original article

Second malignancy risk associated with treatment of Hodgkin’s lymphoma: meta-analysis of the randomised trials


1German Hodgkin Study Group, University of Cologne, Germany; 2Rigshospitalet, Copenhagen, Denmark; 3University of La Sapienza, Rome, Italy; 4Oncology Hospital National Medical Center, Mexico City, Mexico; 5University of Florence, Florence, Italy; 6Medical Radiological Research Centre, Obninsk, Russia; 7Istituto Nazionale Tumori, Milan, Italy; 8Institut Bergonie, Bordeaux; 9Hospital Saint-Louis, Paris; 10Centre Regional Francoise Baclesse, Caen, France; 11Stanford University School of Medicine, Stanford, CA; 12St Jude's Children's Research Hospital, Memphis, TN, USA; 13Hamilton Regional Cancer Center, Ontario, Canada; 14Duke University School of Medicine, Durham, NC, USA; 15FUNDAEU, Buenos Aires, Argentina; 16Christie Hospital, Manchester, UK; 17University Medical Center Nijmegen, Nijmegen, The Netherlands; 18Memorial Sloan-Kettering Cancer Center, New York, NY, USA

Received 20 March 2006; revised 3 July 2006; accepted 11 July 2006

Background: Despite several investigations, second malignancy risks (SMR) following radiotherapy alone (RT), chemotherapy alone (CT) and combined chemoradiotherapy (CRT) for Hodgkin’s lymphoma (HL) remain controversial.

Patients and Methods: We sought individual patient data from randomised trials comparing RT versus CRT, CT versus CRT, RT versus CT or involved-field (IF) versus extended-field (EF) RT for untreated HL. Overall SMR (including effects of salvage treatment) were compared using Peto’s method.

Results: Data for between 53% and 69% of patients were obtained for the four comparisons. (i) RT versus CRT (15 trials, 3343 patients): SMR were lower with CRT than with RT as initial treatment (odds ratio (OR) = 0.78, 95% confidence interval (CI) = 0.62–0.98 and \(^P = 0.03\)). (ii) CT versus CRT (16 trials, 2861 patients): SMR were marginally higher with CRT than with CT as initial treatment (OR = 1.38, CI 1.00–1.89 and \(^P = 0.05\)). (iii) IF-RT versus EF-RT (19 trials, 3221 patients): no significant difference in SMR (\(^P = 0.28\)) although more breast cancers occurred with EF-RT (\(^P = 0.04\) and OR = 3.25).

Conclusions: Administration of CT in addition to RT as initial therapy for HL decreases overall SMR by reducing relapse and need for salvage therapy. Administration of RT additional to CT marginally increases overall SMR in advanced stages. Breast cancer risk (but not SMR in general) was substantially higher after EF-RT. Caution is needed in applying these findings to current therapies.

Key words: chemotherapy, Hodgkin’s lymphoma, meta-analysis, radiotherapy, second malignancies

introduction

Hodgkin’s lymphoma (HL) occurs predominantly in young adults and is one of the most curable malignancies. With current treatment approaches, most patients achieve a lasting complete remission, but secondary malignancies (SM) remain a serious late effect of treatment [1].

Evidence concerning the influence of treatment modality on SM risk is provided by numerous retrospective cohort studies based on large, often pooled datasets [2–22], as well as case–control studies [23–30]. Results, especially concerning solid tumours (ST), vary considerably. In most reports, the analysed treatment categories are based on both first-line and salvage modalities. Some, e.g. Boivin et al. [26], used time-dependent covariates to allow for the effects of later treatments. A few studies, such as Biti et al. [9], evaluated only initial treatment, but censored patients at relapse. Thus, such reports do not enable the ‘overall’ SM risks (i.e. due to both first-line and possible salvage therapy) associated with first-line treatment strategies to be compared directly. Only the analyses by Dores et al. [20] and Ng et al. [21] investigated the effect of initial treatment strategy on overall SM risk.

