of our poor risk patients achieved a CR suggests that the majority can be effectively cytoreduced with chemotherapy. Furthermore, when we applied our model to patients age \( \leq 50 \), a small subset of patients with suitably poor prognosis (two year survival approximately 20%) could be identified. Thus, analysis of dose intensity within the first 12 weeks of treatment may allow modification of therapy, including alternative experimental regimens or bone marrow transplantation.

Multiple pre-treatment clinical factors prognostic for survival have been reported in DLCL; however, many have failed to retain significance in subsequent multivariate analyses, and the variability of factors has led to some confusion in analyzing and comparing clinical trials (8-15). Age, performance status, and factors which reflect tumor burden and growth characteristics, such as stage, the number and specific sites of extranodal disease (e.g. marrow, gastrointestinal tract), bulk, and serum LDH, consistently emerge as significant. Our results are in agreement with these studies, although bulk, measured as the largest diameter of a single mass, and analyzed as a continuous variable, was not found to be significant in our analysis. It should be noted that bulk, measured in this way, was not a significant variable in several other analyses. (5,9,11,12,15). This may relate to the difficulty in assessing this parameter when multiple sites of bulk are involved. In this regard, a method of assessing bulk which takes into consideration more than one site and accounts for both nodal and extranodal disease, such as that suggested by Jagannath and colleagues, may more accurately reflect tumor burden (14). In addition, the lack of a significant adverse effect of bulk on survival in our analysis may, in part, be due to the significant proportion of patients with initial bulky disease who received radiation therapy following the completion of chemotherapy (57%). The significant survival advantage demonstrated among patients who received consolidative radiation therapy to sites of initial bulky disease supports this explanation (Table 3). Age, evaluated as a continuous variable, also was not significant for survival in our analysis. A significant number of patients were age 60 or older (29%), although only 11% of the patient group was older than age 65. The relatively small number of elderly patients may, in part, explain the lack of impact of age. Nonetheless, age did emerge as a surrogate covariate at two points in the tree-structured survival analysis.

The comparison of nonconcurrent groups, variation in length of follow-up, and differences in patient characteristics among the three chemotherapy treatment groups may obscure the impact of treatment on survival. In our study, no significant differences in survival among the three treatment regimens could be demonstrated by univariate, multivariate recursive partitioning or Cox regression analysis. Valid comparison of these three regimens is the goal of a large prospective randomized study being conducted by the Southwest Oncology Group.

Our analysis confirms the importance of drug dose in achieving a maximal therapeutic effect and sets forth dose intensity, at least as measured during the first 12 weeks of therapy, as a strong predictor of overall survival in DLCL. If these results can be confirmed in other similar groups of
patients, then it will be clear that actual dose intensity should be included in the analysis and reporting of future clinical trials of DLCL and that dose intensity should be considered in the selection of alternative or consolidative therapies for DLCL.
ACKNOWLEDGEMENTS

The authors thank Linda Aronson for her assistance in preparation of the manuscript. The authors also express their gratitude to Saul A. Rosenberg, MD for critical review of the manuscript.
References


LEGENDS

Figure 1. (A) Actuarial Survival of 115 Patients with DLCL.
(B) Actuarial Survival and Complete Remission Rates of 115 Patients with DLCL by Regimen.

Figure 2. Results of Multivariate Recursive Partitioning (Tree-Structured) Survival Analysis of 115 Patients with DLCL.
RDI = relative dose intensity. LDH = lactate dehydrogenase, 1.4 x upper limit of normal value. ECOG Performance = Eastern Cooperative Oncology Group performance status. Numbers in squares indicate number of patients in each subgroup. Roman numerals designate terminal subgroups.

Figure 3. Actuarial Survival Curves of the Subgroups of Patients Generated at Each Sequential Split Point in the "Tree" Shown in Figure 2.
Split points: (A) RDI Adriamycin, (B) LDH, (C) ECOG performance status, (D) number of extranodal sites of disease.

Figure 4. Actuarial Survival Curves and Complete Response Rates for Favorable-, Intermediate-, and Poor-Prognosis Patients Identified by Tree-Structured Survival Analysis in Figure 2.
The respective three-year survival rates are 89%, 63%, and 18%. The favorable group consists of terminal subgroup IV, the intermediate group consists of subgroups I and V, and the poor prognosis group consists of subgroups II and III.

