

Review

Model based development of the BEACOPP regimen for advanced stage Hodgkin's disease

M. Loeffler,¹ D. Hasenclever¹ & V. Diehl² for the German Hodgkin's Lymphoma Study Group

¹Institute for Medical Informatics, Statistics and Epidemiology (IMISE), University of Leipzig; ²Clinic I for Internal Medicine, University of Cologne, Germany

Summary

In this article we summarize the theoretical arguments which led us in the German Hodgkin's Disease Study Group to introduce the BEACOPP-regimen and to initiate a large randomised trial to investigate the role of moderate dose escalation in the treatment of advanced stage Hodgkin's disease. Although some indications for a role of dose were available in the early 1990s no prospective randomised trial had been undertaken.

In order to obtain an impression of the shape of the essential dose response characteristic we developed a novel statistical model that could be used to analyse a set of data in which dose variations had occurred. The model took tumour growth and chemotherapy effects into account. The model could be applied to clinical data on tumour control and treatment given in a patient population. The model was fitted to the data of 706 patients which had received COPP/ABVD-like regimens. It revealed considerable heterogeneity in chemosensitivity and a positive slope of the doseresponse relationship. The model was used to simulate the effect of various treatment strategies with dose escalation and schedule changes. On the basis of such

simulations we predicted that shortening cycle intervals from 4 to 3 weeks should lead to small benefits (about 3% in five-year tumour control rates). In contrast, we predicted that a moderate average dose escalation by 30% of a standard chemotherapy would lead to a potential benefit in the order of 10%–15% in tumour control at five years. Subsequently we searched for a treatment scheme that would permit such a dose escalation.

The BEACOPP-scheme was invented to allow the three major myelotoxic substances (cyclophosphamide, adriamycin and etoposide) to be given in the beginning of a cycle. These three substances were then subject to dose escalation in a dose finding trial. G-CSF was introduced to compensate for the myelotoxic effects. The dose level found feasible for a large multicentre setting turned out to be in the required magnitude. The HD9 trial of the GHSG was then initiated to examine whether the predicted dose response curve really exists.

Key words: chemotherapy, dose response relationship, Hodgkin's disease, randomised clinical trials, statistical models, tumour growth kinetics

Introduction

About 60% of the patients with advanced stage Hodgkin's disease can be cured with standard chemotherapy-like ABVD, MOPP/ABVD or MOPP/ABV. In 1991 the GHSG started to design the HD9 trial for advanced Hodgkin's disease. At that time we felt that one should try to improve this rate by some kind of treatment intensification of established regimens. This was deemed possible as growth factor support should permit to reduce the myelotoxic side effects of some of the anti-tumour drugs. It was, however, unclear how best to achieve this goal. Some proponents were pursuing the idea to select a small group of patients with very poor prognosis and treat them with high dose chemotherapy and autologous bone marrow or peripheral stem-cell transfer. However, detailed analyses of the prognostic factors had not resulted in factors that would permit to reliably discriminate such a group [7]. Hence an alternative option was to escalate standard treatment regimens to a moderate extent and to apply them to all

advanced stage patients. However, it was not clear whether it would be better to achieve treatment intensification by increasing the dose or by shortening intervals between cycles or both. At that time there was no concept available to decide which of these strategies should be most promising for a specific tumour entity. It became the objective of a biometric project to provide a model based approach to this problem.

Basic model concepts

The starting point of our considerations was the Skipper model of chemotherapy [4, 9] (Figure 1). The time course of the tumour volume during treatment depends on two factors: the tumour regrowth during the treatment intervals and the chemosensitivity of tumour cells at each treatment shot.