*Correspondence to: Dr J. Franklin, German Hodgkin Study Group, Herder Strasse 52-54, 50931 Koeln, Germany. Tel: +49-221-478-5894; Fax: +49-221-478-6311; E-mail: jeremy.franklin@uk-koeln.de

© 2006 European Society for Medical Oncology
The main objective of the present investigation was to compare overall SM risks in HL patients following first-line treatment with radiotherapy alone (RT), chemotherapy alone (CT) or combined chemoradiotherapy (CRT). Involved-field (IF) and extended-field (EF) RT were also compared. The investigation was carried out and electronically published as a Cochrane Collaboration systematic review [31].

methods

We aimed to collect individual patient data (IPD) from all randomised trials comparing RT, CT and/or CRT or comparing IF with EF or subtotal or total nodal RT (with or without CT) in newly diagnosed HL patients which enrolled at least 30 patients and which finished recruitment before or during 2000. Trials were sought in electronic literature databases, relevant conference proceedings from 1980 to 2001, lists of clinical trials, reference lists of all relevant retrieved publications and previous meta-analyses of HL.

IPD were requested from investigators from each eligible trial, including date of birth, sex, date of (first) HL diagnosis, stage of disease, presence or absence of B symptoms, treatment arm by randomisation, date of randomisation, remission status at the end of first-line treatment (with date), occurrence and date of relapse, occurrence, date and type of SM, occurrence and date of death and date of last follow-up information. All data were checked for completeness and consistency.

All patients who were randomly assigned in each trial were included (intention-to-treat), unless the first diagnosis of HL was reported as erroneous. A small number of patients for whom the relevant data fields were missing had to be excluded. As a preparatory step, each trial was analysed separately, comparing the treatment arms with respect to recruitment times, patient characteristics, complete remission rate, length of follow-up, overall survival (OS), event-free survival and time to SM.

This step investigates the comparability of the treatment arms and the consistency of the data with previous publications of the trial. Each trial was assessed for the following aspects of trial quality: randomisation method, adherence to the intention-to-treat principle, reliability of SM follow-up methods, completeness of follow-up and completeness of SM reporting. To assess completeness of follow-up, the median follow-up time was calculated using the Kaplan–Meier method. The distribution of last information dates was quantified; both high variability (large interquartile range) in relation to the median follow-up time and significant differences between treatment arms indicate less reliable follow-up.

Furthermore, observed SM incidence was compared with that expected in an age- and sex-matched cohort from the general population using data from USA and European cancer registries.

Randomised comparisons of RT versus CRT, CT versus CRT, RT versus CT and IF–RT versus EF–RT (with or without CT), respectively, were combined across the appropriate trials as follows. First, a measure of the ratio of SM incidence between the treatment arms of each trial was calculated separately, together with an estimate of the variance of this quantity, according to the method of Peto et al. [33, 34]. This method makes comparisons within each time period separately and thus takes account of the varying lengths of follow-up among the various trials. These measures were combined across trials relevant to the comparison being made to obtain a pooled Peto odds ratio (OR) for SM rate, i.e. the estimated SM treatment effect. For the OR, 95% confidence intervals (CIs) are also given. The three classes of SM [acute leukaemia (AL), non-Hodgkin’s lymphoma (NHL) and ST] were also analysed separately, censoring at occurrence of the other two classes, as were lung and breast ST.

Subgroup analyses were performed to investigate whether the SM treatment effect depended upon patient characteristics or treatment type.

The following subgroups were employed: stage (Ann Arbor) (early stage = I and II and advanced stage = III and IV), age (0–15 years, 16–39 years, 40–59 years and 60 years and older) and sex. Treatment-related subgroups were extent of RT (IF, more than IF) for CT versus CRT, type of CT (antracycline containing, others) for RT versus CRT and CT versus CRT. Results are displayed chronologically by recruitment period in order to reveal any time period effects.

Separate analyses were conducted with and without including SM that occurred after salvage therapy. For the latter, follow-up times were censored at HL progression/relapse and subsequent SM were ignored.

Sensitivity analyses were performed to check that the results were not crucially dependent on selection criteria or analysis methods. First, analyses were repeated with the exclusion of the less complete follow-up periods in each trial: for each trial, follow-up times were censored at the calendar date at which 75% of surviving patients in that trial were still being followed (‘cut-off’). Secondly, SM risk comparisons were analysed together by Cox proportional hazards regression [35] including relevant covariates and stratified by trial. Thirdly, the analysis was rerun excluding confounded trials. Fourthly, SM and ST analyses were repeated excluding non-melanoma skin cancers (NMSC) (as in many previous investigations of SM; some included trials did not record such cancers). Finally, the cumulative incidence method [36, 37], which allows for competing risks (deaths from other causes compete with second malignancies), was employed and the results qualitatively compared with those of the main analysis.

results

Seventy-six trials were identified as eligible, of which one was excluded after correspondence with the authors because it was not randomised. Although trials with <30 patients were to be excluded, we included a number of small Stanford University trials by amalgamating those with similar design and simultaneous or successive recruitment, and counted and analysed these as a single ‘trial.’ One dataset (St Jude Children’s Research Hospital) received as a single trial was split into two trials for the analysis, each with <30 patients, since two distinct study designs were applied to two groups defined by stage. One submitted dataset was excluded because no second malignancies were recorded.

