Early response to chemotherapy: a surrogate for final outcome of Hodgkin’s disease patients that should influence initial treatment length and intensity?

P. Carde1, S. Koscielný1, J. Franklin2, U. Axiphorm3, J. Raemaekers4, V. Diehl5, B. Aleman5, O. Brosteanu6, D. Hasenclever7, O. Oberlin1, N. Bonvin1 & M. Björkholm3

1Department of Medicine, Institut Gustave Roussy, Villejuif, France; 2Working Group for Biometry, Department of Internal Medicine 1, University of Cologne, Cologne, Germany; 3Department of Medicine, Karolinska Hospital, Stockholm, Sweden; 4University Medical Center Nijmegen, Nijmegen; 5Radiation Oncology, The Netherlands Cancer Institute, Antoni van Leeuwenhoek Huis, Amsterdam, The Netherlands; 6KKSL and 7IMISE, University of Leipzig, Leipzig, Germany

Background: Early adjustment of treatment may benefit the patient. In order to guide treatment adjustment, use of early response (ER) or early complete response (ECR), judged after the few initial cycles of chemotherapy, is common in pediatric and also adult Hodgkin’s and non-Hodgkin’s studies. Paradoxically, almost no data support this strategy.

Patients and methods: The influence of ECR on outcome was evaluated in three series of advanced Hodgkin’s disease (HD), leading to a series of questions.

Results: The 1982 EORTC study assessed prospectively the time frame needed to reach an apparent complete response (CR) through repeated tumor measurements. In patients assessed at mid-treatment before the fifth cycle, both 15 year freedom from progression (FFP) and overall survival (OS) were superior in ECR patients compared with other patients continued on the same treatment (61% versus 37%; P <0.001). A series of questions arise from these observations. Question 1: is the shortening of treatment detrimental? In a randomized Swedish trial, in one arm treatment was shortened in patients evaluated from the fifth cycle as ECR as compared with the standard eight cycles arm, 10 year cause-specific-survival (CSS) was 53 versus 69% [not significant (ns)]; 10 year OS 49 versus 58% (ns). Conversely, in the EORTC 20884 study, ECR patients given only six cycles did as well as patients entering CR later and, for this reason, given eight cycles (identical 6 year event-free survival 75%). Question 2: is early treatment adaptation in patients who failed to reach ER beneficial? In the French MDH 90 trial, 15% of children failed to reach ECR after four cycles; in these children only, anthracyclines plus alkylating agents were given and the dose of radiotherapy increased, improving the results observed in the previous trial. In the EORTC 20884 study, patients who failed to reach an ECR were switched earlier to involved field RT: their results matched those of ECR patients, at the difference of the previous trial. Question 3: is ER a predicting factor that can be used with any type of treatment? Probably not, based on the German Hodgkin’s Lymphoma Study Group trial HD 9: ECR is highly dependent on specific interval from treatment start and on treatment intensity.

Discussion: More general questions stem from these results. Question 4: is the definition of ER secured? With conventional imaging, the different methods for response assessment at end treatment also lead to different response rates; the assessment in the middle of treatment itself and the use of newer imaging techniques may further increase the variation. Indeed, question 5 is: is ER a concept based on any biology? Correlation to markers, 99mTc uptake, PET and hematological tolerance might help to pinpoint how and why ER represents a surrogate for final outcome.

Conclusion: ER is a surrogate for final outcome, reflecting both tumor burden and activity. This predictability may, and possibly should, impact on treatment.

Key words: CT scan, early response, imaging, Hodgkin, minimal residual disease, PET scan, prognosis, response, tumor markers, (99mTc)-MIBI scintigraphy

*Correspondence to: Dr P. Carde, Department of Medicine, Institut Gustave Roussy, F-94805 Villejuif, France. Tel: +33-1-42-11-43-21; Fax: +33-1-42-11-52-70; E-mail: carde@igr.fr

© 2002 European Society for Medical Oncology
chemotherapy regimen, the number of chemotherapy cycles, the field sizes and the dose of radiation within these fields are subjects of debate.

