

Autologous Stem Cell Transplantation in Follicular Lymphoma: a Systematic Review and Meta-analysis

Murtadha Al Khabori, John R. de Almeida, Gordon H. Guyatt, John Kuruvilla, Michael Crump

Manuscript received August 25, 2010; revised September 13, 2011; accepted September 14, 2011.

Correspondence to: Michael Crump, MD, FRCPC, Department of Medical Oncology and Hematology, Princess Margaret Hospital, 610 University Ave, 5-110, Toronto, ON M5G 2M9, Canada (e-mail: michael.crump@uhn.on.ca)

Background The impact of high-dose therapy and autologous stem cell transplantation (ASCT) vs conventional-dose chemotherapy in the initial management of adults with advanced follicular lymphoma (FL) on overall survival remains uncertain. We performed a systematic review of the randomized clinical trials addressing this question.

Methods We searched MEDLINE, EMBASE, CENTRAL, American Society of Hematology, American Society of Clinical Oncology, BIOSIS, PAPERSFIRST, PROCEEDINGS, clinical trials registries, and bibliographies of relevant studies for randomized clinical trials comparing myeloablative chemotherapy with ASCT to any chemotherapy in adults with untreated advanced FL. We performed a meta-analysis using random effects models to estimate overall survival, event-free survival, and risks of adverse outcomes. Statistical heterogeneity was calculated by using the I^2 statistic.

Results Seven trials proved eligible, four of which provided data from 941 patients that could be included in a meta-analysis and three of which remain unpublished. In two of the trials, patients in both arms received rituximab during the induction treatment. Moderate quality evidence from the three trials that reported overall survival ($n = 701$ patients) suggests that ASCT did not result in improved overall survival (hazard ratio of death = 0.99, 95% confidence interval [CI] = 0.73 to 1.33). Low-quality evidence from the four trials of 941 patients suggests improvement in event-free survival in favor of ASCT (hazard ratio of death = 0.54, 95% CI = 0.36 to 0.82) with substantial heterogeneity ($I^2 = 80\%$). Adverse outcomes of treatment-related mortality, myelodysplastic syndrome, acute myeloid leukemia, and solid tumors were not different between the two arms (relative risk [RR] of treatment-related mortality = 1.04, 95% CI = 0.29 to 3.70; RR of myelodysplastic syndrome/acute myeloid leukemia = 2.19, 95% CI = 0.45 to 10.55; $P = 48\%$; and RR of solid tumors = 1.30, 95% CI = 0.33 to 5.08). The absolute risk of death from treatment was 14 per 1000 patients for those who received chemotherapy and 15 per 1000 for those who received ASCT (range = 4–52).

Conclusions Available evidence suggests that high-dose therapy and ASCT as part of FL initial treatment does not improve overall survival. Future trials of ASCT in the context of current chemoimmunotherapy approaches to FL are needed.

J Natl Cancer Inst 2012;104:18–28

Follicular lymphoma (FL) is the most common indolent non-Hodgkin lymphoma (1,2) and generally has a long natural history with multiple remissions and relapses (3). Randomized control trials (RCTs) of multiple chemotherapy regimens have resulted in varying event-free survival (EFS) but have not indicated a survival advantage attributed to one regimen compared with another (4–6). The addition of interferon alpha to chemotherapy improves overall survival (OS) but is associated with substantial toxicity and is not generally accepted as standard therapy (7). Several RCTs of patients with FL have demonstrated the benefit of adding rituximab to primary therapy (8) with regard to OS and to maintenance therapy (9) with regard to progression-free survival.

Several RCTs have explored the possibility that intensification of chemotherapy beyond hematological limiting toxicity with autologous stem cell rescue (10–13) may improve EFS. With the recent RCTs of chemoimmunotherapy showing an OS benefit (8,14), the impact of any toxic therapy on OS should be questioned. Although narrative reviews are available (3,15–17), to our knowledge, no systematic review and meta-analysis has addressed this issue. Herein, we report the results of a systematic review and meta-analysis of the available RCTs examining the impact of high-dose chemotherapy with autologous stem cell transplantation (ASCT) vs any chemotherapy in the primary management of adults with advanced FL on OS, EFS, treatment-related mortality (TRM),

incidence of secondary myelodysplasia/acute myeloid leukemia (MDS/AML) and secondary solid tumors. Our main hypothesis is that ASCT improves OS, and previously reported trials may not have enough statistical power to detect differences in survival.

