

Effect of initial treatment strategy on survival of patients with advanced-stage Hodgkin's lymphoma: a systematic review and network meta-analysis



Nicole Skoetz*, Sven Trelle*, Michaela Rancea, Heinz Haverkamp, Volker Diehl, Andreas Engert, Peter Borchmann

Summary

Background Several treatment strategies are available for adults with advanced-stage Hodgkin's lymphoma, but studies assessing two alternative standards of care—increased dose bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone (BEACOPP_{escalated}), and doxorubicin, bleomycin, vinblastine, and dacarbazine (ABVD)—were not powered to test differences in overall survival. To guide treatment decisions in this population of patients, we did a systematic review and network meta-analysis to identify the best initial treatment strategy.

Methods We searched the Cochrane Library, Medline, and conference proceedings for randomised controlled trials published between January, 1980, and June, 2013, that assessed overall survival in patients with advanced-stage Hodgkin's lymphoma given BEACOPP_{baseline}, BEACOPP_{escalated}, BEACOPP variants, ABVD, cyclophosphamide (mechlorethamine), vincristine, procarbazine, and prednisone (C[M]OPP), hybrid or alternating chemotherapy regimens with ABVD as the backbone (eg, COPP/ABVD, MOPP/ABVD), or doxorubicin, vinblastine, mechlorethamine, vincristine, bleomycin, etoposide, and prednisone combined with radiation therapy (the Stanford V regimen). We assessed studies for eligibility, extracted data, and assessed their quality. We then pooled the data and used a Bayesian random-effects model to combine direct comparisons with indirect evidence. We also reconstructed individual patient survival data from published Kaplan-Meier curves and did standard random-effects Poisson regression. Results are reported relative to ABVD. The primary outcome was overall survival.

Findings We screened 2055 records and identified 75 papers covering 14 eligible trials that assessed 11 different regimens in 9993 patients, providing 59 651 patient-years of follow-up. 1189 patients died, and the median follow-up was 5·9 years (IQR 4·9–6·7). Included studies were of high methodological quality, and between-trial heterogeneity was negligible ($\tau^2=0\cdot01$). Overall survival was highest in patients who received six cycles of BEACOPP_{escalated} (HR 0·38, 95% credibility interval [CrI] 0·20–0·75). Compared with a 5 year survival of 88% for ABVD, the survival benefit for six cycles of BEACOPP_{escalated} is 7% (95% CrI 3–10)—ie, a 5 year survival of 95%. Reconstructed individual survival data showed that, at 5 years, BEACOPP_{escalated} has a 10% (95% CI 3–15) advantage over ABVD in overall survival.

Interpretation Six cycles of BEACOPP_{escalated} significantly improves overall survival compared with ABVD and other regimens, and thus we recommend this treatment strategy as standard of care for patients with access to the appropriate supportive care.

Funding None.

Introduction

Hodgkin's lymphoma is one of the most common malignancies in young adults. Survival in patients with this disease has increased substantially throughout the past few decades, even for patients in advanced stages. However, how to balance risks and benefits of different treatment strategies is still a matter of controversy. The key question is whether intensified chemotherapy should be applied upfront or reserved for relapsing patients.¹ The early intensification approach aims to cure as many patients as possible with aggressive first-line chemotherapy.² This approach is the standard of care used in European Hodgkin's lymphoma study groups such as the German Hodgkin Study Group (GHSG), the Lymphoma Study Association, and the European

Organisation for Research and Treatment of Cancer. A standard regimen is increased doses of bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone (BEACOPP_{escalated}), which is associated with high progression-free and overall survival rates, but it exposes patients to substantial acute chemotherapy-related toxicity.² By contrast, the late intensification approach uses doxorubicin, bleomycin, vinblastine, and dacarbazine (ABVD) as first-line therapy. ABVD induces less acute chemotherapy-related toxicity than does BEACOPP_{escalated} and still cures most patients. However, about 30% of patients with extensive disease will relapse after initial treatment with ABVD.³ These patients might be cured by second-line treatment with high-dose chemotherapy and autologous stem-cell

Lancet Oncol 2013; 14: 943–52

Published Online

August 13, 2013

[http://dx.doi.org/10.1016/S1470-2045\(13\)70341-3](http://dx.doi.org/10.1016/S1470-2045(13)70341-3)

See Comment page 911

*Both authors contributed equally

Cochrane Haematological Malignancies Group
(N Skoetz MD, M Rancea, Prof A Engert MD), German Hodgkin Study Group
(H Haverkamp, Prof V Diehl MD, Prof A Engert, Prof P Borchmann MD), and Department I of Internal Medicine (Prof A Engert, Prof P Borchmann), University Hospital of Cologne, Cologne, Germany; Institute of Social and Preventive Medicine and CTU Bern, Department of Clinical Research, University of Bern, Bern, Switzerland (S Trelle MD)

Correspondence to:
Prof Peter Borchmann,
Department I of Internal Medicine, University Hospital of Cologne, Cologne 50924, Germany
peter.borchmann@uk-koeln.de

transplantation. Thus, initial treatment with ABVD is advocated and widely used as standard of care.¹

Only one trial⁴ has assessed these two different treatment strategies directly, and reported an overall survival difference of 5% in favour of initial treatment with the BEACOPP variant of four cycles of BEACOPP_{escalated} followed by four cycles of BEACOPP_{baseline} (similar to BEACOPP_{escalated}, but with substantially lower doses of most of the drugs). Unfortunately, this trial was not powered to test overall survival differences.⁴ A head-to-head comparison of ABVD with six cycles of BEACOPP_{escalated}, the current standard treatment for the early intensification approach, has never been done.⁵ Thus, the best initial treatment strategy in terms of overall survival in this population of patients is not known.

To investigate this important question, we used a network meta-analysis approach. In network meta-analyses, the information available from within-study comparisons of regimen A and regimen B is combined with indirect comparisons of A and B derived from studies that compare either of the two regimens with a common comparator C. Additionally, the method allows the quantification of the relative efficacy of regimens that have not been compared directly within a controlled randomised trial. This approach allows for a unified, coherent analysis of all relevant randomised controlled trials while fully respecting randomisation of the included trials, and overcomes the limitation of having few available direct comparisons. The network meta-analysis presented here aims to summarise the direct and indirect evidence for different first-line chemotherapy variants and intensities to provide the highest level of evidence for treatment

decisions in patients with advanced-stage Hodgkin's lymphoma.