The question of whether EF-RT can be substituted by IF-RT is answered by the GHSG HD 8 trial, which started in 1993 [3]. After two cycles of cyclophosphamide–vincristine–procarbazine–prednisone (COPP/ABVD), patients with unfavorable stages I and II disease and selected patients with stage IIIA disease were randomized between the standard arm 2× COPP/ABVD + 30 Gy EF (+10 Gy bulk) and the experimental arm 2× COPP/ABVD + 30 Gy IF (+10 Gy bulk). The final analysis of this trial was conducted in July 2001. All 1136 qualified patients were evaluable. Overall results showed an FFTF of 83% and a survival of 91% (at 5 years). For reasons such as progressive disease or protocol violations during the induction chemotherapy, 70 patients were not informative for the arm comparison. All of the 1066 informative patients were evaluated according to their randomized treatment arm. No statistically significant difference has been observed with regard to CR, PR, no change (NC), progress and relapse rate. Conclusively, a statistically significant difference has not been apparent between both treatment arms for FFTF, HL-specific FFTF and overall survival.

Consequently, combined-modality therapy including IF-RT was chosen as a concept for the current HD 11 trial for intermediate stage HL. Using a 2 × 2 factorial design, the optimal chemotherapy regimen and the optimal dose of radiotherapy will be investigated. Thus, the HD 11 trial compares ABVD with bleomycin–etoposide–doxorubicin–cyclophosphamide–vincristine–procarbazine–prednisone (BEACOPP) and, in addition, aims to answer the question of whether to further reduce the radiotherapy dose (4× ABVD + 30 Gy IF versus 4× ABVD + 20 Gy IF versus 4× BEACOPP + 30 Gy IF versus 4× BEACOPP + 20 Gy IF). The second interim analysis of HD 11 was carried out in June 2001. A total of 480 (83%) of 580 randomized and qualified patients were evaluable. Overall results showed an FFTF of 92% and a survival of 99% (median observation time 18 months).

Conclusions

The last two GHSG study generations for treatment of early and intermediate stage HL have been strongly influenced by the need to drop EF-RT. The rationale was to reduce the formerly high relapse rate in early stage disease treated only with radiotherapy and the fatal long-term toxicity. In the HD 7 trial, the GHSG was able to show that two cycles of ABVD prior to EF-RT can significantly reduce the relapse rate and thus improve disease-free survival. EF-RT has therefore been excluded in the current HD 10 trial, and the main question is to determine the number of ABVD cycles necessary before IF-RT. In intermediate stage HL the HD 8 trial has proven that after two cycles of COPP/ABVD, EF-RT can be substituted by IF-RT without loss of efficacy.

The current GHSG HD 11 trial is investigating the question of what is the optimal chemotherapy regimen prior to IF-RT by introducing the new BEACOPP regimen in its baseline version with a reduction of radiotherapy dose.

References

Introduction

The interval between minimal treatment required to control advanced Hodgkin’s disease (HD) and the maximal tolerated treatment is narrow. Prediction of outcome would be important to adjust treatment before it ends, when too late. Moves in two opposite directions could be beneficial: treatment made shorter or lighter to prevent morbidity or late toxicity; or treatment modified or intensified to improve HD control. Initial patient and disease characteristics are not ideal prognostic factors for treatment adaptation as they are obtained too early [1]; conversely, the outcome is already quite gloomy when failure to achieve complete remission (CR) or early relapse are evident [2, 3]. Non-Hodgkin’s lymphoma (NHL) patients who entered early CR (ECR), in just three cycles of chemotherapy (CT), were shown to have more durable remissions than those entering CR after five cycles [4]. In the advanced HD Stanford series [5], the dose rate delivered in the first three cycles of CT was the main factor predicting CR and outcome. In patients who achieved a CR in three cycles, another retrospective study showed a large overall survival (OS) advantage [6]. These observations prompted two cooperative groups, the European Organisation for Research and Treatment of Cancer (EORTC) (1981–86 trial) [7] and the Karolinska Hospital (1981–89 trial) [8] to assess the optimal number of CT cycles in early response (ER) patients. ECR has been assessed prospectively in the subsequent EORTC 20884 [9] and German Hodgkin’s Lymphoma Study Group (GHSG) HD 9 trials [10], allowing a comparison with the former two trials on the basis of dose intensity.