Methods

Search Strategy

We searched records from OVID MEDLINE (<http://www.ovid.com/site/catalog/DataBase/901.jsp>, for the period 1950–January 2009), OVID EMBASE (<http://www.ovid.com/site/catalog/DataBase/903.jsp>, for the periods 1947–1979 and 1980–2009 week 6), the Cochrane Central Register of Controlled Trials (CENTRAL, http://onlinelibrary.wiley.com/o/cochrane/cochrane_clcentral_articles_fs.html), and conference proceedings from the American Society of Hematology (<http://bloodjournal.hematologylibrary.org/search/>) and the American Society of Clinical Oncology (<http://www.asco.org/ascov2/Meetings/Abstracts>) for the time period 2000–2008. Our search included BIOSIS (<http://webofknowledge.com>, ISI Web of Knowledge-BIOSIS, for the time period 1980–February 12, 2009), PapersFirst (<http://www.oclc.org/ca/en/default.htm>, for the time period 2000–2008, through First Search), Proquest and Proceedings (<http://www.proquest.com/en-US/catalogs/databases/detail/cpi-set-c.shtml>, for the time period 2000–2008, through First Search; for abstracts and dissertations), and clinical trial registries; the National Cancer Institute (<http://clinicaltrials.gov>) and the World Health Organization International Clinical Trials Registry (www.who.int/trialsearch) for ongoing trials and searched the bibliographies of relevant publications. We provide the search terms in the Supplemental Materials (available online).

Study Selection

We included all RCTs of previously untreated FL adult patients that compared high-dose myeloablative chemotherapy with autologous peripheral blood or bone marrow stem cell transplantation, whether published or unpublished, in any language, which enrolled patients of any histological grade (World Health Organization grades I, II, and III) and clinical risk factors. Acceptable treatment regimens included any form of chemotherapy with or without radiation or immunotherapy during or after chemotherapy. Only trials with available data, either published or retrieved through personal communication, were included in the meta-analysis. Two reviewers (M. Al Khabori, J. R. de Almeida) independently screened all citations. Additional information (gathered from the full text and by contacting the authors) was retrieved for citations that were deemed potentially eligible. Disagreement was resolved by discussion.

Data Abstraction

Two reviewers (M. Al Khabori, J. R. de Almeida) abstracted the data independently using a common data collection form that had been piloted specifically for this review. Any disagreement on the data extracted from the studies was resolved by discussion and review of the full text or the original source. Corresponding authors of all eligible trials were contacted for missing and updated data on OS, EFS, TRM, MDS/AML, secondary malignancies, and

CONTEXT AND CAVEATS

Prior knowledge

Different chemotherapies have been combined with autologous stem cell transplantation (ASCT) in randomized clinical trials for follicular lymphoma, a type of non-Hodgkin lymphoma. The effect of these combinations on event-free survival varies, and a comparison of the survival benefits elicited by these different immune-chemotherapies has not been reported.

Study design

A systematic review and meta-analysis of randomized control trials comparing chemotherapy alone with chemotherapy and ASCT for follicular lymphoma were performed. Overall survival, event-free survival, and other adverse patient outcomes were compared for patients who received chemotherapy alone vs chemotherapy with ASCT. The quality of evidence for each analysis was also assessed.

Contribution

Seven randomized clinical trials met the predetermined eligibility criteria. Moderate quality evidence of three trials that reported overall survival indicated that high-dose chemotherapy with ASCT did not improve the overall survival of adult follicular lymphoma patients. On the basis of low-quality evidence from four trials, event-free survival was longer for patients who received chemotherapy with ASCT. Other adverse outcomes and the absolute risk of death from treatment did not differ between the two treatment arms.

Implications

Long-term follow-up of patients in trials is needed to better assess the effect of ASCT on event-free survival. Future clinical trials measuring the effect of combining ASCT with high-dose chemotherapy on overall survival are needed.

Limitations

Because of low-quality evidence, the suggested improvement in event-free survival for patients treated with chemotherapy, and ASCT should be investigated in additional trials with longer follow-up. Some data from unpublished trials were unavailable for analysis and may have led to potential publication bias.

From the Editors

questions regarding methodological issues and to confirm results and resolve questions on methodology of the trials. We received further data by personal communication for three trials from Dr C. Sebban (OS, EFS, TRM, method of sequence generation, allocation concealment, and number of patients who were lost to follow-up), Dr M. Ladetto (OS, EFS, TRM, MDS/AML, and secondary malignancies), and Dr Deconinck (OS, EFS, TRM, MDS/AML, solid tumor incidence, method of sequence generation, allocation concealment, and number of patients who were lost to follow-up).

Risk of Bias Assessment

The two reviewers independently abstracted information from six design elements related to risk of bias, including sequence generation, concealment of allocation, blinding, completeness of follow-up, selective reporting and other biases using categories of low, unclear, and high risk (18). We calculated the Cohen kappa coefficient (κ) to assess the agreement between the two reviewers

when categorizing the six elements and reported it as overall agreement. Any disagreement on the quality rating between the two reviewers was resolved by discussion and consensus. The lack of blinding was not considered to result in a high risk of bias, given the objective nature of the outcomes used in this systematic review (OS, EFS, TRM, MDS/AML, and secondary malignancies).

Outcome Measures

OS and EFS were defined according to the International Workshop criteria (19). OS was defined for all patients from the date of entry into the trial to the time of death from any cause. For patients with complete and partial response to chemotherapy, EFS was defined as the date of entry into the trial until death from any cause or progression of disease. We also reported definitions by authors not using the International Workshop criteria (19). TRM was defined as deaths related to chemotherapy or ASCT within 1 year of completion of treatment. Secondary malignancies were defined as any MDS, AML, or solid tumors that developed after chemotherapy or ASCT.