Clearly interval shortening becomes more relevant the more rapid a tumour grows, i.e., it depends on the slope of the tumour regrowth. The importance of esca-

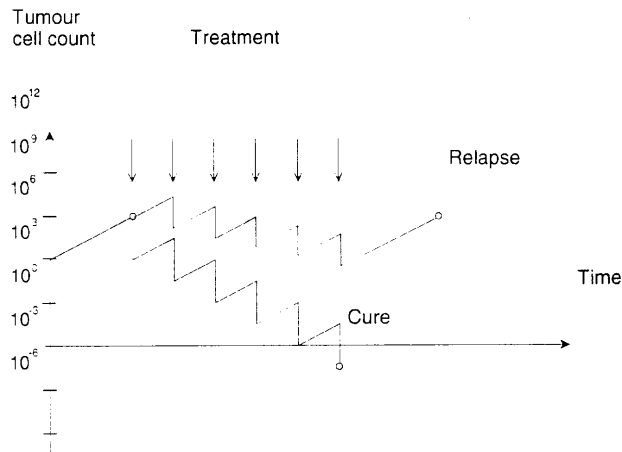


Figure 1. The classical model of chemotherapy effects and regrowth during treatment intervals and following treatment.

lated doses depends on the sensitivity of the tumour to chemotherapy over the entire treatment duration. If there is qualitative resistance higher doses or more cycles will not help.

For most tumour entities cell counts, doubling times and chemosensitivity are difficult to measure directly in individual patients. In addition, these characteristics clearly vary considerably between patients. We will show here that one can nevertheless obtain a meaningful description of tumour growth and chemosensitivity on a population scale if one accepts some simplifying model assumptions. Subsequently, we will show how to estimate the distribution of tumour latency times (i.e., the time a tumour requires to grow from one cell to clinical detection) and the distribution of (relative) chemosensitivity from clinical data.

Model assumptions

(1) Assumptions on tumour growth

We assume that the tumours grow exponentially in each patient with a fixed growth rate during treatment intervals and after treatment. To model the heterogeneity between patients we assume that the growth rates vary according to a lognormal distribution. This is in line with the results of measurements of tumour doubling times in the literature [1].

(2a) Assumption on chemotherapy effect

We assume that the efficacy of a treatment depends on two variables: relative average dose given and treatment duration. We measure treatment duration as the time from the first to the last shot of a chemotherapy regimen. We quantify the relative average dose (RAD) as the total dose given relative to the total doses of the standard treatment averaged with equal weights over all drugs used.

(2b) The concept of net treatment efficacy

In one patient, the overall treatment results in a certain net tumour cell kill at the end of the treatment. If the net

tumour cell kill is greater than the initial tumour cell burden the patient is cured. Otherwise the remaining tumour cells grow until a relapse is detected. Since the absolute tumour cell counts are difficult to measure, we introduce the ratio:

$$TE = \log(\text{net tumour cell kill}) / \log(\text{initial tumour cell count})$$

This ratio quantifies the treatment efficacy of a treatment in a patient. It has a simple clinical interpretation. If the treatment efficacy is zero, treatment has no effect, if it is less than one the patient is destined to relapse. If the net treatment efficacy is greater than 1, the patient is cured and if it is considerably greater than 1, the patient is overtreated by the given therapy.

(2c) Potential treatment efficacy

For a given tumour growth rate the treatment duration determines the amount of regrowth during treatment intervals. To achieve a certain net treatment efficacy the dose must be sufficient in order to eliminate the regrowth occurring during treatment intervals. Thus the net treatment efficacy results as a difference of the potential treatment efficacy (PTE) of the therapy (i.e., as if there were no regrowth) and the regrowth (measured on a log scale relative to the initial tumour cell burden). This fundamental relationship quantifies the balance between the effect of dose and regrowth.

(2d) Assumptions on chemosensitivity

We express the chemosensitivity of a tumour by the potential treatment efficacy of the standard treatment. We take the heterogeneity between patients into account by assuming that the chemosensitivity varies according to a gamma distribution in the population. This distribution is assumed to be independent of the distribution of the latency times.

(2e) Total dose and potential treatment efficacy

In order to take the dose response relationships into account, we need to specify the relation between relative average dose and potential treatment efficacy. We assume a simple power law introducing the residual chemosensitivity (RCS) as the slope of this function at the standard dose level ($PTE = RAD^{RCS}$). $RCS = 0$ implies no chemosensitivity. $RCS = 1$ implies a linear relationship, while $RCS < 1$ implies a weaker impact.