In total, IPD from 37 trials could be analysed [38–71], including four which contributed to more than one treatment comparison. The earliest trial began accrual in 1962 and the latest ended accrual in 2000. Trial size ranged from 24 to 1136 patients. Data could not be obtained for a large number of otherwise eligible studies [72–103] as noted in Tables 1–4.

The analysed dataset included 9312 patients and 705 cases of SM: 92 AL, 103 NHL, 494 ST and 16 unspecified. The most common sites of ST were lung (97), skin (87: melanoma 19, non-melanoma 57 and unspecified 11), female breast (65), small intestine/colon/rectum (33) and stomach (20). Median OS times following occurrence of SM were as follows: AL 7 months, NHL 34 months and ST 36 months.

RT versus CRT

In total, 15 studies with 3343 patients (68% of those in all the eligible identified trials for this comparison), recruited from 1966 to 1998, were included (Table 1). IPD could not be obtained for 12 other trials. Most trials were for stage I–II patients only, while three trials also enrolled stage III patients.
Eight trials had an unconfounded design, i.e. the same radiotherapy was planned in each treatment arm. Seven trials were confounded, with (sub)total nodal irradiation [(S)TNI] in the RT arm and IF or mantle field in the CRT arm. The nine earlier trials used a regimen without anthracycline [a combination chemotherapy with mechlorethamine, vincristine, procarbazine and prednisone (MOPP) typically six times], whereas the six more recent trials included an anthracycline doxorubicin bleomycin vinblastine dacarbazine (ABVD) typically six times.

There was a higher overall risk of SM with RT alone compared with CRT ($P = 0.03$, OR = 0.78 for all stages together)—see Figure 1. This effect was most marked in stage III patients ($P = 0.02$, OR = 0.45) and did not reach significance in patients with stage I/II disease ($P = 0.13$, OR = 0.83).

The treatment effect of higher SM risk with RT alone was also seen when considering ST only ($P = 0.05$, OR = 0.78) and when considering NHL only ($P = 0.03$, OR = 0.46). AL risk was higher (though not significantly so) with CRT ($P = 0.21$, OR = 1.55 for early stages). No treatment effects were seen when considering either lung cancer or breast cancer alone.

If follow-up was censored at progression/relapse of HL, the SM treatment effect largely disappeared in both early and advanced stages ($P = 0.51$, OR = 1.11 for all stages together). Similar results were seen when considering ST only or NHL only; whereas, for AL only, there was a significantly higher risk with CRT ($P = 0.01$, OR = 3.40) when censoring at progression/relapse.

**CT versus CRT**

Sixteen studies with 2861 patients in total (53% of those in all the eligible identified trials), recruited from 1966 to 2000, were included (Table 2). Data could not be obtained for 12 other trials. There were 696 patients with early-stage (I–II) and 2165 with advanced-stage (III–IV) disease.

Ten trials were unconfounded, i.e. identical chemotherapy in each treatment arm, typically six cycles of MOPP or ABVD. Only the four most recent trials included an anthracycline (doxorubicin). Three trials were partially and three fully confounded, specifying more cycles in the CT arm than the CRT arm. The earliest five trials (1966–1974) used EF/RT or TNI, whereas the majority of subsequent trials used IF/RT.

SM risk was higher with CRT than with CT alone ($P = 0.05$, OR = 1.38 for all stages together)—see Figure 2. In early stages alone, no significant effect ($P = 0.73$, OR = 1.17) was seen. A modest treatment effect was seen in AL alone for all stages together ($P = 0.07$, OR = 1.82). There was no treatment effect for NHL or ST alone.

The effects are largely unchanged, but favour CT somewhat more strongly, if follow-up is censored at progression/relapse.