Analysis and Data Synthesis

For binary outcomes, the treatment effect was expressed as relative risk (RR) with its 95% confidence interval (CI). For the time to event outcomes, the treatment effect is expressed as hazard ratio (HR) with its 95% confidence interval. If the hazard ratio is not reported in the article and no information could be collected from the primary authors, we used Parmar methods (20,21) to estimate the hazard ratio from the available data. Cohen kappa coefficient (22) was calculated to measure the agreement between the two reviewers (M. Al Khabori, J. R. de Almeida) on the relevant studies identified at the screening and full text stages. All meta-analyses used random effect models. Heterogeneity was assessed by Cochrane χ^2 analysis and by measuring I^2 (23). We explored the following a priori hypotheses to explain heterogeneity: baseline risk [high risk defined as ≥ 3 factors on the FLIPLI or IPS (24,25)]; use of rituximab; follow-up (≤ 5 vs > 5 years); and type of analysis (intention-to-treat and per protocol). For the baseline risk, the subgroup analysis of interest was not reported in the published trials; therefore, we were only able to conduct the subgroup comparison between trials. Trials with high-risk patients comprising more than a third of the study population were considered as high risk. For the OS and EFS outcomes, we expected the benefit of ASCT to be greater in high-risk patients and in per protocol analyses and lower with use of rituximab and longer follow-up. For the TRM, MDS/AML, and secondary malignancy outcomes, we expected higher event rates with longer follow-up. Other analyses and χ^2 tests of subgroup interaction were performed using RevMan 5 (Review Manager 5 computer program, version 5.0. Copenhagen: The Nordic Cochrane Center, The Cochrane Collaboration, 2008) according to methods described by Deeks et al. (26). The GRADE criteria (27) were used to summarize the evidence gathered in the systematic review for each outcome in the following domains: risk of bias, inconsistency, indirectness, imprecision, and publication bias. The results were presented in an evidence profile. All statistical tests were two-sided and carried out at $\alpha = .05$. The P value for comparing heterogeneity between subgroups was calculated using Cochrane χ^2 (23).

Results

Search Results

The search strategy identified 1661 citations (Figure 1). After screening titles and abstracts, 1625 were excluded for the following reasons: duplicate publications, not RCT, and disease not FL. Full text or further details were retrieved for the remaining 36 citations. Of those, 23 duplicate reports and six non-RCT studies were excluded. Three RCTs (28–30) included in the systematic review did not have data available for pooling and were excluded from this meta-analysis. The κ statistics for the agreement between the two reviewers (M. Al Khabori, J. R. de Almeida) for the screening stage and the full text stage of the search were 0.72 and 0.82, respectively.

Characteristics of the Included Studies

Seven RCTs were identified (Table 1) and included in the systematic review, but only four RCTs with data from 941 patients were included in this meta-analysis (11–13,31) (Table 2), as outcome data were not available for pooling from three of the trials. All trials were multicenter (information was not available for Portlock, NCI-V89-0192), two-arm, parallel prospective RCTs. The sample sizes ranged from 67 (30) to 469 (29) adult participants with previously untreated FL. In the four trials with available data, the median age was approximately 50 years. High-risk patients constituted less than a third of the patients enrolled except in one trial (11) in which they represented 58% of the population. A variety of induction, mobilization, and high-dose therapy regimens were used in the transplant and chemotherapy arms of the four trials, as detailed in Table 1. In two of the trials, patients in both arms received rituximab during the induction treatment (11,30). The median follow-up duration was up to 5 years for two trials (11,12) and approximately 9 years for the other two trials (13,31). OS was

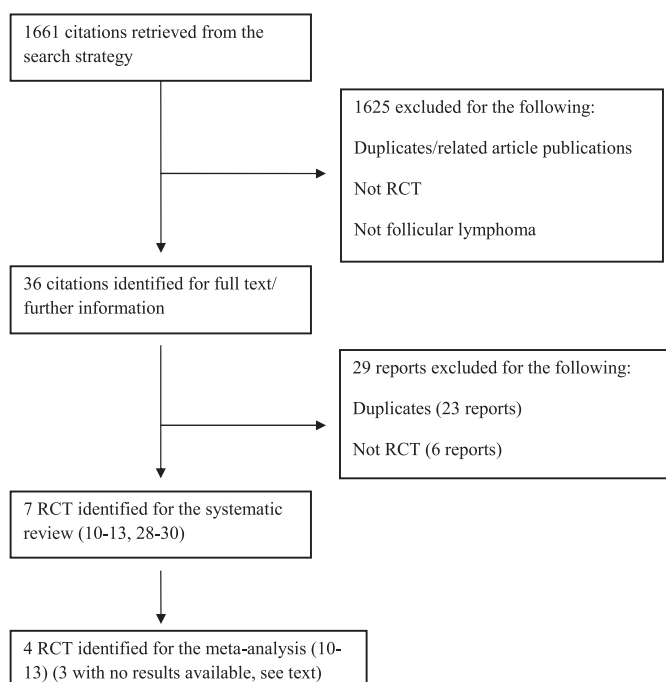


Figure 1. Study selection for the systematic review and meta-analysis. RCT = randomized clinical trial.