(3) Model estimation

The above assumptions define a statistical model of tumour growth and chemotherapy with five unknown parameters (two parameters for the lognormal distribution of latency times, two for the gamma distribution of chemosensitivity, and the slope parameter RCS). These parameters can be estimated from clinical data by maximum likelihood methods. For details see [5].

Table 1. Relative dose and duration of treatment in the HD3 and HD6 trials of the GHLSG.

	Relative total dose	Treatment duration (months)
Minimum	0.12	0.5
25% quartile	0.85	7.4
Median	0.94	8.3
75% quartile	0.98	9.4
Maximum	1.30	14.0

Model results

The model was fitted to the data of 706 stage IIIB-IV Hodgkin's disease patients treated with COPP/ABVD or COPP/ABV/IMEP polychemotherapy in the trials HD3 and HD6 of the German Hodgkin's Lymphoma Study Group.

There was considerable heterogeneity in treatment given with respect to the relative average dose and treatment duration (Table 1). Figure 2 shows the estimated latency time distribution. It has a median of 2 years a mean of 3 years and spreads out to about 10 years.

Figure 3 shows the estimated distribution of the net treatment efficacy (TE) with standard treatment. Negative values of TE correspond to progression during therapy (12%). Values between 0 and 1 correspond to PR and CR patients which are not cured but will experience a relapse (30%). A net treatment efficacy TE larger than 1 corresponds to patients that enter a continuous complete remission (cure, 58%). It is remarkable that the mode of the distribution is near the cure threshold. This implies that a relevant percentage of patients could be cured by some kind of treatment intensification. On the other hand about 20% of the patients are estimated to have a treatment efficacy beyond 2, which suggests that they may be overtreated.

The major model finding was a significant slope of the dose response relationship. More precisely the slope of the potential treatment efficacy as a function of the relative average dose (at standard therapy dose levels) was estimated to be 0.98. A lower 95% confidence limit for the slope parameter RCS is (0.48, 1.48). Thus there was a very clear indication of a positive slope of the dose response in this data set. Given these parameter estimates one can calculate how the rate of patients free of disease after five years depends on the relative average dose. The shape of this relationship is given in Figure 4. Two curves are presented which are obtained by attributing the values 1.0 or 0.7 to the parameter RCS. The steeper curve thus corresponds to the dose response estimated from the data analysis while the second curve describes the slope for a more conservative value of RCS. Both curves are clearly bended flattening somewhat at higher doses. However, an increase of the relative average dose by 30% should permit to obtain clear increases in the cure rates.

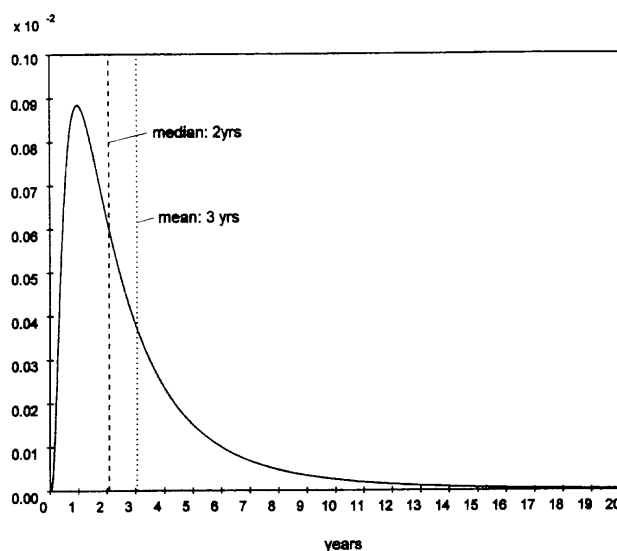


Figure 2. Model estimate of the tumour latency time distribution derived from a fit to data of 706 patients with advanced stage Hodgkin's disease.