Patients and methods

Basic characteristics of the trials and statistics

Initial treatment design, patient numbers, characteristics and analysis concerning achievement of CR at the end of treatment is similar to that originally reported for each trial. Survival Kaplan–Meier estimations and comparisons by the log-rank test were performed in a similar manner.

The EORTC 1981–86 trial (IIIIB-IV). This was conducted in 207 HD patients with stage III/IV disease to assess the percentage of patients attaining CR at 2-monthly intervals from MOPP (melphalan, vincristine, procarbazine and prednisone) or MOPP/ABVD (doxorubicin, bleomycin, vinblastine and dacarbazine) initiation. In patients responding before the third cycle (n = 193), the proportion of ECR (apparent CR before cycle 5) was 81 of 193 [7]. Follow-up is now 15 years.

The Stockholm 1981–89 trial (Karolinska Hospital). This was conducted in 87 HD patients with stage II/IV disease [8]. Patients were randomized to a fixed number of MOPP/ABVD cycles (8 months) (FT group; n = 47) or to an individual/response-adapted treatment (minimum 4 months) (IT group, n = 41). Seventeen of 41 patients attained ECR (apparent clinical, biological and radiological CR before cycle 5). Follow-up is 14 years.

The EORTC 1989–2001 trial (20884). This was conducted in 763 HD patients with stage III/IV disease. ECR status before the fifth cycle was used to allow treatment alleviation in ECR patients, from eight to six cycles of MOPP/ABV. Conversely, patients still in partial remission (PR) before the seventh cycle were switched to a salvage irradiation program [9].

The French Society of Pediatric Oncology trial (MDH82). This trial demonstrated the response to primary chemotherapy as the only predictor of survival. To reduce long-term treatment complications without compromising efficacy, the MDH90 trial was based on a chemotherapy devoid of both alkylating agents and anthracyclines, followed by radiotherapy (RT) at 20 Gy for good responders, while poor responders received more intense/prolonged treatment [11].

The GHSG 1993–99 (HD 9 trial). This was conducted in 1195 patients with stage IIB/IV, comparing eight cycles of COPP (cyclophosphamide, vincristine, procarbazine and prednisone/ABVD, baseline BEACOPP (bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine and prednisone) or escalated BEACOPP [10]. In the June 2001 update for the 1072 evaluable patients, the freedom from treatment failure (FFTF) was compared between patients in CR before the fifth cycle, and those in PR. The analysis was stratified by treatment arm.

Results

In the EORTC 1981–86 trial, ECR patients (42% of total) demonstrated a markedly superior outcome in all respects compared with the other patients not yet in CR before the fifth cycle: final achievement of CR 96% versus 31%, freedom from progression (FFP) 61% versus 37% (P <0.001) and OS 61% versus 41% (P = 0.001). This advantage was not due to patient-related confounding factors: besides similar patient characteristics, ECR patients’ survival advantage came from less HD progression-related deaths (HD-specific survival = 85% versus 60%; P <0.001) rather than from other deaths (non-HD-specific survival 74% versus 71%; P = 74) (Figure 1). Setting apart these ECR patients, who were not otherwise identifiable, may allow opposed strategic moves: (i) ECR patients may benefit from an alleviated treatment, i.e. less dose intense, shorter, and/or deprived of irradiation or alkylating agents; (ii) non-ECR patients may be proposed to continue the initial treatment longer or to switch earlier to an alternative treatment. However, several questions arise.

Question 1: is the shortening of treatment detrimental?