Methods

Search strategies and selection criteria

We adapted search strategies from those suggested in the *Cochrane Handbook for Systematic Reviews of Interventions*.⁶ We searched the following databases without any language restrictions: the Cochrane Library, including the Cochrane Central Register of Controlled Trials (CENTRAL; issue 5, 2013), Medline (Ovid; January, 1980, to June, 2013), and conference proceedings of annual meetings of the American Society of Hematology, American Society of Clinical Oncology, European Hematology Association, and International Symposium on Hodgkin Lymphoma from 2000 to 2012, if not included in CENTRAL.

We included randomised controlled trials in adult patients with newly diagnosed Hodgkin's lymphoma at an advanced stage, as defined by the investigators, assessing any of the following prespecified interventions: BEACOPP_{baseline}, BEACOPP_{escalated}, BEACOPP variants (ie, four cycles of BEACOPP_{escalated} followed by two or four cycles of BEACOPP_{baseline}), ABVD, cyclophosphamide (mechlorethamine), vincristine, procarbazine, and prednisone (C[M]OPP), hybrid or alternating chemotherapy regimens with ABVD as the backbone (eg, COPP/ABVD, MOPP/ABVD), and doxorubicin, vinblastine, mechlorethamine, vincristine, bleomycin, etoposide, and prednisone combined with radiation therapy (Stanford V). Details of the regimens are described in the appendix. Trial reports needed to provide data on overall survival to be eligible for inclusion, irrespective of follow-up length.

Study selection and data extraction

Study selection, data extraction, and quality assessment were done independently by two investigators (NS and MR). Both investigators screened the titles and abstracts of study reports identified by the search strategies for eligibility. Disagreements were resolved by discussion with a third author (PB). Full-text versions of all eligible study reports were used for quality assessment (risk of bias) and for data collection. Data were extracted into a form that included a prespecified set of variables (study quality, study characteristics, patient characteristics, interventions, and outcome data). A domain-based approach was used for quality assessment. We assessed the following domains: generation of randomisation list, concealment of allocation, blinding, and whether intention-to-treat analysis was used. We estimated log hazard ratios (HR) and corresponding standard errors from overall survival curves of treatment effects with methods described by Tierney and colleagues.⁷ If outcome data were missing, we asked the contact author of the individual study for additional information. To allow for more flexible modelling, we

See Online for appendix

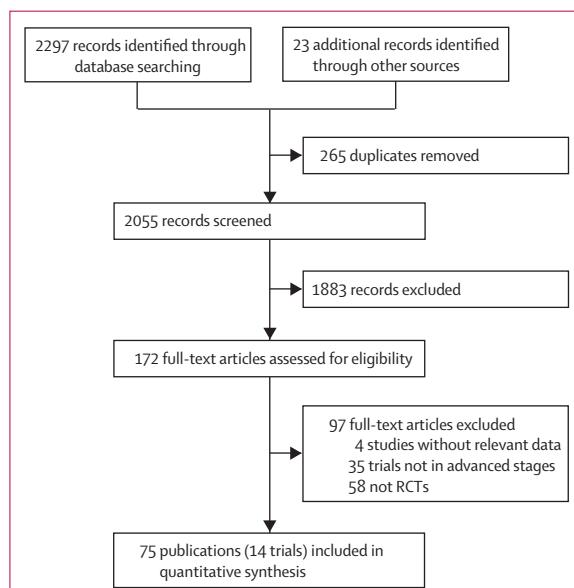


Figure 1: Study flow chart
RCT=randomised controlled trial.

reconstructed individual patient survival data from published overall survival curves by digitising published Kaplan-Meier curves.⁸

Outcome measures

We prespecified overall survival as the primary outcome. Secondary outcome measures were secondary neoplasia, secondary leukaemia, and freedom-from-treatment failure as defined by the included trials. If more than one related outcome was reported, we selected the one closest to time from randomisation to any of the following events (whichever occurred first): progression during treatment,

lack of complete remission at the end of primary therapy, relapse, or death from any cause. Patients without an event should have been censored at the longest follow-up available. To address serious short-term toxicities, we also analysed mortality during treatment, defined as any death during the planned chemotherapy administration up to 28 days after the end of treatment.

Statistical analyses

For the network meta-analysis of study-level overall survival and freedom-from-treatment failure data, we applied a random-effects approach as previously

Study start	Duration of recruitment (years)	Randomised patients (analysed)	Regimen (n [analysed])	Median follow-up (years)	Median age (years)	Eligible Ann Arbor stages	Patients at each Ann Arbor stage†		
							<III	III	IV
GHSG HD9 ²	1993	5	1282 (1196)	9·3	32	IIB, III, IV	155 (13%)	646 (54%)	395 (33%)
GITIL/IIL ⁴	2000	7	331	5·1	..	IIB, III or IV
ECOG2496 ^{3‡}	1999	7	854 (794)	5·3	33	I-IIA/B with bulky disease, III, IV	281 (35%)	301 (38%)	203 (26%)
GHSG HD9elderly ^{16§}	1993	5	68	6·7	70	IIB with bulky disease, III, IV	3 (4%)	44 (65%)	21 (31%)
IIL HD9601 ^{17§}	1996	4	354 (335)	7·2	34	IIB, III, IV	112 (33%)	144 (43%)	79 (24%)
GISL HD2000 ¹⁸	2000	7	307 (295)	3·3	32	IIB, III, IV, bulky disease	91 (31%)	133 (45%)	69 (23%)
BNLI (ISRCTN64141244) ¹⁹	1998	8	520	4·3	..	IA-IIA with bulky disease, IB, IIB, III, IV	253 (49%)	153 (29%)	114 (22%)
CALGB ²⁰	1982	4	400 (361)	6·0	34	III, IV	60 (17%)¶	132 (37%)	167 (46%)
NCIC ²¹	1984	5	332 (301)	5·8	30	IIIB, IV, first relapse after wide-field RT	75 (25%)¶	104 (35%)	121 (40%)
CALGB 8952 ²²	1990	5	856 (852)	6·0	..	III, IV, first relapse after definitive RT	381 (45%)
Intergroup ²³	1987	2	737 (691)	7·3	31	III, IV, first relapse after RT	33 (5%)	379 (55%)	279 (40%)
GHSG HD12 ^{24§}	1999	4	1670 (1574)	6·5	36	IIB with bulky disease, III, IV	255 (16%)	769 (49%)	549 (35%)
GHSG HD15 ^{5**}	2003	5	2182 (2126)	4·8	33	IIB with bulky disease, III, IV	331 (16%)	1065 (50%)	727 (34%)
EORTC 20012 ^{25§}	2003	7	550 (549)	3·9	..	III, IV	0	143 (26%)	406 (74%)