Table 1. Description of randomized clinical trials included in the systematic review*

First author, year published (reference)	Design	No. of patients	Trial description	Comparison
Lenz et al. 2004 (12)	Multicenter two arms parallel prospective controlled	240	Initial treatment: two cycles of CHOP or MCP chemotherapy. Two further cycles if not in CR. Mobilization: Dexa-BEAM. High-dose chemotherapy: combined total body irradiation (12 Gy in six applications) and cyclophosphamide (60 mg/kg body weight, i.v., on days -3 and -2).	Initial treatment: four cycles of CHOP or MCP chemotherapy. Two further cycles if not in CR. All treatments are supplemented by two extra cycles (to balance mobilization of chemotherapy in the other arm). Maintenance: subcutaneous IFN- α (5 mU) was administered three times weekly until disease progression was observed. Initial treatment: a six-course induction phase of CHVP regimen was administered monthly. This is followed, for responders and patients presenting with stable disease, by a maintenance phase that consisted of one cycle every 2 mo for 1 y. Concomitant subcutaneous IFN- α 2b was administered at 5 mU three times per week for 18 mo.
Deconinck et al. 2005 (10)	Multicenter two arms parallel prospective controlled	166	Initial treatment: VCAP regimen was administered every 3 wk. For patients in CR, VGPR, or PR after the second or third cycle VCAP, they continued stem cell harvesting. Before transplantation, patients received one course of IMVP16. Patients with less than PR after the VCAP chemotherapy received salvage therapy of two to three courses of DHAP. If at least a PR was obtained after DHAP, stem cells were harvested; otherwise patients went off study. High-dose chemotherapy: combined total body irradiation (12 Gy in six applications) and cyclophosphamide (60 mg/kg for 2 d).	Initial treatment: CHVP chemotherapy was administered for six cycles. During the study, teniposide was replaced by etoposide (100 mg/m ²) on day 1. IFN- α was given subcutaneously at a dosage of 5 mU three times a week. Patients achieving a CR or PR received six courses of CHVP plus IFN- α every 2 mo for 1 y.
Sebban et al. 2006 (13)	Multicenter two arms parallel prospective controlled	401	Initial treatment: CHOP chemotherapy was administered every 3 wk for four cycles. Mobilization: responding patients (CR or PR) received a single course of cyclophosphamide (4500 mg/m ²), etoposide (450 mg/m ²), and granulocyte colony-stimulation factor (300 μ g) from days 4–12 followed by peripheral blood stem cell harvest. High-dose chemotherapy: cyclophosphamide (60 mg/kg/d) and etoposide (150 mg/m ²) were administered from day -6 to -5. Total body irradiation was then performed delivering 10 Gy in five fractions.	Initial treatment: CHOP chemotherapy (11) for six courses followed by four courses of 375 mg/m ² rituximab. Patients in PR or who remained PCR positive received two final rituximab courses at the end of the program. Radiation therapy included 30–36 Gy on bulky sites or on residual masses 2 mo after finishing the treatment.
Ladetto et al. 2008 (11)	Multicenter two arms parallel prospective controlled	134	Initial treatment: doxorubicin, vincristine, prednisone chemotherapy, totaling four 75 mg/m ² doxorubicin administrations, was given. Patients not achieving CR received two additional DHAP treatments. Mobilization: etoposide (2 g/m ²) was administered followed by a chemotherapy-free interval of 40 d during which patients received two courses of rituximab (375 mg/m ²). Cyclophosphamide (7 g/m ²) was then delivered. Two further rituximab doses were given on the day after cyclophosphamide administration and on the first day the patient had a whole blood cell count greater than 1000 per μ L for the in vivo purging. Patients in PR or who remained PCR positive received two final rituximab courses at the end of the program. High-dose chemotherapy: mitoxantrone (60 mg/m ²) on day -5 and melphalan (180 mg/m ²) on day -2. Radiation therapy included 30–36 Gy delivered to bulky sites or to residual masses 2 mo post-ASCT.	Initial treatment: CHOP chemotherapy (11) for six courses followed by four courses of 375 mg/m ² rituximab. Patients in PR or who remained PCR positive received two final rituximab courses at the end of the program. Radiation therapy included 30–36 Gy on bulky sites or on residual masses 2 mo after finishing the treatment.

(Table continues)

Table 1 (Continued).