---

**Table 1.** Trials analysed for the comparison RT versus CRT

<table>
<thead>
<tr>
<th>Trial</th>
<th>Recruitment period/median follow-up (years)</th>
<th>Stage and other criteria</th>
<th>No. of patients in dataset</th>
<th>RT</th>
<th>CRT</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stanford H2-H6, K1, R1</td>
<td>1968–1979, 27</td>
<td>I–IV</td>
<td>269</td>
<td>EF/TNI</td>
<td>6MOPP + EF/TNI</td>
<td></td>
</tr>
<tr>
<td>St Jude IIIB</td>
<td>1972–1975, 27</td>
<td>IIA, IIA</td>
<td>24</td>
<td>IF</td>
<td>VCP + EF</td>
<td></td>
</tr>
<tr>
<td>Mexico 82HO31</td>
<td>1983–1988, 16</td>
<td>IIA</td>
<td>103</td>
<td>IF</td>
<td>6MOPP</td>
<td></td>
</tr>
<tr>
<td>Rome RT versus CRT</td>
<td>1983–1993, 11</td>
<td>IIA</td>
<td>333</td>
<td>IF</td>
<td>6MOPP + IF</td>
<td></td>
</tr>
<tr>
<td>Manchester VAPEC-B</td>
<td>1989–1997, 8</td>
<td>IA, IIA</td>
<td>124</td>
<td>IF</td>
<td>VAPE-C + IF</td>
<td></td>
</tr>
<tr>
<td>EORTC-GELA H8F</td>
<td>1993–1998, 6</td>
<td>I–II</td>
<td>543</td>
<td>IF</td>
<td>6MOPP + MVPP</td>
<td></td>
</tr>
<tr>
<td>GHSG HD7</td>
<td>1994–1998, 5</td>
<td>I–II</td>
<td>627</td>
<td>IF</td>
<td>2ABVD + EF</td>
<td></td>
</tr>
<tr>
<td>CALGB 6604</td>
<td>1966–1971, 27</td>
<td>III</td>
<td>40</td>
<td>TNI</td>
<td>Mechlorethamine + vinblastine + IF/TNI</td>
<td></td>
</tr>
<tr>
<td>CALGB 7451</td>
<td>1974–1981, 17</td>
<td>III</td>
<td>168</td>
<td>TNI</td>
<td>6BOPP + TNI</td>
<td></td>
</tr>
</tbody>
</table>

*Excluded trials (eligible, but individual patient data not obtained): BNLI TNI versus LOPP + TNI ($n = 85$), Chicago EF versus COPP + EF ($n = 49$), Eastern Cooperative Oncology Group 2475 EF versus C/MOPP + IF ($n = 54$), GEMH H9-69 RT versus MOPP + RT ($n = 198$), IHDCS (Pediatric Oncology Group) IF/EF versus MOPP + IF ($n = 220$), Lyon LM809a EF versus 6MOPP + IF ($n = 48$), Moscow EF versus CVPP + EF/IF ($n = 95$), NCI EF versus MOPP + EF ($n = 87$), Roswell Park IF/TNI versus CHVPP + IF/TNI ($n = 165$), SWOG 7811 EF versus 6MOPP + IF ($n = 235$), SWOG 9133 STNI versus (doxorubicin + vinblaste) + STNI ($n = 348$), Western CSG 135 TNI versus MOPP + TNI ($n = 40$).

ABVD, doxorubicin bleomycin vinblastine dacarbazine; BOPP, BCNU vincristine procarbazine prednisone; CALGB, Cancer and Leukaemia Group B; CRT, combined chemoradiotherapy; CVPP, cyclophosphamide vinblastine procarbazine prednisone; EF, extended field; EBVP, epirubicin bleomycin vinblastine prednisone; EORTC, European Organisation for Research and Treatment of Cancer; GELA, Groupe d’Etudes des Lymphomes de l’Adulte; GHSG, German Hodgkin Study Group; IF, involved field; MOPP, combination chemotherapy with mechlorethamine, vincristine, procarbazine and prednisone; MVPP, mechlorethamine vinblastine procarbazine prednisone; RT, radiotherapy alone; (S)TNI, (sub)total nodal irradiation; VAPEC-B, doxorubicin cyclophosphamide etoposide vincristine bleomycin prednisone; VBM, vinblastine bleomycin methotrexate; VCP, vinblastine cyclophosphamide procarbazine.