Data are number, median, or number (%), unless otherwise specified. Regimens are described in the appendix. The number followed by an asterisk preceding each BEACOPP regimen is the number of cycles of treatment in the regimen. GHSG=German Hodgkin Study Group. C(M)OPP=cyclophosphamide (mechlorethamine), vincristine, procarbazine, prednisone. ABVD=doxorubicin, bleomycin, vinblastine, dacarbazine. BEACOPP_{baseline}=bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone. BEACOPP_{escalated}=increased dose bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, prednisone. GITIL=Gruppo Italiano di Terapie Innovative nei Linfomi. IIL=Intergruppo Italiano Linfomi. ECOG=Eastern Cooperative Oncology Group. Stanford V=doxorubicin, vinblastine, mechlorethamine, vincristine, bleomycin, etoposide, and prednisone combined with radiation therapy. EBV=epidoxorubicin, bleomycin, vinblastine. CAD=lomustine, doxorubicin, vindesine. GISL=Gruppo Italiano per lo Studio dei Linfomi. BNLI=British National Lymphoma Initiative. CALGB=Cancer and Leukemia Group B. NCIC=National Cancer Institute Canada. RT=radiotherapy. EORTC=European Organisation for the Research and Treatment of Cancer. ..=data not available. ¶Numbers of patients were recalculated from percentages; minor discrepancies in numbers are due to rounding in the original reports. ‡Ann Arbor stage unclear for nine patients. §Additional information from investigators received. ||Assumption only. ||Ann Arbor stage unclear for 1 patient. **Ann Arbor stage unclear for three patients.

Table: Characteristics of included studies

described.⁹ The model uses group-level information and models (log) HRs to preserve randomisation and accounts for within-trial correlation of multigroup trials. The model was implemented within a Bayesian framework using WinBUGS (version 1.4.3). The median of the posterior distribution was used as a point estimate for the treatment effect. After ensuring that posterior distributions were roughly normally distributed, a 95% credible interval (CrI) was derived from the 2.5th and 97.5th percentiles. In the presence of minimally informative priors, CrIs can be interpreted like conventional CIs. We assessed model fit using three criteria based on the deviance and node-based residuals. Inconsistency, defined as the difference between the pooled direct and indirect evidence of a particular comparison, was assessed using inconsistency factors based on a modified back-calculation approach.¹⁰

We calculated HRs with ABVD as the baseline regimen to act as the effect measure. Additionally, we derived probabilities to be superior from the posterior distribution; for each of the iterations, regimens were ranked according to the estimated log HR. The probability of a regimen being superior was then defined as the proportion of times a regimen ranked first.

To assist interpretation of HRs we also provide an absolute measure of treatment effect for overall survival—the difference in 5 year survival.¹¹ We did standard random-effects meta-regression of 5 year survival in all ABVD trial groups to arrive at a reasonable

baseline risk, to which we then applied HRs derived from the network meta-analysis. To reconstruct individual patient survival data an iterative algorithm was then applied to solve the inverted Kaplan-Meier equations originally used to produce the digitised graphs.⁸ The algorithm uses the digitised curve data—ie, data on survival probabilities and time and the total number of patients and events per group—to find numerical solutions to the inverted Kaplan-Meier equations. The algorithm assumes constant censoring—ie, the censoring mechanism is non-informative. We used the algorithm as provided by Guyot and colleagues,⁸ implemented in R (version 2.13.0).

We used standard random-effects Poisson regression to analyse the reconstructed individual patient data.¹² The modelling approach introduces treatment-by-time interactions to allow for non-proportional hazards. We used fractional polynomials so that time could be flexibly modelled.¹³ We did two sensitivity analyses: in one we included year of trial (defined as the middle of the recruitment period) as a covariate in the model to control for possible trends over time, in another we pooled regimens including BEACOPP_{escalated} and introduced a covariate to account for their different intensities.

Incidence of second neoplasia was analysed with a Poisson random-effects model,¹⁴ and treatment-related mortality with a logistic random-effects model.¹⁵ These models were implemented within a Bayesian framework with WinBUGS (version 1.4.3).

The appendix provides additional details of the statistical methods used.

Role of the funding source

No specific funding was provided for this project. The corresponding author had full access to all data in the study and had final responsibility for the decision to submit for publication.

Results

We identified 2320 potentially relevant references from database searches (2055 records without duplicates; figure 1). Of these records, we excluded 1883 at the initial stage of screening because they did not fulfil our predefined inclusion criteria. The remaining 172 publications were retrieved as full-text or abstract publications for detailed assessment. We excluded 97 records, and included 75 publications (all of which were associated with 14 trials) in the analysis, although not all the publications reported overall survival. We used the latest publication of each trial for the network meta-analysis, as cited in the main publication.

The characteristics of the 14 included trials are summarised in the table. All trials were multicentre and done by cooperative study groups. The oldest study began in 1982, the newest in 2003, and the median recruitment period was 5 years (IQR 4.3–7.0). 10443 patients were randomly assigned to a treatment regimen

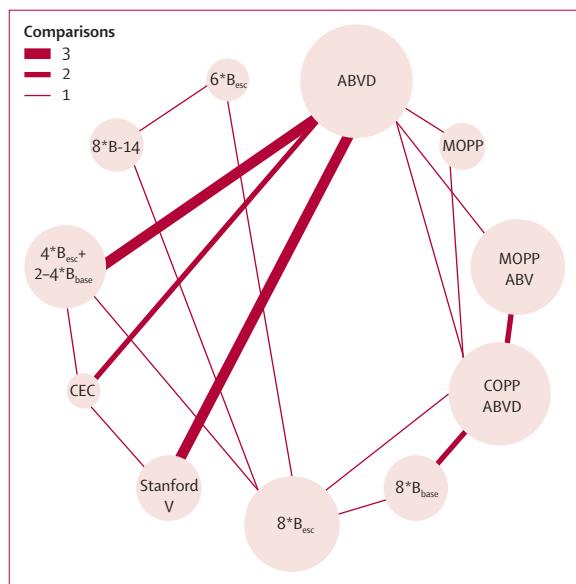


Figure 2: Network of analysed comparisons

Each circle corresponds to a regimen included in the analysis, with the area proportional to the statistical information (cumulative number of events). Each line represents direct comparisons between regimens, with the thickness corresponding to the number of available direct within-trial comparisons. Regimens are described in the appendix. The number followed by an asterisk preceding each BEACOPP regimen is the number of cycles of treatment in the regimen. $B_{\text{base}} = \text{BEACOPP}_{\text{baseline}}$. $B_{\text{esc}} = \text{BEACOPP}_{\text{escalated}}$. B-14 = BEACOPP-14. CEC = C(M)OPP/EBV/CAD.