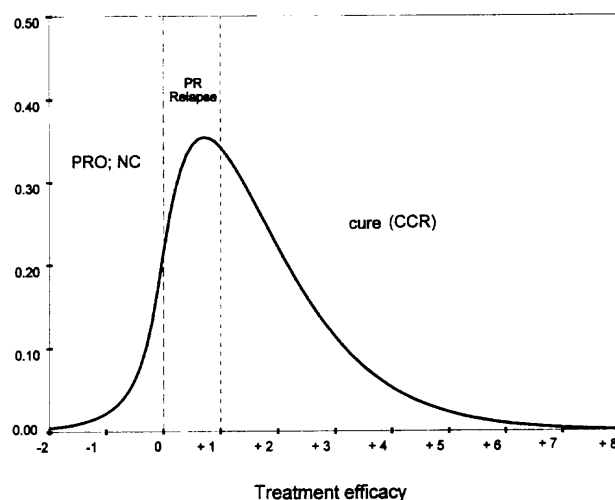


Figure 3. Model estimate of the net treatment efficacy derived from a fit to data of 706 patients with advanced stage Hodgkin's disease.

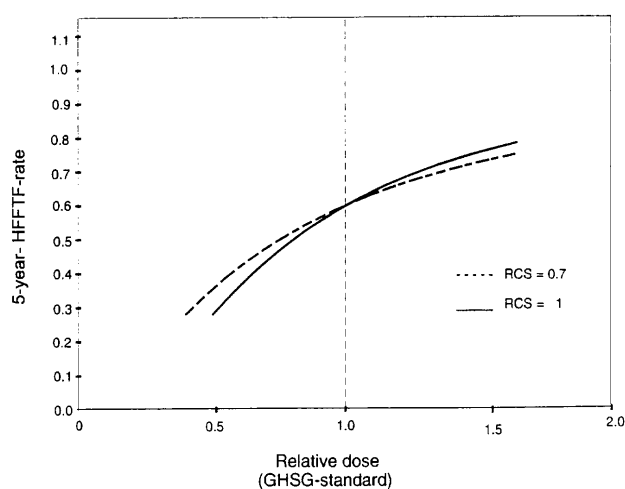


Figure 4. Model estimate of the dose response relationship for the five-year tumour control rate depending on the relative average dose (COPP/ABVD = 1.0).

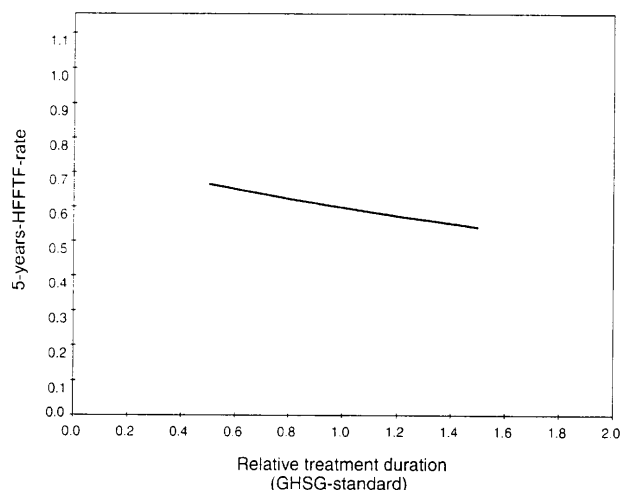


Figure 5. Model prediction for the importance of treatment duration on treatment outcome. It is noticeable that shortening the duration should permit a moderate gain while prolongation should be associated with a loss of tumour control. The effect is, however, less pronounced than the dose effect shown in Figure 4.

Model predictions for moderately escalated treatment strategies

In order to discriminate between the more promising intensification strategies model simulations were undertaken. For this purpose all estimated parameters were kept fixed and only the relative average dose of the standard chemotherapy and the treatment intervals were modified. The model predicted a benefit in the order of only 3% in five-year tumour control rates if one would selectively shorten treatment duration by 25% (i.e., cycle intervals from 4 to 3 weeks; Figures 5 and 6). In contrast, a moderate dose escalation of 30% was predicted to increase long term tumour control rates by 10% in the conservative case (RCS = 0.7) and 14% in the realistic case (RCS = 1.0). For the latter case Figure 6 gives the model predictions, which were already published earlier [2, 5, 6].