First author, year published (reference)	Design	No. of patients	Trial description	Comparison
NCI-V89-0192 (28)	Two arms parallel prospective controlled	106 (planned)	Marrow ablation and autologous bone marrow therapy will be administered.	Consolidative extended field radiotherapy will be administered.
EORTC NCT00003152 (29)	Multicenter two arms parallel prospective controlled	469 (planned)	High-dose chemotherapy with cyclophosphamide and total body irradiation followed by autologous stem cell transplantation will be performed.	IFN- α will be delivered subcutaneously three times per week for a maximum of 3 y.
Meckenstock et al. 2005 (30)	Two arms parallel prospective controlled	67 (64 evaluable)	Three courses of conventional CHOP followed by two courses of HAM preceding high-dose therapy with BEAM, and autologous peripheral blood stem cell transplantation. Rituximab was added to every cycle after the first cycle.	Six courses of fludarabine (25 mg/m ² on days 1–3) with mitoxantrone (10 mg/m ² on day 1) were administered. Rituximab was added to every cycle after the first cycle.

* ASCT = autologous stem cell transplantation; BEAM = carmustine, etoposide, cytarabine, and melphalan; CHOP = cyclophosphamide, doxorubicin, vincristine, prednisone; CHVP = cyclophosphamide, doxorubicin, teniposide, prednisone; CR = complete response; Dexam-BEAM = dexamethasone, carmustine, etoposide, cytarabine, melphalan; DHAP = dexamethasone, cytarabine, cisplatin; EORTC = European Organization for Research and Treatment of Cancer; HAM = cytarabine, mitoxantrone; IFN = interferon; IMVP16 = ifosfamide, methotrexate, etoposide; i.v. = intravenous; MCP = mitoxantrone, chlorambucil, prednisolone; NCI = National Cancer Institute; PCR = polymerase chain reaction; PR = partial response; VCAP = cyclophosphamide, high-dose doxorubicin, vincristine, prednisone; VGPR = very good partial response.

reported by three trials (11,13,31). The trial described by Lenz et al. (12) did not report OS, and the author did not provide the data. EFS was reported for three trials (11,13,31) with definitions consistent with that of the International Workshop criteria (19). In the study reported by Lenz et al. (12), only progression-free survival was reported as the disease control outcome.

Two (28,29) of three RCTs with no available data were found on the National Cancer Institute website (www.cancer.gov): Portlock (NCI-V89-0192) and the European Organization for Research and Treatment of Cancer (NCT00003152). These studies started in 1990 and 1997, respectively, and both terminated enrollment. Available details of the design and interventions are presented in Table 1. The study by Portlock (NCI-V89-0192) planned to include 106 patients with non-Hodgkin lymphoma (stage 3 or 4) in complete remission following ProMACE-MOPP chemotherapy (cyclophosphamide, doxorubicin, etoposide, mustargen, oncovirin, procarbazine, and prednisone). The NCT00003152 study planned to include 469 patients with advanced FL (stage 3 or 4) in partial or complete remission after eight 3-weekly courses of CVP chemotherapy (cyclophosphamide, vincristine, and prednisone). Both studies have not been published, and further information could not be retrieved from the corresponding investigators. The results of the third trial, reported by Meckenstock et al. (30), were reported at an interim analysis in a conference abstract. The available data on the patients and interventions are detailed in Table 1, although no further information could be retrieved from the corresponding author. Participants were adults aged less than 60 years with advanced FL. With a median follow-up of 23 months, both the EFS and OS were similar between the two arms (64 evaluable patients of 67 total enrolled).

Risk of Bias

Two reviewers (M. Al Khabori, J. R. de Almeida) independently assessed the reports with a very good overall agreement ($\kappa = 0.72$). The lowest agreement was in the other bias category in which either reviewer felt that limited reporting of quality domains might have biased the results. Of four trials with available data (11–13,31), one trial reported their sequence generation (11) and two others provided sequence generation by personal communication (13,31); all were computer generated and consistent with low risk of bias. Allocation concealment (central randomization) was reported by the four trials (11–13,31). Patients and physicians were not blinded in any of the trials and it was not stated (not available through personal communication), if blinding was used at the time of data collection, analysis, or article preparation. Only the trial reported by Ladetto et al. (11) specifically stated the loss to follow-up and methods used to deal with the loss to follow-up. All trials performed the analysis as intention to treat except for one trial (12) in which analysis was reported as per protocol, although *P* was stated for the intention-to-treat analysis. One trial (11) was stopped early because benefit in EFS favoring ASCT was observed.

Outcomes

The three studies that reported OS (11,13,31) enrolled 701 patients and reported no statistically significant difference in OS (a pooled HR of OS = 0.99, 95% CI = 0.73 to 1.33, $P_{\text{heterogeneity}} = .65$; $P = 0\%$, 95% CI for $P = 0\%$ to 90%) (Table 3 and Figure 2). The

Table 2. Patient characteristics of published autologous stem cell transplant studies included in the meta-analysis*