In the Swedish trial, the following intention-to-treat results were observed in the experimental arm (IT group), as compared with the FT group, respectively: 10 year cause-specific-survival (CSS) 53 versus 69% [not significant (ns)]; 10 year OS 49% versus 58% (ns). Thus, the trend was unfavorable to the IT group; however, even when restricted to patients achieving CR: respectively, mean number of cycles 2.6 MOPP + 2.5 ABVD versus 3.7 MOPP + 3.5 ABVD (P <0.001); percentage CR 68% versus 81% (ns); 10 year CSS 68% versus 81%
(P <0.05); 10 year OS 63% versus 70% (ns). Nevertheless most of ECR patients (CR<sub>1</sub>) do well at 108 months follow-up: of the 16 patients out of the initial 41 who achieved CR<sub>1</sub> (41%), 10 of these 16 (25% of the initial 41 patients) are long-term remitters given only four cycles, compared with 90% in the FT group. These patients could not be identified by any prognostic feature (Figure 2).

Conversely, in the EORTC trial 20884, ECR patients (32% of the whole cohort) given only six cycles did as well as patients entering CR later, before the seventh cycle (CR<sub>6</sub>, 28% of the whole cohort) and given eight cycles: 6 year event-free survival (EFS) was identical (75%), and the 6 year OS was 82% in ECR versus 85% in CR<sub>6</sub> patients.

**Question 2: is early treatment adaptation in patients who failed to reach ER beneficial?**

In the MDH-90 study, treatment adaptation allowed poor responders to achieve a 91% EFS, in comparison with a lower percentage (18%) in the previous trial [11].

In the EORTC trial III1IBV, 111 of 207 patients not achieving an ECR before the fifth cycle had no treatment adaptation, which resulted in a 43% 5 year FFP and 65% OS; in contrast, in the subsequent EORTC 20884 trial, 460 of 676 patients in the same situation (including 181 CR<sub>0</sub>, 241 PR<sub>0</sub>, 38 failures) who had earlier treatment adaptation (namely 241 PR<sub>0</sub> treated with involved field 30 Gy irradiation) and enjoyed a much better outcome: 79% 5 year FFP and 83% OS; furthermore, PR<sub>0</sub> patients share the same excellent fate as ECR patients in the 20884 trial [12, 13] (Figure 3).

**Figure 1.** Survival curves in 207 stage IIIB and IV patients, EORTC 1981–1986 trial; 81 patients achieved an early complete response (RC), 126 did not (not RC) [12]. (A) Survival from non-Hodgkin’s disease mortality is alike in the RC and not RC patients, suggesting that initial patient characteristics are well balanced in the two groups. (B) Survival from Hodgkin’s disease mortality is 85% in RC patients compared to 60% in the not RC patients, P <0.001.

**Figure 2.** Survival curves in 89 stage IIIB–IV Hodgkin’s patients, Stockholm 1981–1989 trial [8]. Patients were randomized to a fixed 8 cycles/8 months of MOPP/ABVD (upper curve) or to an individual/adapted no. cycles/minimum 4 months (lower curve). Survival was not significantly different.

**Figure 3.** Event-free survival curves in 736 stage IIIB and IV patients, EORTC 1981–1986 trial #20884 [9]. In 243 patients still in partial remission (PR) (arrow) before the seventh cycle who were successfully switched to a salvage irradiation program; Event-free survival (and overall survival not shown) match those of patients reaching CR, whether randomized or not.
Question 3: can ER be used as an outcome-predicting factor with any type of treatment?

Based on the HD 9 GHSG trial, this is not the case. First, one notes that the percentage of ECR does not vary much according to treatment intensity (38%, 31%, 32%, respectively), a percentage close to those observed in the EORTC and Swedish trials. Then, comparing ECR patients before the fifth cycle with those in PR indeed showed a trend for a superior FFTF in ECR patients overall (P = 0.06); however, the trend came almost exclusively from the COPP/ABVD arm data (P = 0.048) (J. Franklin, U. Paulus, V. Diehl et al., personal communication). ECR outcome predictability may be erased, as any other prognostic factor, under conditions of very intense regimens (Figure 4). Finally, although at a given time-point ER selects a similar proportion of favorable ‘ECR’ patients, the value of ECR considered as an outcome predictor depends upon each specific treatment strategy, since some of the patients not reaching ECR may benefit from the more intense regimens.

Discussion

Additional, more general questions stem from the information above.

Question 4: is definition of ER secured?