The architecture of the BEACOPP regimen

Based on the above insights it was apparent that one would primarily try to increase dose in a standard treatment protocol for primary treatment. We therefore examined the COPP/ABVD-scheme that was used in our group for the possibility to introduce G-CSF administration in such a way that one could escalate the doses of the most myelotoxic substances. In this regard the COPP-cycle is not very suitable as procarbazine is given over 14 days. Hence G-CSF, which should not be given simultaneously with cytostatics, could not be introduced before day 15. This would clearly be too late. Furthermore cyclophosphamide is given on days 1 and 8. The second administration would hence be given at the time of the neutropenic nadir of the first administration making dose escalation difficult. Similar difficulties also apply to the ABVD-scheme which has a very tight time

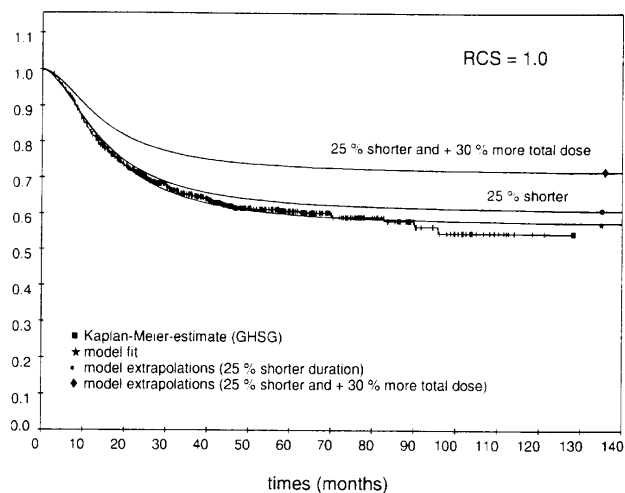


Figure 6. Model fit and model predictions for moderately escalated regimens. The plot is generated under the assumption that the slope parameter (RCS) is 1.0. If one assumes that etoposide in the BEACOPP baseline regime is equipotent to DTIC in the ABVD scheme this figure can be interpreted as a model prediction for the outcome of the HD9 trial. It predicts an increase of the five-year tumour control rate of about 14%.

frame not permitting enough flexibility for escalation of doses. It hence became obvious that the architecture of a scheme suitable for examining the role of dose would have to differ substantially from the conventional schedules. Furthermore it was clear that one should only escalate such substances where the induction of neutropenia would be the major side effect to permit a compensation by G-CSF. In this context we decided to introduce etoposide as a substance with myelotoxic effects and skipped DTIC whose relevance we doubted.

Finally we decided that the major myelotoxic substances should be given in the first three days and that G-CSF should be given following day 8. Hence all other substances would have to be given within the first week. The BEACOPP-scheme in baseline dosage resulted from these considerations (see Table 2). The baseline BEACOPP-scheme was then tested in a feasibility study and proved to be effective and feasible [3].

BEACOPP dose finding trial

It was then apparent that we had to undertake a dose finding trial to titrate the maximum acceptable dose of the BEACOPP scheme. As we were interested in finding a scheme that could be used in a multicenter setting it was important to apply rather strict criteria to select the acceptable dose level. The dose finding trial was initiated to find the dose for which the frequency of unacceptable toxic events was not higher than 1/3. As an unacceptable toxic event we defined a neutropenia below 1000 leukocytes/ml for more than three days or a thrombocytopenia below 50,000 platelets/ml or the occurrence of an infection with WHO grade 3 or 4. The dose finding trial was then initiated with a novel technique to permit recruitment of several patients simultaneously. Toxic events after each cycle were monitored carefully

Table 2. Treatment schedules and dose levels.