First author, year published (reference)	Total No. of patients (men, women)	Median age, y (range)	No. of high-risk patients (%)	No. of patients with B symptoms (%)†	No. of patients with poor performance status (%)‡	No. of patients with bulky disease (%)	Stem cell harvest	Study sponsor(s)
Lenz, 2004 (12)§	240 (118, 122)	49.1 (29–59)	41 (17.1)	87 (36.3)	13 (5.4)	NR	At least 2 × 10 ⁶ CD34 ⁺ cells per kg of body weight (and 2 × 10 ⁶ CD34 ⁺ cells per kg body weight as backup) were required	Supported in part by a grant of the Deutsche Krebshilfe (project number 70-2208-Hi 2)
Deconinck, 2005 (10)¶	166 (85, 81)	51 and 50 (29–61)	46 (27.7)	NR	10 (6.0)	53 (31.9)	NR	French Ministere de la Sante et de la Solidarite Sociale and the Schering-Plough.
Sebban, 2006 (13)#	401 (221, 180)	49	120 (29.9)	102 (25.4)	29 (7.2)	266 (66.3)	4 × 10 ⁶ mononuclear cells per kg of body weight were required	French Programme Hospitalier de Recherche Clinique (PHCRC-1994) and by a grant from Schering-Plough.
Ladetto, 2008 (11)**	134 (78, 56)	51 (25–59)	78 (58.2)	63 (47.0)	80 (59.7)	75 (55.9)	A minimum of 5 × 10 ⁶ CD34 ⁺ cells per kg of body weight was required for autologous PBSCs only (plus at least 3 × 10 ⁶ CD34 ⁺ cells per kg of body weight or a bone marrow harvest as backup) were required.	Roche (Milan, Italy), Compagnia di San Paolo (Torino, Italy), Ministero Italiano Universita e Ricerca (Rome, Italy), and Regione Piemonte (Torino, Italy)

* NR = not reported; PBSC = peripheral blood stem cells.

† Fever >38°C, weight loss, night sweats.

‡ Patients with an Eastern Cooperative Oncology Group score of 2 or more were designated as having poor performance status.

§ Patients with an International Prognostic Index or Follicular Lymphoma International Prognostic Index score of greater than 2 were designated high risk.

¶ Median age of patients was 51 years for the transplant arm and 50 years for the chemotherapy arm. For this study, patients with a Follicular Lymphoma International Prognostic Index score of greater than 2 were designated high risk.

For this study, bulky disease was defined as tumors greater than 7 cm in diameter.

** For this study, bulky disease was defined as tumors greater than 5 cm in diameter.

Table 3. Evidence profile of autologous stem cell transplantation studies used for meta-analysis*

Study characteristics	Data quality assessment							Summary of findings			
	Type of study	No. of studies	No of patients	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Relative effect† (95% CI)	Estimation of absolute event rates (derived from observed events in study)	
										Event rate with chemotherapy only	Event rate with ASCT (95% CI)
Overall survival	3	701	No serious limitations, one RCT stopped early for EFS benefit	Consistent	Direct	Imprecise (95% CI includes important benefit and harm)	Unlikely (tends to bias toward positive)	HR = 0.99 (0.73 to 1.33)	Not available	Not available	Moderate (imprecision)
EFS	4	941	Serious limitation, one study stopped early, another study PP serious limitation, one RCT stopped early	Inconsistent (explained by a priori hypotheses) Consistent	Direct	Precise	Likely	HR = 0.54 (0.36 to 0.82)	Not available	Not available	Low (inconsistency, risk of bias, publication bias)
Treatment-related mortality	3	701	Serious limitation, one RCT stopped early	Consistent	Direct	Imprecise (95% CI includes benefit and harm)	Likely	RR = 1.04 (0.29 to 3.70)	14 per 1000	15 per 1000 (4 to 52)	Low (imprecision, risk of bias, likely publication bias)
MDS/AML	3	701	Serious limitation, one study stopped early	Inconsistent (inconsistency not fully explained by a priori hypotheses) Consistent	Direct	Imprecise (95% CI includes benefit and harm)	Unlikely	RR = 2.19 (0.45 to 10.55)	17 per 1000	37 per 1000 (8 to 170)	Low (inconsistency, imprecision, risk of bias)
Secondary solid tumors	3	701	Serious limitation, one study stopped early	Consistent	Direct	Imprecise (95% CI includes benefit and harm)	Unlikely	RR = 1.30 (0.33 to 5.08)	34 per 1000	44 per 1000 (11 to 173)	Low (risk of bias, imprecision)

* ASCT= autologous stem cell transplantation; CI = confidence interval; CR = complete response; EFS = event-free survival; HR = hazard ratio; MDS/AML = secondary myelodysplasia/acute myeloid leukemia; PP = per protocol; RCT = randomized controlled trial; RR = relative risk.

† The pooled survival probability for the chemotherapy arm across all studies is not available given the lack of individual patient data.