---

Eight trials had an unconfounded design, i.e. the same radiotherapy was planned in each treatment arm. Seven trials were confounded, with (sub)total nodal irradiation [(S)TNI] in the RT arm and IF or mantle field in the CRT arm. The nine earlier trials used a regimen without anthracycline [a combination chemotherapy with mechlorethamine, vincristine, procarbazine and prednisone (MOPP) typically six times], whereas the six more recent trials included an anthracycline doxorubicin bleomycin vinblastine dacarbazine (ABVD) typically six times.

There was a higher overall risk of SM with RT alone compared with CRT ($P = 0.03$, OR = 0.78 for all stages together)—see Figure 1. This effect was most marked in stage III patients ($P = 0.02$, OR = 0.45) and did not reach significance in patients with stage I/II disease ($P = 0.13$, OR = 0.83).

The treatment effect of higher SM risk with RT alone was also seen when considering ST only ($P = 0.05$, OR = 0.78) and when considering NHL only ($P = 0.03$, OR = 0.46). AL risk was higher (though not significantly so) with CRT ($P = 0.21$, OR = 1.55 for early stages). No treatment effects were seen when considering either lung cancer or breast cancer alone.

If follow-up was censored at progression/relapse of HL, the SM treatment effect largely disappeared in both early and advanced stages ($P = 0.51$, OR = 1.11 for all stages together). Similar results were seen when considering ST only or NHL only; whereas, for AL only, there was a significantly higher risk with CRT ($P = 0.01$, OR = 3.40) when censoring at progression/relapse.
Table 2. Trials analysed for the comparison CT versus CRT\(^a\)

<table>
<thead>
<tr>
<th>Trial</th>
<th>Recruitment period/median follow-up (years)</th>
<th>Stage and other criteria</th>
<th>No. of patients in dataset</th>
<th>CT</th>
<th>CRT</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mexico 82H031</td>
<td>1983–1988, 16</td>
<td>I–IV</td>
<td>201</td>
<td>6ABVD</td>
<td>6ABVD + mantle</td>
<td>[57]</td>
</tr>
<tr>
<td>CALGB 6604</td>
<td>1966–1971, 27</td>
<td>III</td>
<td>67</td>
<td>Methotrexate + vinblastine + TNI</td>
<td>Methotrexate + vinblastine + TNI</td>
<td>[38]</td>
</tr>
<tr>
<td>Stanford K7, S8</td>
<td>1969–1980, 22</td>
<td>IV</td>
<td>58</td>
<td>6MOPP</td>
<td>6MOPP + TNI</td>
<td></td>
</tr>
<tr>
<td>St Jude IBC</td>
<td>1972–1975, 27</td>
<td>III, IIIB, IV</td>
<td>24</td>
<td>VCP</td>
<td>VCP + EF</td>
<td>[67]</td>
</tr>
<tr>
<td>Stanford C7–10, C12–15</td>
<td>1980–1987, 14</td>
<td>III–IV</td>
<td>74</td>
<td>6(MOPP or PAVe) + ABVD</td>
<td>6PAVe + TNI</td>
<td></td>
</tr>
<tr>
<td>GHSG HD3(^b)</td>
<td>1982–1988, 14</td>
<td>IIIB–IV</td>
<td>100</td>
<td>4(COPP + ABVD)</td>
<td>3(COPP + ABVD) + IF</td>
<td>[48]</td>
</tr>
<tr>
<td>EORTC 20884(^b)</td>
<td>1989–2000, 6</td>
<td>III–IV</td>
<td>333</td>
<td>6/8MOPP/ABV</td>
<td>6/8MOPP/ABV + IF</td>
<td>[40]</td>
</tr>
<tr>
<td>GELA H89</td>
<td>1989–1996, 6</td>
<td>IIIB–IV</td>
<td>419</td>
<td>8(MOPP/ABV or ABVPP)</td>
<td>6(MOPP/ABV or ABVPP) + (S)TNI</td>
<td>[47]</td>
</tr>
</tbody>
</table>

\(^a\)Excluded trials (eligible, but individual patient data not obtained): CCG 521 6MOPP/ABVD versus 6ABVD + IF (n = 111), CCG 5942 COPP/ABV versus COPP/ABV + IF (n = 501), ECOG EST1476 6Bleo-MOPP + 3ABV versus 6Bleo-MOPP + IF (n = 232), ECOG EST1481 (BCVPP or MOPP/ABV) versus BCVPP + IF (n = 319), Lyon LMS80b 12MOPP/CVPP versus 6MOPP + EF (n = 58), Mexico Ho8326 6BVD versus 6BVD + IF (n = 118), NCI 6MOPP versus 6MOPP + EF (n = 58), POG 8625 3MOPP/ABV versus 2MOPP/ABV + IF (n = 247), POG 8725 4MOPP/ABV versus 4MOPP/ABV + TNI (n = 181), SWOG 7518 10MOPP-Bleo versus 3MOPP-Bleo + TNI (n = 118), SWOG 7747/775 MOPP(MOPP/Bleo) versus MOPP(Bleo) + IF (n = 254), SWOG 8805 6MOPP + IF versus 6MOPP-BAP + IF (n = 278).