BEACOPP regimen		Baseline dose/day	Relative dose levels						
			Baseline level	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
Cyclophosphamide	i.v. d1	650 mg/m ²	1	1.23	1.46	1.69	1.93	2.16	2.39
Adriamycin	i.v. d1	25 mg/m ²	1	1.40	1.40	1.40	1.40	1.40	1.40
Etoposide	i.v. d1-3	100 mg/m ²	1	1.25	1.50	1.75	2.00	2.25	2.50
Procarbazine	p.o. d1-7	100 mg/m ²	1	1	1	1	1	1	1
Vincristine	i.v. d8	2 mg	1	1	1	1	1	1	1
Bleomycin	i.v. d8	10 mg/m ²	1	1	1	1	1	1	1
Prednisone	p.o. d1-7	40 mg/m ²	1	1	1	1	1	1	1
Average relative dose			1.00	1.13	1.19	1.26	1.33	1.40	1.47

and continuously such that any new patient entering to the trial was allocated to the highest dose level considered safe up to that time point. The dose levels were prespecified as indicated in Table 2. Sixty patients were entered into the trial Figure 7 shows the sampling path of these patients indicating that a dose level between 4 and 5 should be feasible and safe [10]. An analysis of the dose reductions and treatment delays in this cohort of patients also revealed that an elevated relative average dose could be maintained throughout the entire treatment. Hence it was obvious that a BEACOPP scheme with dose level 4 would fulfill the requirements for a prospective clinical trial to investigate the dose response relationship predicted by the statistical model introduced above.

The HD9 trial

Based on these considerations the German Hodgkin's Disease Study Group initiated the HD9 trial. It compares BEACOPP in baseline dose with BEACOPP in escalated dose (level 4). In each scheme eight cycles were scheduled in three-week intervals. To control for the changes made in the BEACOPP-regimen we decided to keep our internal standard treatment COPP/ABVD as a control arm. This would permit to examine to some

extent whether the shortening of the treatment duration and the introduction of etoposide would create any differences.

In 1993 the GHSg started the threearmed HD9 trial. The trial was designed to include 1000 patients in order to have the statistical power to detect the predicted benefit of about 12%–13%. Results of an interim analysis of this trial are given in another article in this issue.

Discussion

Here we summarized the model based rationale which led us to the design of the BEACOPP scheme for advanced stage Hodgkin's disease. A statistical model of chemotherapy and tumour growth provided strong suggestions for the existence of a relevant dose response relationship for standard chemotherapy. Based on a fit to clinical trial data of patients who had received the same treatments in slightly varying doses and durations we anticipated that a moderate increase of the average dose by 30% should yield a remarkable increase in long-term cure rates. The estimated dose response curve would suggest that such a dose escalation would be feasible for a broad cohort of patients and no discrimination for unfavourable prognostic subgroups would be necessary.

Clearly a model is a simplification of reality. No model substitutes for a well planned prospective randomized clinical trial. It is therefore necessary to outline the most essential model limitations.

There are several conceptual limitations. Thus it appears rather arbitrary to assume that all drugs in the standard treatment have the same antitumour effect (i.e. equal relative weights). We decided to consider etoposide given in three doses of 100 mg/m² to be equally effective as 25 mg/m² of adriamycin. If one accepts this assumption the model predictions given above (Figure 6) can be considered as predictions for the HD9 trial. If one assumes that etoposide has a higher weight one would expect that the BEACOPP schemes would be even more

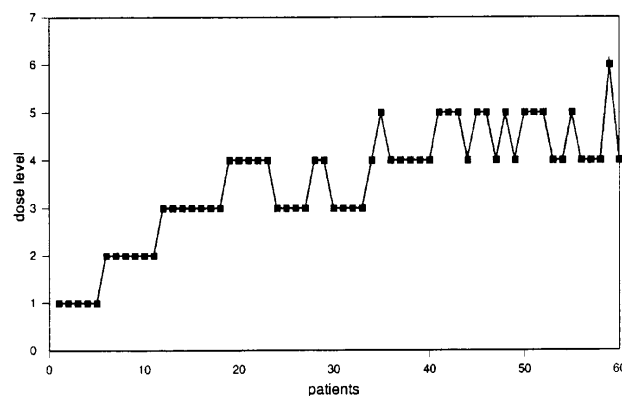


Figure 7. BEACOPP dose finding trial. Here we show the sampling path of the dose finding trial which included 60 patients. Each symbol represents the dose level one patient was allocated to.

effective in the HD9 trial leading to larger differences compared with the COPP/ABVD standard.