There are many ways to assess response, including as a tumor burden volume or as tumor ‘activity’. First, there is no consensus, either about the standard initial measurement of the disease, or the later measurement of the response. Indeed, although the Cotswolds criteria [14] are supposed to be used, they do not fit well with the most appropriate criteria for response evaluation [15]. Indeed, response rates varying from 6% to 28% have been reported at end treatment, depending on the definition of ‘normal’ lymph node [16]. Secondly, these percentages are overestimated by up to two-fold, if response confirmation is not delayed from day 28 onwards, an observation pertinent to ER evaluation, apparently more influenced by the time elapsed since treatment start than by treatment type or intensity. In fact, definition of ER all depends on the initial definition of tumor burden (TB) and activity, neither of which is standardized. The best attempt to-date to normalize definition of TB has been the addition of the tumor volumes through angio-computed tomography (CT) and ultrasound (US) normalized to body size in cm²/m² [17]. A much simpler method, using unidimensional criteria, is yet to be assessed in lymphoma [18].

Attempts have been made to assess response as tumor ‘activity’. Gallium scan helps, after treatment has ended, to assess residual mediastinal masses [19], but not for prognosis evaluation or strategy selection during treatment. Fluorine-18-fluoro-2-deoxy-D-glucose positron emission tomography (PET) scan compares favorably with CT and gallium scans [20]; however, as understood very early [21], a more interesting issue is to evaluate the timing of PET conversion during treatment in the prospect of long-term outcome. Such investigations are in progress at several institutions. While not yet confirmed with PET scan, it has been suggested that chemotherapy response could be predicted by technetium 99m (99mTc)-MIBI scintigraphy. This, observed in breast carcinoma patients through the snapshot provided by the 99mTc scintigraphy, suggests that TB volume, tumor activity and likelihood of tumor response are inter-related inside the mech-

**Figure 4.** Freedom from treatment failure (FFTF) curves in 1195 IIb-IV patients, GHSG 1993-1999 (#HD-9) trial [10; J. Franklin et al. personal communication 2001]. FFTF was compared between patients in early complete response (CR₄) before the fifth cycle and those in PR (PR₄). (A) When stratified by treatment arm, the analysis favored the CR₄ only in the less intense standard COPP/ABVD arm. (B) In the more intensive escalated BEACOPP arm, FFTF was alike in CR₄ and PR₄ patients.
mildly intense initial treatment, either to keep using low-toxicity treatment in ECR patients, to intensify treatment or to switch treatment. Treatment switch when 50% to 75% of treatment has already been given may benefit most PR patients, even though some of these PRs would have later achieved a CR. Validation of ER for HD, NHL and other malignancies requires taking into account the characteristics of treatment in terms of timing and intensity. It also requires the normalization of iterative tumor burden measurement and the modeling of regression rate in correlation with biological tumor markers, drug delivery and hematological tolerance parameters.

References

12. Carde P, Koscielny S, Baars JW et al. Mid-treatment response is a powerful surrogate to predict outcome and MOPP/ABVD confirms

Figure S. Freedom from treatment failure ( FFTF) curves in 266 IIb–IV patients, GHSG 1988–1993 (HD6) trial [23]. The low hematotoxicity profile patients (lower curve), despite higher cumulative dose and intensity, demonstrated a reduced disease control and survival compared with patients with high toxicity.

Question five: is ER a concept based on any biology?

From what precedes, it appears that ER, or ECR, may be a stronger predictor than the initial patient- or disease-related characteristics usually attributed prognostic value. Not only during CT, but also during RT, at the National Cancer Institute, cumulative percentage of tumor regression has been correlated to outcome in mediastinal HD [25]. ECR is also correlated to histological response [26] and to quicker marker decrease in solid tumors [27], higher 99mTc and PET uptake, and poorer hematological tolerance, probably through various drug-resistance mechanisms, and possibly pharmacogenetic characteristics. Correlation with promising biological markers [28] has yet to be proven in Hodgkin’s lymphoma.

Conclusion

ER is a surrogate for final outcome, reflecting both tumor burden and ‘activity’. This predictability may be used with