\(^b\)Randomisation only if complete remission after chemotherapy.

ABV, doxorubicin bleomycin vinblastine; ABVD, doxorubicin bleomycin vinblastine dacarbazine; ABVPP, doxorubicin bleomycin vinblastine procarbazine prednisone; BOPP, BCNU vincristine procarbazine prednisone; CALGB, Cancer and Leukaemia Group B; COPP, cyclophosphamide vincristine procarbazine prednisone; CT, combined chemotherapy alone; CVPP, cyclophosphamide vinblastine procarbazine prednisone; ECOG, Eastern Cooperative Oncology Group; EF, extended field; EORTC, European Organisation for Research and Treatment of Cancer; GATLA, Grupo Argentino de Tratamiento de la Leucemia Aguda; GELA, Groupe d’Etudes des Lymphomes de l’Adulte; GHSG, German Hodgkin Study Group; IF, involved field; MOPP, combination chemotherapy with methotrexate, vincristine, procarbazine and prednisone; MSKCC, Memorial Sloan-Kettering Cancer Center; MOPP, methotrexate vinblastine procarbazine prednisone; PAVe, procarbazine melphalan vinblastine; POG, Pediatric Oncology Group; (S)TNI, (sub)total nodal irradiation; VCP, vincristine cyclophosphamide procarbazine.

Table 3. Trials analysed for the comparison RT versus CT\(^a\)

<table>
<thead>
<tr>
<th>Trial</th>
<th>Recruitment period/median follow-up (years)</th>
<th>Stage and other criteria</th>
<th>No. of patients in dataset</th>
<th>RT</th>
<th>CT</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rome, Florence</td>
<td>1979–1982, 16</td>
<td>IA, IIA</td>
<td>94</td>
<td>EF</td>
<td>6MOPP</td>
<td>[64]</td>
</tr>
<tr>
<td>Mexico 82H031</td>
<td>1983–1988, 16</td>
<td>I–II</td>
<td>205</td>
<td>Mantle</td>
<td>6ABVD</td>
<td>[57]</td>
</tr>
</tbody>
</table>

\(^a\)Excluded trials (eligible, but individual patient data not obtained): BNLI TNI versus 6MOPP (n = 165), NCI EF versus 6MOPP (n = 86), St Bartholomew’s, London, TNI versus 6MOPP (n = 60).

ABV, doxorubicin bleomycin vinblastine dacarbazine; BOPP, BCNU vincristine procarbazine prednisone; CALGB, Cancer and Leukaemia Group B; CT, chemotherapy alone; EF, extended field; MOPP, combination chemotherapy with methotrexate, vincristine, procarbazine and prednisone; RT, radiotherapy alone; TNI, total nodal irradiation.

(P = 0.01, OR = 1.60 for all stages together, no difference for early stages alone). In this case, there are somewhat more ST (P = 0.07, OR = 1.60) and significantly more leukemias (P = 0.01, OR = 2.75) with CRT than with CT (but no difference in NHL).

RT versus CT

Three studies with 415 patients in total (57% of those in eligible identified trials), recruited from 1974 to 1988, were included (Table 3). IPD could not be obtained for three other trials. Two trials recruited stages I–II and one recruited stage III. Mantle
field, EF and TNI radiotherapy were compared with six cycles of MOPP, BCNU, vincristine, procarbazine and prednisone (BOPP) and ABVD, respectively. There were non-significantly more SM with CT than RT ($P = 0.13$, OR = 2.12 for all stages); this difference was more pronounced for the early stages ($P = 0.05$, OR = 3.37)—see Figure 3. There were insufficient events to analyse each type of SM. In the analysis censoring follow-up at progression and relapse, no significant treatment effect was seen ($P = 0.30$, OR = 1.99).