Figure 5. Actuarial Survival Curves of the Three Prognostic Groups of Patients Restricted to Patients Who Achieved Complete Remission (96 patients).

Figure 6. (A) Recursive Partitioning Survival Analysis Applied to Patients Age \( \leq 50 \) (58 patients). Numbers in squares indicate number of patients in each subgroup.
(B) Actuarial Survival Curves of the Four Terminal Subgroups Grouped by Prognosis.
The majority of patients exhibited favorable prognosis (49 patients). Characteristics of the poor prognosis group (9 patients) were RDI of Adriamycin \( \leq 85\%\) together with LDH > 1.1 times normal or ECOG performance status 3-4 despite RDI of Adriamycin > 85%.
Figure 7. Actuarial Survival Curves of the Three Prognostic Groups in Figure 4 with Dose Intensity Calculated Relative to the Optimum Dose Defined for Each Regimen. The three groups remain distinct when dose intensity is standardized to regimens other than CHOP.
<table>
<thead>
<tr>
<th></th>
<th>CHOP</th>
<th>(M)BACOD</th>
<th>MACOP-B</th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Patients</td>
<td>68</td>
<td>31</td>
<td>16</td>
<td>115</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 50 years</td>
<td>30</td>
<td>17</td>
<td>11</td>
<td>58 (51)</td>
</tr>
<tr>
<td>51-65 years</td>
<td>26</td>
<td>13</td>
<td>5</td>
<td>44 (38)</td>
</tr>
<tr>
<td>&gt; 65 years</td>
<td>12</td>
<td>1</td>
<td>0</td>
<td>13 (11)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>43</td>
<td>16</td>
<td>6</td>
<td>65 (57)</td>
</tr>
<tr>
<td>Female</td>
<td>25</td>
<td>15</td>
<td>10</td>
<td>50 (43)</td>
</tr>
<tr>
<td>ECOG Performance Status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-2</td>
<td>60</td>
<td>26</td>
<td>14</td>
<td>100 (87)</td>
</tr>
<tr>
<td>3-4</td>
<td>8</td>
<td>5</td>
<td>2</td>
<td>15 (13)</td>
</tr>
<tr>
<td>B Symptoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>51</td>
<td>18</td>
<td>7</td>
<td>76 (66)</td>
</tr>
<tr>
<td>Present</td>
<td>17</td>
<td>13</td>
<td>9</td>
<td>39 (34)</td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I, II</td>
<td>35</td>
<td>8</td>
<td>7</td>
<td>50 (43)</td>
</tr>
<tr>
<td>III, IV</td>
<td>33</td>
<td>23</td>
<td>9</td>
<td>65 (57)</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 12</td>
<td>16</td>
<td>11</td>
<td>5</td>
<td>32 (28)</td>
</tr>
<tr>
<td>≥ 12</td>
<td>52</td>
<td>20</td>
<td>11</td>
<td>83 (72)</td>
</tr>
<tr>
<td>Lactate dehydrogenase</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal Range</td>
<td>29</td>
<td>8</td>
<td>4</td>
<td>41 (36)</td>
</tr>
<tr>
<td>Elevated</td>
<td>39</td>
<td>23</td>
<td>12</td>
<td>74 (64)</td>
</tr>
<tr>
<td>Bone Marrow Involvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>3</td>
<td>7</td>
<td>4</td>
<td>14 (12)</td>
</tr>
<tr>
<td>No</td>
<td>65</td>
<td>24</td>
<td>12</td>
<td>101 (88)</td>
</tr>
<tr>
<td>Number of Extranodal Sites</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-1</td>
<td>57</td>
<td>18</td>
<td>10</td>
<td>85 (74)</td>
</tr>
<tr>
<td>≥2</td>
<td>11</td>
<td>13</td>
<td>6</td>
<td>30 (26)</td>
</tr>
<tr>
<td>Tumor Bulk</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 10 cm</td>
<td>53</td>
<td>15</td>
<td>10</td>
<td>78 (68)</td>
</tr>
<tr>
<td>≥ 10 cm</td>
<td>15</td>
<td>16</td>
<td>6</td>
<td>37 (32)</td>
</tr>
<tr>
<td>Histology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immunoblastic</td>
<td>12</td>
<td>11</td>
<td>6</td>
<td>29 (25)</td>
</tr>
<tr>
<td>Non-immunoblastic</td>
<td>56</td>
<td>20</td>
<td>10</td>
<td>86 (75)</td>
</tr>
<tr>
<td>Consolidative Radiotherapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>22</td>
<td>12</td>
<td>10</td>
<td>44 (38)</td>
</tr>
<tr>
<td>No</td>
<td>46</td>
<td>19</td>
<td>6</td>
<td>71 (62)</td>
</tr>
</tbody>
</table>
Table 2. **UNIVARIATE ANALYSIS OF ACTUAL DI FOR SURVIVAL**