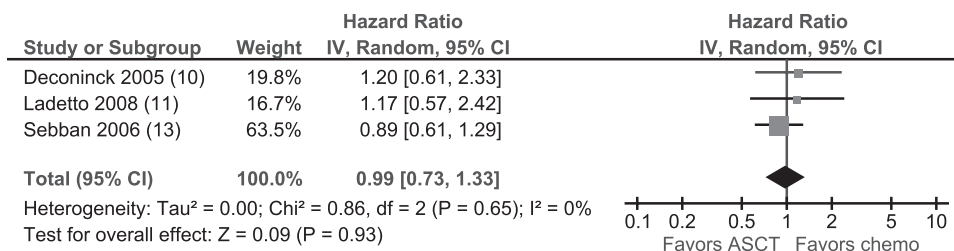


Figure 2. Forest plot of overall survival. P was calculated according to the method described by Deeks et al. (23). The first author, year of publication, and reference are given for each study. The hazard ratios (boxes) with 95% confidence intervals (CIs, horizontal lines) were calculated, and the pooled hazard ratio (diamond) was estimated with generic inverse variance (IV) using random effect analysis. The

weight of each study's contribution to the overall effect calculated from random effect analyses is given and is also reflected by the size of the box. All statistical tests were two-sided. The P value for comparing heterogeneity between subgroups was calculated using Cochrane χ^2 (23). ASCT = autologous stem cell transplantation; chemo = chemotherapy.

quality of evidence was rated moderate because of imprecision (Table 3). Also, an estimation of the absolute event rates (derived from observed events in the study) for patients who received chemotherapy only and ASCT were unavailable (Table 3).

All four trials, with a total of 941 patients, suggested an improvement in EFS in favor of ASCT (11–13,31) (pooled HR of EFS = 0.54, 95% CI = 0.36 to 0.82; $P_{\text{heterogeneity}} = .002$; $F = 80\%$, 95% CI for $F = 48\%$ to 93%) (Figure 3). This heterogeneity was because of the duration of follow-up and a statistically significant subgroup effect ($P < .001$). The median follow-up was 4–5 years for two trials (11,32) (pooled HR of the two trials = 0.38, 95% CI = 0.28 to 0.52, $F = 0\%$) and 9 years for the other two (13,31) (pooled HR for the two trials = 0.77, 95% CI = 0.61 to 0.97, $F = 14\%$) (Figure 3). Comparing the trial using rituximab as part of the initial therapy vs those that did not, the EFS was statistically significantly decreased (HR of trials with rituximab = 0.38, 95% CI = 0.25 to 0.59 vs pooled HR of trials not employing rituximab = 0.60, 95% CI = 0.38 to 0.95; $F = 79\%$; P for subgroup effect = .02). The quality of evidence (27) was rated low because of inconsistency, early termination, and potential publication bias (Table 3).

Three studies reported TRM (11,13,31) with five treatment-related deaths among 346 patients who underwent ASCT and five deaths among 355 patients who received chemotherapy. The difference in TRM between the two arms was not statistically significant (pooled RR = 1.04, 95% CI = 0.29 to 3.70, $P_{\text{heterogeneity}} = .51$; $F = 0\%$, 95% CI for $F = 0\%$ to 90%). The absolute risk across the chemotherapy arms was 14 deaths per 1000 patients, whereas the calculated absolute risk with ASCT on the basis of the above relative risk is 15 in 1000 (range = 4–52 deaths per 1000 patients). The quality of evidence (27) was rated low because of imprecision, early termination of one RCT, and potential publication bias (Table 3).

Data regarding the development of MDS/AML were reported for three trials (11–13,31). In the transplant arm, 13 of 346 patients developed MDS/AML during follow-up compared with six of 355 patients in the chemotherapy arm. The difference in MDS/AML rates between the two arms was not statistically significant (pooled RR = 2.19, 95% CI = 0.45 to 10.55, $P_{\text{heterogeneity}} = .14$; $F = 48\%$, 95% CI for $F = 0\%$ to 85%) (Figure 4). The absolute risk across the chemotherapy arms was 17 in 1000, whereas the calculated absolute risk with ASCT strategy on the basis of the above

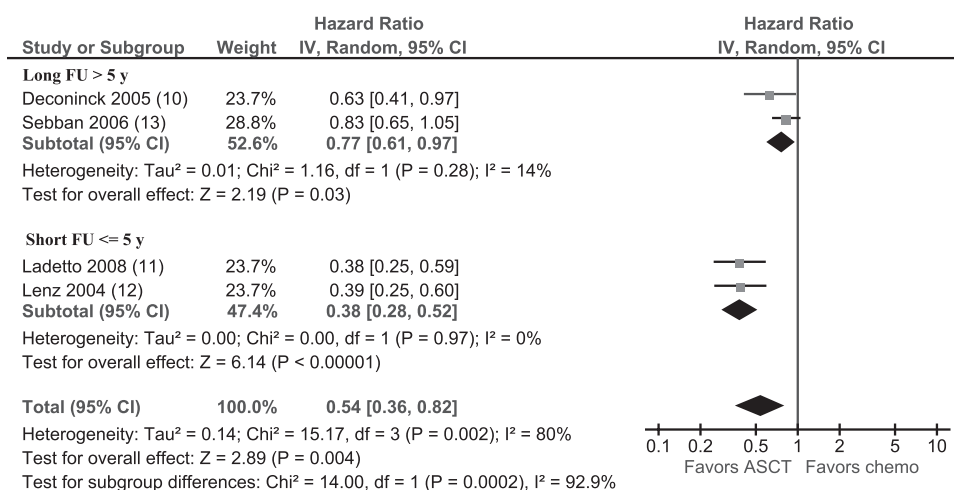


Figure 3. Forest plot of event-free survival comparing studies with long- vs short-term follow-up (FU). P was calculated according to the method described by Deeks et al. (23). The first author, year of publication, and reference are given for each study. The hazard ratios (box) with 95% confidence intervals (CIs, horizontal lines) were calculated for each study, and the pooled hazard ratio (diamond) was estimated with