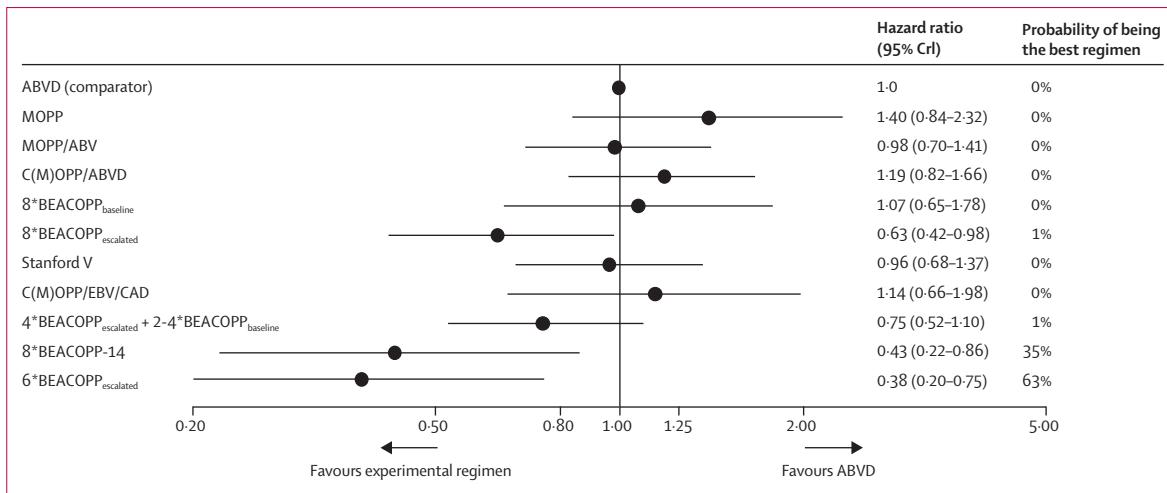


Figure 3: Overall survival for each regimen, compared with overall survival for ABVD

Regimens are described in the appendix. The number followed by an asterisk preceding each BEACOPP regimen is the number of cycles of treatment in the regimen. CrI=credible interval.

in the included studies, 9993 of whom were analysed, (median 535 patients randomised per study, IQR 338–856) with a median follow-up of 5·9 years (IQR 4·9–6·7). Five of the studies had three study groups, all others had two groups. Definitions of advanced stages varied between trials, but more than 60% of patients were in Ann Arbor stages III or IV in each study.

The methodological quality of trials included was high overall (appendix). Random sequence generation was adequate in all trials. Allocation concealment was not reported in nine trials, and was adequate in the remaining five trials. None of the trials was blinded. However, for the endpoint of overall survival, we think that bias is unlikely because death is an endpoint not susceptible to patient, physician, or outcome assessor bias.

Figure 2 shows all the comparisons analysed within the network. ABVD was the most frequently investigated regimen (11 comparisons). It was compared to Stanford V in three trials and to four cycles of BEACOPP_{escalated} plus two to four cycles of BEACOPP_{baseline} in three trials.

Overall survival was reported as a secondary outcome in all 14 trials. Results of the individual direct within-trial comparisons are shown in the appendix. Most comparisons showed differences, but the statistical power of the trials was generally too low for the differences in the secondary outcome of overall survival to reach conventional levels of significance. Pooling of HRs for overall survival for individual regimens compared with ABVD in the network meta-analysis showed a statistically significant advantage for six cycles of BEACOPP_{escalated} and eight cycles of BEACOPP-14 (HR 0·38, 95% CrI 0·20–0·75 and HR 0·43, 0·22–0·86, respectively). These two BEACOPP regimens had a combined 98% probability of being superior; individual probabilities were 63% for six cycles of BEACOPP_{escalated} and 35% for eight cycles of

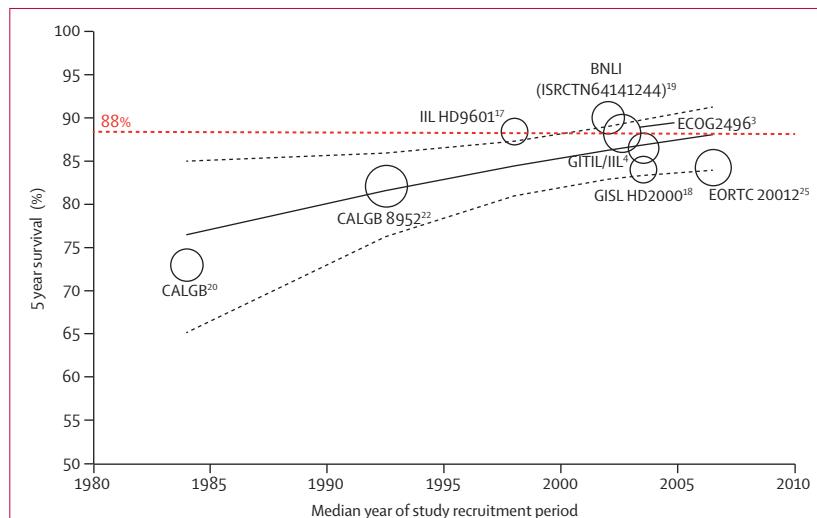


Figure 4: 5 year survival rates in patients assigned to ABVD, according to year of recruitment

The position of each circle represents the proportion of patients achieving 5 year survival in each study that assessed ABVD. The size of the circle corresponds to the (random-effects) weight in the meta-regression. The dashed horizontal straight line shows the proportion of patients given ABVD who survived for 5 years, which we used as a reference value. The solid curved line shows the pooled 5 year survival for ABVD; the dotted curved lines show the corresponding 95% CI. BNLI=British National Lymphoma Initiative. CALGB=Cancer and Leukemia Group B. ECOG=Eastern Cooperative Oncology Group. EORTC=European Organisation for the Research and Treatment of Cancer. GISM=Gruppo Italiano per lo Studio dei Linfomi. GITIL=Gruppo Italiano di Terapie Innovative nei Linfomi. IIL=Intergruppo Italiano Linfomi.

BEACOPP-14. Eight cycles of BEACOPP_{escalated} also significantly improved overall survival; however, the effect was less pronounced than with the two other BEACOPP-based regimens (figure 3). Overall survival in patients receiving other regimens such as Stanford V or MOPP/ABV did not differ in any clinically relevant way from patients receiving ABVD. In the network comparison of Stanford V with the two best performing regimens (six cycles of BEACOPP_{escalated} or eight cycles of

BEACOPP-14), BEACOPP significantly increased overall survival (BEACOPP_{escalated}, HR 0·39, 95% CrI 0·18–0·86; BEACOPP-14, HR 0·44, 0·20–0·97).