### IF-RT versus EF-RT

Ten studies with 3221 patients in total (69% of those in eligible identified trials), recruited from 1962 to 1999, were included (Table 4). IPD could not be obtained for three other trials. Eight trials were mainly for early- and two mainly for advanced-stage disease.

Two studies (259 patients in total) planned no chemotherapy, six planned identical chemotherapy in each arm and two specified, in certain cases, more cycles in the IF-RT arm than in the EF-RT arm (i.e. partially confounded). Chemotherapy was MOPP like, MOPP/ABVD like or ABVD except in one study, usually with three to six cycles.

Neither were there significant differences in the rate of SM between EF-RT and IF-RT ($P = 0.28$, OR = 1.17 for all stages)—see Figure 4—nor were there significant differences in AL, NHL or ST rates. For breast cancers alone, there was a significantly greater risk with EF-RT ($P = 0.04$, OR = 3.25) but no significant difference for lung cancers ($P = 0.22$).

When follow-up was censored at progression/relapse, the tendency to more SM with EF-RT was stronger but still not significant ($P = 0.09$, OR = 1.54).

### Discussion

This systematic review is one of the largest investigations of SM yet performed. To the authors’ knowledge, it is the only large study of SM risk employing randomised comparisons, except for the meta-analyses by Loeffler et al. [104] and (concerning deaths from SM) Specht et al. [77].

### Conclusions

RT as a first-line treatment strategy for stage I–III patients leads to a higher overall rate of all SM, ST and NHL, respectively, than a combined modality strategy. This appears to be due to the significantly greater rate of progression/relapse and therefore intensive salvage therapy after RT alone, since the effect vanishes in an analysis censored at progression/relapse.

Administration of RT in addition to CT in first-line treatment of advanced-stage patients appears to increase the overall rate of SM and ST, respectively (borderline significance). There was no evidence of such an effect in early stages, but data were limited.

Perhaps surprisingly, our analysis did not convincingly demonstrate a higher rate of SM due to EF-RT rather than...
IF-RT, with the exception of breast cancers (OR 3.25 observed). An analysis censored at progression/relapse indicated a higher SM rate with EF (borderline significance), which may have been offset in the overall analysis by the significantly higher rate of progression/relapse [31], and therefore salvage therapy, after first-line IF.

limitations of included studies

The large majority of studies used adequate methods of randomisation resulting in treatment cohorts well balanced with respect to patient characteristics. In many studies, certain randomised patients were excluded from the trialists’ own analyses, against our intention-to-treat principle. We could not always reinclude such patients in our analysis: either they were omitted from the dataset or outcome data were missing.

Twelve of 37 trials had a median follow-up of at least 20 years (Tables 1–4). Several of the largest trials, however, had a median follow-up of between 4 and 6 years, too short for detection of ST in particular. Loss to follow-up resulted in a dispersion of dates of last information in most trials (the largest interquartile range was nearly 10 years). This is a potential source of bias in estimating event rates. We, however, found no evidence of differences in follow-up pattern between the treatment arms in any trial, so a bias in treatment effect is not suggested.

The greatest source of uncertainty, in our opinion, is the reliability of reporting of SM. Particularly, the earlier trials were not designed to provide information on SM risk; underreporting is likely. Most trialists assessed their SM information as ‘probably incomplete.’ Few had cross-checked with death or cancer registries. Comparison of observed SM rates with those expected on the basis of cancer registry data implied serious underreporting in a few trials. On the other hand, SM rates may be overreported in the sense that patients without an event are more likely to be lost to follow-up. A bias in the estimation of treatment effect would result only if such
reporting biases differed between treatment arms. Furthermore, some trialists did not record NMSC.

**meta-analysis methods**

In order to obtain adequate numbers of SM events for reliable comparisons, we included trials spanning four decades with varying diagnostic and therapeutic methods. Treatment differences may vary widely according to chemotherapy regimen and radiotherapy technique. Irradiation techniques have advanced considerably since the 1960s and this may have reduced SM risk. The use of new chemotherapy drugs, avoiding alkylating agents and favouring anthracycline-containing combinations have been shown to reduce the risk of second AL [105]. Subgroup analyses generally lack power to elucidate these variations reliably. Development of more effective salvage treatment strategies will also have altered survival rates and SM risk after first-line treatment failure.

It was decided in advance to count all SM as events, including NMSC which were not counted in some previous investigations of SM after HL. Some contributing trialists did not submit data on NMSC. We, however, performed sensitivity analyses not counting NMSC as events, which led to ORs closely consistent with the main analysis.