<table>
<thead>
<tr>
<th>RDI for</th>
<th>Relative Dose Intensities&lt;sup&gt;0&lt;/sup&gt;</th>
<th>CHOP</th>
<th>(M)BACOD</th>
<th>MACOP-B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Projected Median Actual</td>
<td>Projected Median Actual</td>
<td>Projected Median Actual</td>
<td>p-Value*</td>
</tr>
<tr>
<td>Four-drug average</td>
<td>1.00 0.96</td>
<td>0.84 0.84</td>
<td>1.18 1.17</td>
<td>0.04</td>
</tr>
<tr>
<td>Cyclophosphamide</td>
<td>1.00 0.93</td>
<td>0.80 0.79</td>
<td>0.70 0.67</td>
<td>0.03</td>
</tr>
<tr>
<td>Adriamycin</td>
<td>1.00 0.95</td>
<td>0.90 0.86</td>
<td>1.50 1.45</td>
<td>0.02</td>
</tr>
<tr>
<td>Vincristine</td>
<td>1.00 1.00</td>
<td>0.85 0.90</td>
<td>1.50 1.44</td>
<td>0.28</td>
</tr>
<tr>
<td>Prednisone</td>
<td>1.00 0.99</td>
<td>0.80 0.79</td>
<td>1.00 1.00</td>
<td>0.34</td>
</tr>
</tbody>
</table>

<sup>*p-values were determined by an approximate likelihood analysis of a Cox proportional hazards model. Actual RDI was analyzed for survival as a continuous variable across groups (115 patients), separately for each drug.  

<sup>0</sup>Standardized to CHOP
Table 3. FACTORS PROGNOSTIC FOR SURVIVAL

<table>
<thead>
<tr>
<th>Variable</th>
<th>p-value**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance status</td>
<td>&lt; 0.01*</td>
</tr>
<tr>
<td># Extranodal sites</td>
<td>&lt; 0.01*</td>
</tr>
<tr>
<td>Stage</td>
<td>0.02*</td>
</tr>
<tr>
<td>Marrow involvement</td>
<td>0.02</td>
</tr>
<tr>
<td>B symptoms</td>
<td>0.04</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>0.06*</td>
</tr>
<tr>
<td>Sex</td>
<td>0.07</td>
</tr>
<tr>
<td>LDH</td>
<td>0.08*</td>
</tr>
<tr>
<td>Age</td>
<td>0.17*</td>
</tr>
<tr>
<td>Histology</td>
<td>0.68</td>
</tr>
<tr>
<td>Bulk</td>
<td>0.98*</td>
</tr>
<tr>
<td>Consolidative radiotherapy (n=53)</td>
<td>0.01°</td>
</tr>
<tr>
<td># Cycles to CR (n = 96)</td>
<td>0.61*</td>
</tr>
</tbody>
</table>

** p-values were determined by an approximate likelihood ratio statistic for a Cox proportional hazards model.

* Indicates when the corresponding factor was analyzed as a continuous or ordered variable.

° Analysis included patients with initial bulky mass ≥ 5 cm who did (n=30) or did not (n=23) receive radiation therapy following achievement of complete remission.
A

PROBABILITY (
%
)

TIME (YEARS)

SUMC May-88

115 PTS

B

1. CHOP 87% CR
2. (M)BACOD 68% CR
3. MACOP-B 94% CR

Gehan P-values
1 vs. 2 0.968
1 vs. 3 0.501
2 vs. 3 0.347
A

RDI ADRIAMYCIN

≤ .85

10

LDH

< 1.1 X

4

≥ 1.1 X

6

>.85

48

ECOG PERFORMANCE

0-2

45

3-4

3

B

1. RDI-A > 85% AND ECOG 0-2 OR RDI-A ≤ 85% AND LDH < 1.1 X
2. RDI-A ≤ 85% AND LDH ≥ 1.1 X OR RDI-A > 85% AND ECOG 3-4

Gehan P-values
1 vs. 2 < 0.001

SUMC May-88