We used standard random-effects meta-regression of absolute overall survival rates to estimate a 5 year overall survival rate of 88% (95% CI 84–91) for ABVD, which we used as a reference value (figure 4). On the basis of this

survival rate, the reported results would translate into a 7% (95% CrI 3–10) benefit for 5 year overall survival for BEACOPP_{escalated}, and a 7% (2–9) benefit for BEACOPP-14 resulting in 95% 5 year overall survival for both regimens.

Results for both BEACOPP-based regimens were confirmed when we analysed reconstructed individual patient data (figure 5). We included 1189 deaths in

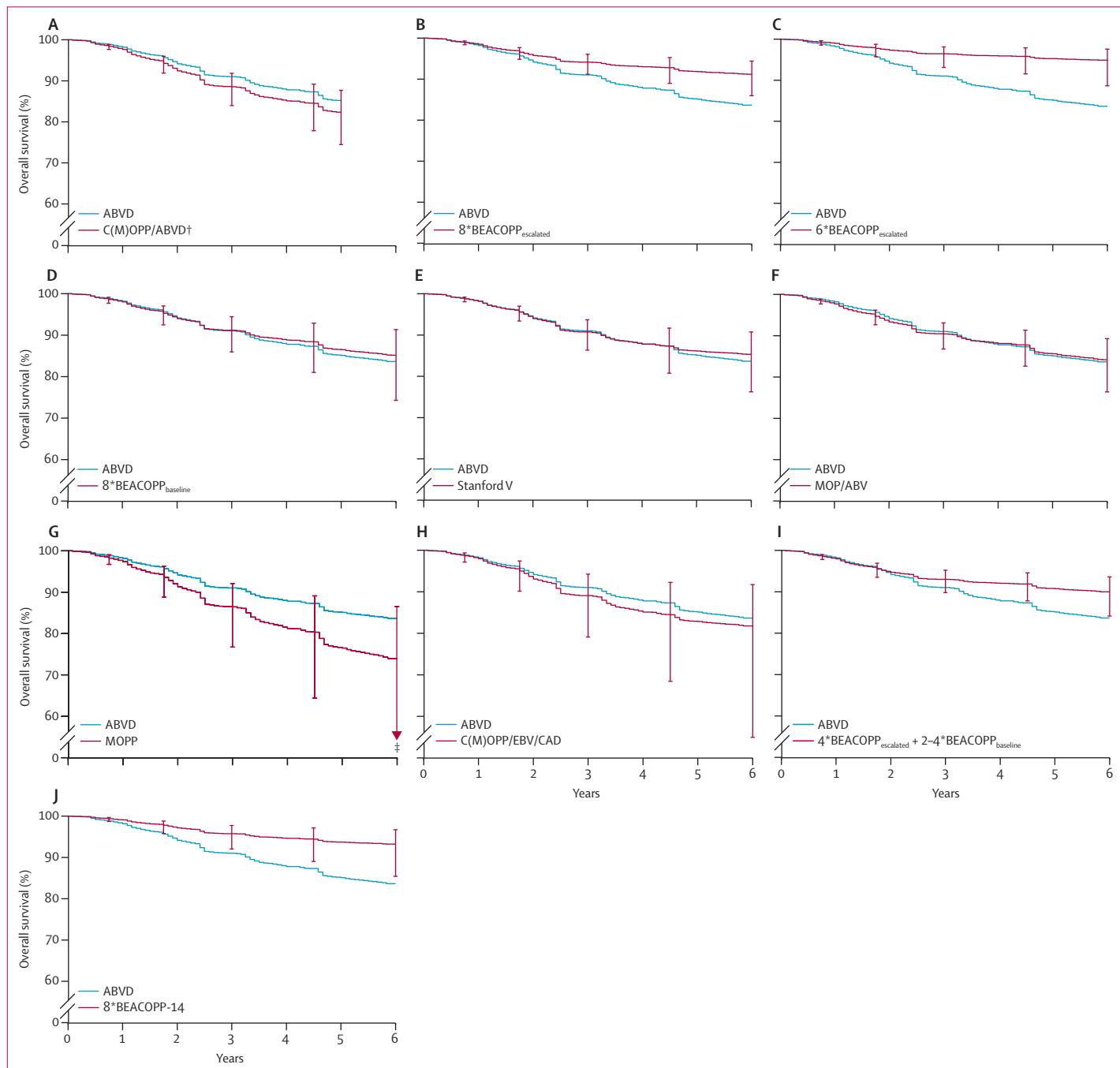


Figure 5: Overall survival based on reconstructed individual patient data

Regimens are described in the appendix. The number followed by an asterisk preceding each BEACOPP regimen is the number of cycles of treatment in the regimen. Curves show survival probabilities; error bars show 95% CIs. †The maximum follow-up time for C(M)OPP/ABVD was 5 years. ‡95% CI lower limit is 48%.

10 011 patients over 47 033 reconstructed patient-years of follow-up in the analysis. Figure 5 shows that the HRs appeared to vary over time, especially for four regimens, which showed non-proportional hazards: six cycles of BEACOPP_{escalated}, eight cycles of BEACOPP_{escalated}, four cycles of BEACOPP_{escalated} plus two to four cycles of BEACOPP_{baseline}, and eight cycles of BEACOPP-14. The corresponding global statistical test for all analysed regimens confirmed our visual impression ($p=0.0078$). Overall survival with six cycles of BEACOPP_{escalated} at 5 years was 10% better (95% CI 3–15) than with ABVD. Both sensitivity analyses confirmed these results (data not shown).

Between-trial heterogeneity was low in all analyses ($\tau^2=0.01$ [95% CrI 0.00–0.08] for the Bayesian analysis and $\tau^2=0.00$ for the analysis of reconstructed individual patient data). The network structure for the overall survival outcome allowed for six inconsistency factors (appendix). Data were sparse and we were able to estimate only four of them. None suggested relevant inconsistency, but CrIs were wide (appendix). Model fit was adequate according to all three criteria (appendix).