Although each type of SM (AL, NHL and ST), as well as lung and breast cancers, was analysed separately, total SM was the main outcome measure, since the larger number of events thus available permitted a more powerful analysis. One must, however, be aware that the counted events differ in severity.

In the present analysis, early (stages I–II) and advanced (stages III–IV) disease were analysed both together and separately. The comparisons RT versus CRT and IF-RT versus EF-RT were based largely on early-stage patients, while the majority of patients in the CT versus CRT comparison had advanced-stage disease. For reasons of statistical power, we defined the combined analysis as the main one. Heterogeneity between stages in the comparison of overall second malignancy risk, however, is likely because both (i) the treatment intensity of each modality and (ii) the progression/relapse rate and thus the frequency of salvage treatment are stage dependent. For this
reason, firm conclusions for a particular comparison have been drawn only for those stages where the data for this comparison are adequate.

We found no evidence of differences in treatment comparison of SM risk according to age or sex, but again, subgroup analyses are underpowered and might miss real differences. No data were available on many potential SM risk factors, such as smoking habits.

**previous evidence**

The meta-analysis by Loeffler et al. [104] reported significantly more deaths due to second AL with CRT compared with CT alone in predominantly advanced-stage patients (hazard ratio 2.48, \( P = 0.038 \)). In the present analysis, concordant results were obtained (OR = 1.82, \( P = 0.07 \); censoring at progression/relapse: OR = 2.57, \( P = 0.02 \)). The meta-analysis by Specht et al. [77] did not detect any difference in SM-related death rates between IF-RT versus EF-RT or between RT versus CRT.

Several previous studies [5, 6, 9, 13, 21, 23] reported higher SM risk with CRT than with RT alone, none found higher risk with RT alone as in the present review. This presumably reflects the fact that the (many) patients relapsing after RT and receiving salvage CT were classified in the CRT group in most previous studies. The report by Ng et al. [21], which classified patients according to first-line treatment only, also obtained a significantly higher relative risk of SM with initial CRT than with initial RT alone; however, this analysis included all stages IA–IVB, so patient groups with greatly differing treatment intensities were compared.

A few studies showed higher risk with CRT than with CT alone, but most found no difference. The present analysis demonstrated a higher SM rate with CRT.

Various studies have demonstrated a higher SM or AL risk with CT than with RT. The present analysis, based on an inadequate number of patients, demonstrated a non-significant trend in this direction. Two studies [21, 28] have reported a higher risk with EF-RT than with IF-RT. In the present analysis, the trend in this direction was not significant.

With two exceptions [26, 20], no significant differences between treatment modalities in the risk of ST had been previously observed, nor was a significant difference in NHL risk.
obtained [van Leeuwen et al. [11] reported a trend (P = 0.06) to more NHL with CRT than with either modality alone].

implications for research

Assessment and comparison of SM risk requires reliable long-term follow-up data from large numbers of patients, ideally those enrolled in randomised trials making the relevant treatment comparisons. Although the present analysis was able to detect significant differences in SM rates, the P values obtained were not small enough to confirm these differences beyond reasonable doubt. CIs for relative SM risks are wide. Routine follow-up documentation should include questions concerning the occurrence, type and site of second malignancies. Update campaigns should aim to fill in missing information 20–30 years after the recruitment period. Where possible, trialists should collaborate with death and cancer registries in order to record mortality and SM as completely as possible. This meta-analysis must be updated with longer follow-up from the included studies, as well as further eligible trials, if possible restricted to currently relevant treatment regimens.

contributors

J. Franklin was responsible for conception and planning, M. Paus and L. Specht were responsible for search for trials and contact with trialists, A. Pluetschow was responsible for data checking, A. Pluetschow and J. Franklin were responsible for data analysis and J. Franklin and A. Pluetschow constituted the writing committee.

acknowledgements

We are grateful to the following people for advice on methods and content: K. Wheatley, S. Richards, B. Djulbegovic, R. Meyer and M. Loeffler. The Deutsche Forschungsgemeinschaft
references

28. Brusamolino E, Anselmo AP, Knysh G et al. The risk of acute leukemia in patients treated for Hodgkin’s disease is significantly higher after combined modality programs than after chemotherapy alone and is correlated with the extent of radiotherapy and type and duration of chemotherapy: a case-control study. Haematologica 1998; 83: 812–823.
Annals of Oncology

original article