Another model limitation is that we assumed homogenous behaviour of the population with regard to chemosensitivity. It is, however, biologically more likely that there are subgroups of tumour entities which have a higher chemosensitivity than others. Thus the model may underestimate the sensitivity of some tumours while it overestimates the sensitivity of others. It is not clear whether this on average leads to the same overall outcome.

Furthermore the model was fitted to a dataset in retrospect. Variations of dose and duration were not planned but occurred for a variety of reasons (e.g., dose reductions due to toxicity, delays due to patients wish, etc). Thus selection processes may have occurred but we believe that most of them should tend to reduce the estimate of the slope of the dose response relationship.

All these limitations need to be reconsidered if the data of the HD9 trial are compared with the above model predictions. It is very likely that the model prediction was too conservative and that more pronounced effects actually occur. In this case a more sophisticated model would have to account for the real weights of different drugs and for the correlation between chemosensitivity and prognostic subgroups. Furthermore fitting the model to data of the prospective HD9 trial should then give a more reliable estimate of the dose response relationship.

On the other hand the model was very helpful to discriminate different trial strategies. The main result was to focus on moderate dose escalation for all advanced stage patients rather than shortening of treatment duration or entering high dose strategies with stem cell transfer in primary treatment.

References

1. Charbit A, Malaise EP, Tubiana M. Relation between the pathological nature and the growth rate of human tumors. *Eur J Cancer* 1971; 7: 307-15.
2. Diehl V. Doseescalation study for treatment of Hodgkin's disease. The German Hodgkin Study Group. *Ann Hematol* 1993; 66 (3): 139-40.
3. Diehl V, Sieber M, Ruffer U et al. for the German Hodgkin's Lymphoma Study Group (GHSG). BEACOPP: An intensified chemotherapy regimen in advanced Hodgkin's disease. *Ann Oncol* 1997; 8: 143-8.
4. Frei E, Canellos GP. Dose: A critical factor in cancer chemotherapy. *Am J Med* 1980; 69: 585-94.
5. Hasenclever D. Tumorwachstum und Wirkung von Chemotherapie. Dissertation. University of Cologne, Germany 1995.
6. Hasenclever D, Loeffler M, Diehl V for the German Hodgkin's Lymphoma Study Group. Rationale for dose escalation of first-line conventional chemotherapy in advanced Hodgkin's disease. *Ann Oncol* 1996; 7 (4): 95-8.
7. Loeffler M, Mauch P, MacLennan K et al. Review on prognostic factors. *Ann Oncol* 1992; 3 (Suppl 4): 63-6.
8. Redmond C, Fisher B, Wieand HS. The methodologic dilemma in retrospectively correlating the amount of chemotherapy received in adjuvant therapy protocols with disease-free survival. *Cancer Treat Rep* 1983; 67 (6): 519-26.
9. Skipper HE. Kinetic behavior *versus* response to chemotherapy. *Natl Cancer Inst Monogr* 1971; 34: 2-14.
10. Tesch H, Diehl V, Lathan B et al. for the German Hodgkin's Lymphoma Study Group (Chairman Prof. V. Diehl). Moderate dose escalation for advanced stage Hodgkin's disease using the BEACOPP scheme and adjuvant radiotherapy - a study of the German Hodgkin's Lymphoma Study Group (GHSG). *Blood* (in press).

Correspondence to:

M. Loeffler, MD, PhD
 IMISE University of Leipzig
 Liebigstraße 27
 D-04103 Leipzig
 Germany
 E-mail: Loeffler@imise.uni-leipzig.de