Definitions of freedom-from-treatment failure varied substantially between trials (appendix). Twelve trials reported data for this outcome. Results were consistent with the results for overall survival, with six cycles of BEACOPP_{escalated} having a 66% probability of being the superior regimen and eight cycles of BEACOPP-14 having a 15% probability (appendix). Twelve trials provided data for secondary malignancies and leukaemia. Median follow-up ranged from 3.3 to 9.3 years in these trials (median 6 years, IQR 4.3–6.5 for second malignancies and 4.3–6.8 for leukaemia). During this period, 327 secondary malignancies and 109 cases of leukaemia occurred. The numbers of events, according to the underlying regimen, are shown in the appendix. The numbers of cases of leukaemia were too low for a meaningful analysis. Also, numbers of second malignancies were low, resulting in wide CrIs. Regimens did not differ significantly with respect to second malignancies and leukaemia (appendix). Across all trials and regimens, 213 (2%) patients died during treatment. We noted no significant differences in mortality during treatment between regimens, but CrIs were wide (appendix).

Discussion

We derive three principal findings from our analysis: initial treatment with six cycles of BEACOPP_{escalated} provides the highest overall survival benefit when compared with ABVD; BEACOPP variants such as eight cycles of BEACOPP_{escalated} or BEACOPP-14 also significantly improve overall survival when compared with ABVD, but to a lesser extent; and secondary neoplasia was a rare event within the reported observation period and thus between-group differences in occurrences could not be assessed.

The most important finding of this study is the overall survival difference between BEACOPP variants,

especially six cycles of BEACOPP_{escalated}, compared with ABVD when given as an initial treatment strategy for adult patients with advanced stage Hodgkin's lymphoma. Although an overall survival benefit of 4–8% for BEACOPP variants had consistently been reported in all head-to-head comparisons with ABVD, the absence of statistical significance within the trials led to misleading conclusions;^{4,18,25} some readers took these findings to mean there was no evidence of superiority.^{1,26} However, we show that the difference in overall survival between ABVD and BEACOPP regimens is indeed meaningful. We also provide information that is not available from single trials. For example, the best initial treatment according to our analysis, six cycles of BEACOPP_{escalated}, has not been compared with ABVD directly;³ however, our network meta-analysis indirectly compared this regimen with initial ABVD treatment. Finally, we believe this meta-analysis is the largest and most comprehensive study of initial treatment for advanced-stage Hodgkin's lymphoma so far, and provides the highest level of evidence for both physicians and patients (panel).

We chose overall survival as the primary outcome for our analysis because, unlike response rates, progression-free survival, or event-free survival, overall survival is not biased by the outcome definition and assessment. It is, obviously, a very important endpoint for this group of young patients, and crucially depends on the observation

Panel: Research in context

Systematic review

We searched the Cochrane Library, including the Cochrane Central Register of Controlled Trials, and Medline up to June, 2013, with the search term "Hodgkin lymphoma". We did not apply any language restrictions. We identified 172 meta-analyses and reviews, of which only one was a systematic review of patients with Hodgkin's lymphoma in advanced stages.²⁷ The investigators analysed randomised controlled trials that assessed chemotherapy regimens with at least two cycles of BEACOPP_{escalated} (increased dose bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, prednisone) and compared them to chemotherapy with at least four cycles of ABVD (doxorubicin, bleomycin, vinblastine, dacarbazine) as first-line therapy for patients with early unfavourable or advanced-stage Hodgkin's lymphoma. Four trials of 2868 patients fitted the inclusion criteria and were meta-analysed. Although progression-free survival was statistically significantly longer for BEACOPP_{escalated} than for ABVD (HR 0.53, 95% CI 0.44–0.64), this effect did not translate into a statistically significant overall survival difference (HR 0.80, 0.59–1.09). Therefore, we undertook this network meta-analysis to assess the overall survival benefit of BEACOPP_{escalated}.

Interpretation

By contrast with the identified meta-analysis, we were able to include 14 trials in the network meta-analysis that assessed 11 different regimens in 10 042 patients for 47 033 patient-years of follow-up, and 1189 events with an average median follow-up of 5.9 years. To our knowledge, we show for the first time that the already known positive effect of BEACOPP regimens on progression-free survival translates into a significant overall survival advantage. Our results strongly support the use of six cycles of BEACOPP_{escalated} or eight cycles BEACOPP-14 as initial treatment for patients with Hodgkin's lymphoma in advanced stages.

time. At longer than 5 years, the average median observation time of these trials is adequate to cover nearly all deaths caused either by acute toxicity of first-line and second-line treatment or disease progression. With respect to fatal toxicity, acute treatment-related mortality (TRM) is a matter of concern with BEACOPP.²⁸ 1.7% of patients in the BEACOPP_{escalated} group in the GHSG HD9 trial had TRM, which was mainly caused by neutropenic infections (87.5%).^{29,30} Patients 40 years or older with poor performance status or those 50 years or older are at highest risk for TRM and should be given close clinical attention.²⁹ BEACOPP_{escalated} should not be used in patients older than 60 years due to excessive risk of TRM.³⁰ Because assessment of overall survival includes measuring the incidence of TRM, the survival difference that our meta-analysis revealed shows a true benefit of BEACOPP variants. TRM events in our dataset were too rare to allow for meaningful testing. However, when we compared results from the randomised trials, TRM for ABVD ranges from less than 1% to 3.3%, and from 0.8% to 2.2% for BEACOPP variants, respectively.^{4,25} This indirect comparison does not provide reliable evidence; however, it certainly does not show a meaningful difference between the two approaches with respect to TRM.

We also noted very low incidences of secondary malignancies across all regimens within the observation time. Regarding secondary acute leukaemia and myelodysplasia, six cycles of BEACOPP_{escalated} are associated with an incidence of secondary acute leukaemia and myelodysplasia of 0.3%.⁵ The incidences of these diseases after ABVD and BEACOPP variants from direct comparisons range from 0% to 1%, and 0% to 2%, respectively.^{4,18,25} Thus, these data provide no evidence for excessive rates of secondary acute leukaemia and myelodysplasia after chemotherapy with BEACOPP_{escalated}, which also is true for other secondary neoplasia in these trials.

These findings suggest that the risks of TRM and second malignancies associated with BEACOPP variants do not outweigh the survival benefits reported in our analysis. The hazards of BEACOPP regimens versus ABVD are non-proportional—ie, the regimens have increasing survival differences over time—which accords with long-term follow-up data from single trials.^{2,30} Patients who have had several relapses often die many years after not responding to initial ABVD treatment and might account for the increasing difference.^{4,18,25} An increasing difference in progression-free survival and overall survival has also been reported in the long-term follow-up of the GHSG HD9 study comparing COPP-ABVD and BEACOPP variants.³⁰ This result is also in accordance with the poor long-term survival reported for ABVD.³¹ The increasing overall survival difference over time underscores the relevance of the large overall survival difference at 5 years.

Regarding overall survival, treatment with eight cycles of BEACOPP-14 is not significantly different from six

cycles of BEACOPP_{escalated} in our analysis. However, these two regimens were directly compared in the GHSG HD15 trial,⁵ the results of which showed that BEACOPP-14 induces more severe toxicity associated with the nervous system and respiratory tract than does BEACOPP_{escalated}. Therefore, the risk and benefit balance of BEACOPP-14 is less favourable than with BEACOPP_{escalated}.

Our conclusion is strengthened by the methods used. First, the overall survival benefit for the most promising regimens was robust across a range of different analyses. Second, we identified two regimens that were clearly better than the others on the basis of the data, which makes interpretation straightforward from a clinical point of view. Statistical uncertainty was low, which strengthened the recommendations that can be derived.³² Third, although included trials were done in western Europe and North America only, they were undertaken by a range of study groups, which strengthens the external validity. Fourth, we used a comprehensive search strategy, thereby assuring that all available trial data were included.

The limitations of this analysis also need to be acknowledged. First, observation time was too short to consider secondary solid tumours or other late and fatal complications such as cardiovascular disease.³³ These events usually occur decades after initial treatment and are not systematically investigated in clinical trials; therefore, we could not analyse them. Data were also too sparse to assess other safety outcomes such as gonadal toxicity. Unfortunately, only the GHSG HD15 trial has reported data for this important safety variable so far,³⁴ and information about ABVD from the randomised controlled trials in this network meta-analysis is not available. Since much higher doses of alkylating agents are administered with BEACOPP_{escalated} than with ABVD, BEACOPP_{escalated} is very likely to induce more gonadal damage. However, for most patients, relapse-free survival is more important than preservation of fertility.³⁵ Nonetheless, individual patients might prioritise fertility preservation over definite tumour control. Therefore, the overall survival evidence derived from this meta-analysis should be used as part of a shared decision-making process.

Second, we included only a limited number of regimens (n=11) and trials (n=14). Moreover, the best performing regimen (ie, six cycles of BEACOPP_{escalated}) has been investigated in one trial only. Hodgkin's lymphoma is an orphan disease and collaborative study groups or intergroup trials are needed to enrol sufficient numbers of patients in a manageable period of time. Although between-trial heterogeneity was low and could be estimated precisely enough for the primary outcomes of interest, the low number of trials weakens the external validity of our analysis.

Third, the definition of advanced-stage Hodgkin's lymphoma varied across trials. The disease stage of patients included in the present analysis was

heterogeneous. Unfortunately, data did not allow us to explore effect modification by stage of disease. Patients with Ann Arbor stage III or IV were included in all trials and represent most patients in this analysis. Results apply to patients at those stages at least.

Fourth, results might apply only to countries that have a well developed medical infrastructure and an easily accessible health-care system. Acute haematological toxicity associated with BEACOPP variants is relevant, since about half of patients with this toxicity need close blood-count monitoring and dose reductions. Severe neutropenia occurs more frequently with BEACOPP variants than with ABVD, and growth factor support is mandatory.⁵ Acute toxicity such as neutropenic infection can affect overall survival if not appropriately managed. Insufficient supportive care might thus modify overall survival effects. Because sufficient supportive care for all patients was provided in all trials exploring BEACOPP, we could not investigate the potential effect of the health-care setting. Costly supportive measures and prompt access to specialised health-care facilities might be necessary to achieve the survival benefits of BEACOPP.

In summary, previous work has shown that initial treatment with BEACOPP variants has a well known positive effect on progression-free survival,²⁷ and this translates into a significant and meaningful overall survival advantage for patients with advanced-stage Hodgkin's lymphoma compared with initial treatment with ABVD. The survival advantage reaches 7–10% at 5 years when six cycles of BEACOPP_{escalated} are applied. Both the meaningful survival difference and the high-level evidence derived from this meta-analysis lead us to advocate initial treatment with six cycles of BEACOPP_{escalated} as standard of care in patients with advanced-stage Hodgkin's lymphoma in the appropriate health-care settings.

Contributors

NS, MR, AE, VD, and PB designed the study. NS and MR screened studies and extracted data. ST did the statistical analyses and prepared figures. NS, ST, MR, VD, AE, HH, and PB reviewed the results, interpreted data, and wrote the manuscript. All authors saw and approved the final version of the paper.

Conflicts of interest

We declare that we have no conflicts of interest.

Acknowledgments

We thank Monica Bellei, Luigi Marcheselli, and Massimo Federico from the Intergruppo Italiano Linfomi HD9601 trial for providing additional data. We are grateful to Kali Tal for proofreading and constructive feedback on the manuscript.

References

- 1 Connors JM. Hodgkin's lymphoma—the great teacher. *N Engl J Med* 2011; **365**: 264–65.
- 2 Diehl V, Franklin J, Pfreundschuh M, et al. Standard and increased-dose BEACOPP chemotherapy compared with COPP-ABVD for advanced Hodgkin's disease. *N Engl J Med* 2003; **348**: 2386–95.
- 3 Gordon LI, Hong F, Fisher RI, et al. Randomized phase III trial of ABVD versus Stanford V with or without radiation therapy in locally extensive and advanced-stage Hodgkin lymphoma: an intergroup study coordinated by the Eastern Cooperative Oncology Group (E2496). *J Clin Oncol* 2013; **31**: 684–91.
- 4 Viviani S, Zinzani PL, Rambaldi A, et al. ABVD versus BEACOPP for Hodgkin's lymphoma when high-dose salvage is planned. *N Engl J Med* 2011; **365**: 203–12.
- 5 Engert A, Haverkamp H, Kobe C, et al, on behalf of the German Hodgkin Study Group, the Swiss Group for Clinical Cancer Research, and the Arbeitsgemeinschaft Medikamentöse Tumorthterapie. Reduced-intensity chemotherapy and PET-guided radiotherapy in patients with advanced stage Hodgkin's lymphoma (HD15 trial): a randomised, open-label, phase 3 non-inferiority trial. *Lancet* 2012; **379**: 1791–99.
- 6 Higgins J, Green S, eds. Cochrane handbook for systematic reviews of interventions version 5.1.0 , updated March, 2011. Oxford, UK: The Cochrane Collaboration, 2011. <http://handbook.cochrane.org> (accessed July 29, 2013).
- 7 Tierney JF, Stewart LA, Ghersi D, Burdett S, Sydes MR. Practical methods for incorporating summary time-to-event data into meta-analysis. *Trials* 2007; **8**: 16.
- 8 Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol* 2012; **12**: 9.
- 9 Woods BS, Hawkins N, Scott DA. Network meta-analysis on the log-hazard scale, combining count and hazard ratio statistics accounting for multi-arm trials: a tutorial. *BMC Med Res Methodol* 2010; **10**: 54.
- 10 Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010; **29**: 932–44.
- 11 Altman DG, Andersen PK. Calculating the number needed to treat for trials where the outcome is time to an event. *BMJ* 1999; **319**: 1492–95.
- 12 Crowther MJ, Riley RD, Staessen JA, Wang J, Gueyffier F, Lambert PC. Individual patient data meta-analysis of survival data using Poisson regression models. *BMC Med Res Methodol* 2012; **12**: 34.
- 13 Jansen JP. Network meta-analysis of survival data with fractional polynomials. *BMC Med Res Methodol* 2011; **11**: 61.
- 14 Cooper NJ, Sutton AJ, Lu G, Khunti K. Mixed comparison of stroke prevention treatments in individuals with nonrheumatic atrial fibrillation. *Arch Intern Med* 2006; **166**: 1269–75.
- 15 Jonas DE, Wilkins TM, Bangdiwala S, et al. Findings of Bayesian mixed treatment comparison meta-analyses: comparison and exploration using real-world trial data and simulation. Rockville, MD: Agency for Healthcare Research and Quality, 2013.
- 16 Ballova V, Ruffer JU, Haverkamp H, et al. A prospectively randomized trial carried out by the German Hodgkin Study Group (GHSG) for elderly patients with advanced Hodgkin's disease comparing BEACOPP baseline and COPP-ABVD (study HD9elderly). *Ann Oncol* 2005; **16**: 124–31.
- 17 Chisesi T, Bellei M, Luminari S, et al. Long-term follow-up analysis of HD9601 trial comparing ABVD versus Stanford V versus MOPP/EBV/CAD in patients with newly diagnosed advanced-stage Hodgkin's lymphoma: a study from the Intergruppo Italiano Linfomi. *J Clin Oncol* 2011; **29**: 4227–33.
- 18 Federico M, Luminari S, Iannitti E, et al. ABVD compared with BEACOPP compared with CEC for the initial treatment of patients with advanced Hodgkin's lymphoma: results from the HD2000 Gruppo Italiano per lo Studio dei Linfomi Trial. *J Clin Oncol* 2009; **27**: 805–11.
- 19 Hoskin PJ, Lowry L, Horwich A, et al. Randomized comparison of the Stanford V regimen and ABVD in the treatment of advanced Hodgkin's Lymphoma: United Kingdom National Cancer Research Institute Lymphoma Group Study ISRCTN 64141244. *J Clin Oncol* 2009; **27**: 5390–96.
- 20 Canfell GP, Anderson JR, Propert KJ, et al. Chemotherapy of advanced Hodgkin's disease with MOPP, ABVD, or MOPP alternating with ABVD. *N Engl J Med* 1992; **327**: 1478–84.
- 21 Connors JM, Klimo P, Adams G, et al. Treatment of advanced Hodgkin's disease with chemotherapy—comparison of MOPP/ABV hybrid regimen with alternating courses of MOPP and ABVD: a report from the National Cancer Institute of Canada clinical trials group. *J Clin Oncol* 1997; **15**: 1638–45.
- 22 Duggan DB, Petroni GR, Johnson JL, et al. Randomized comparison of ABVD and MOPP/ABV hybrid for the treatment of advanced Hodgkin's disease: report of an intergroup trial. *J Clin Oncol* 2003; **21**: 607–14.

- 23 Glick JH, Young ML, Harrington D, et al. MOPP/ABV hybrid chemotherapy for advanced Hodgkin's disease significantly improves failure-free and overall survival: the 8-year results of the intergroup trial. *J Clin Oncol* 1998; **16**: 19–26.
- 24 Borchmann P, Haverkamp H, Diehl V, et al. Eight cycles of escalated-dose BEACOPP compared with four cycles of escalated-dose BEACOPP followed by four cycles of baseline-dose BEACOPP with or without radiotherapy in patients with advanced-stage Hodgkin's lymphoma: final analysis of the HD12 trial of the German Hodgkin Study Group. *J Clin Oncol* 2011; **29**: 4234–42.
- 25 Carde P, Mounier N. ABVD (8 cycles) versus BEACOPP (4 escalated cycles ≥ 4 baseline) in stage III-IV high-risk Hodgkin lymphoma (HL): first results of EORTC 20012 Intergroup randomized phase III clinical trial. *Proc Am Soc Clin Oncol* 2012; **30** (suppl): 8002 (abstr).
- 26 Borchmann P, Diehl V, Engert A. ABVD versus BEACOPP for Hodgkin's lymphoma. *N Engl J Med* 2011; **365**: 1545–46.
- 27 Bauer K, Skoetz N, Monsef I, Engert A, Brillant C. Comparison of chemotherapy including escalated BEACOPP versus chemotherapy including ABVD for patients with early unfavourable or advanced stage Hodgkin lymphoma. *Cochrane Database Syst Rev* 2011; **8**: CD007941.
- 28 Longo DL. Treatment of advanced Hodgkin lymphoma: the more things change, the more they stay the same. *J Clin Oncol* 2013; **31**: 660–62.
- 29 Wongso D, Fuchs M, Plutschow A, et al. Treatment-related mortality in patients with advanced-stage Hodgkin lymphoma: an analysis of the German Hodgkin Study Group. *J Clin Oncol* 2013; published online June 24. DOI:10.1200/JCO.2012.47.9774.
- 30 Engert A, Diehl V, Franklin J, et al. Escalated-dose BEACOPP in the treatment of patients with advanced-stage Hodgkin's lymphoma: 10 years of follow-up of the GHSG HD9 study. *J Clin Oncol* 2009; **27**: 4548–54.
- 31 Canello GP, Niedzwiecki D. Long-term follow-up of Hodgkin's disease trial. *N Engl J Med* 2002; **346**: 1417–18.
- 32 Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008; **336**: 924–26.
- 33 Aleman BM, van den Belt-Dusebout AW, Klokman WJ, Van't Veer MB, Bartelink H, van Leeuwen FE. Long-term cause-specific mortality of patients treated for Hodgkin's disease. *J Clin Oncol* 2003; **21**: 3431–39.
- 34 Behringer K, Mueller H, Goergen H, et al. Gonadal function and fertility in survivors after Hodgkin lymphoma treatment within the German Hodgkin Study Group HD13 to HD15 trials. *J Clin Oncol* 2013; **31**: 231–39.
- 35 Turner S, Maher EJ, Young T, Young J, Vaughan Hudson G. What are the information priorities for cancer patients involved in treatment decisions? An experienced surrogate study in Hodgkin's disease. *Br J Cancer* 1996; **73**: 222–27.