generic inverse variance (IV) using random effect analysis. The weight of each study's contribution to the overall effect calculated by random effect analyses is given and reflected by the size of each box. All statistical tests were two-sided. The P value for comparing heterogeneity between subgroups was calculated using Cochrane χ^2 (23). ASCT = autologous stem cell transplantation; chemo = chemotherapy.

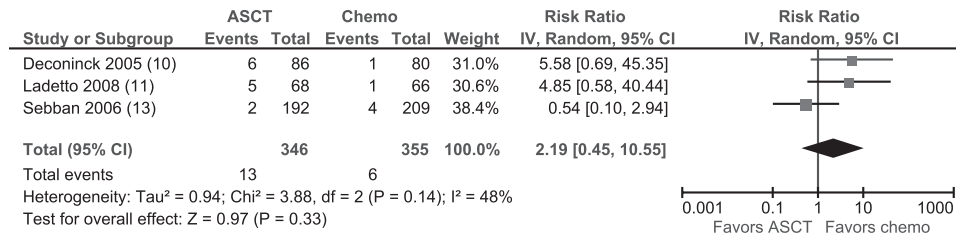


Figure 4. Forest plot of myelodysplastic syndrome and acute myeloid leukemia events comparing the transplantation vs. chemotherapy arms. *P* was calculated according to the method described by Deeks et al. (23). The first author, year of publication, and reference are given for each study. The number of new patients with myelodysplastic syndrome/acute myeloid leukemia and the total number of patients evaluable in the study, respectively, in the ASCT arm are given. The number of new patients diagnosed with myelodysplastic syndrome/acute myeloid leukemia and the total number of patients evaluable in the study in the

chemotherapy arm are also shown. The relative risk (boxes) with corresponding 95% confidence intervals (CIs, horizontal lines) for each study was calculated, and the pooled relative risk (diamond) was estimated with generic inverse variance (IV) by random effect analysis. Also, the weight of the contribution to the overall effect of each study was calculated from random effect analyses and is also reflected by the size of each box. All statistical tests were two-sided. The *P* value for comparing heterogeneity between subgroups was calculated using Cochran χ^2 (23). ASCT = autologous stem cell transplantation; chemo = chemotherapy.

relative risk is 37 deaths in 1000 patients (range = 8–170 in 1000). Most of the heterogeneity was created by the study reported by Sebban et al. (13). Baseline risk, use of rituximab, or follow-up duration did not explain this heterogeneity (for all analyses, $P_{\text{subgroup effect}} = .32$). The quality of evidence (27) was rated as low because of inconsistency, imprecision, and early termination of one RCT (Table 3).

Three trials (11,13,31) reported data on solid tumors. There were 16 and 12 solid tumors, respectively, observed in a sample of 346 ASCT patients and 355 standard-dose therapy patients. The solid tumor rates for the two treatment arms were similar (pooled RR for second solid tumors = 1.30, 95% CI = 0.33 to 5.08, $P_{\text{heterogeneity}} = .20$; $I^2 = 38\%$, 95% CI for $I^2 = 0\%$ to 81%). The absolute risk across the chemotherapy arms was 34 in 1000 patients, whereas the calculated absolute risk with ASCT strategy on the basis of the above relative risk is 44 in 1000 (range = 11–173 in 1000). The quality of evidence (27) was rated low because of imprecision and early termination of one RCT (Table 3).

Discussion

In this systematic review and meta-analysis, we found moderate quality evidence that high-dose chemotherapy and ASCT did not improve OS in adults with previously untreated FL. EFS was superior for the transplant arm compared with the chemotherapy arm, albeit with substantial heterogeneity. This survival benefit comes at the expense of MDS/AML at a relative risk that cannot exclude a large increase risk in the transplant arm. The TRM and rates of secondary solid tumors were similar between those who received chemotherapy only vs ASCT.

The estimate of the hazard ratio for OS was rated moderate on evidence quality because of imprecision. Although OS comparisons were not reported in all of the trials, we do not anticipate publication bias as this would likely bias the pooled estimate toward a positive treatment effect. Use of chemoimmunotherapy including rituximab in subsequent treatment for the non-transplanted patients may have contributed to the lack of survival difference (9). The higher rate of MDS/AML in the transplant arm in trials with longer follow-up may contribute to the lack of a difference in survival that was observed.

The EFS was rated low on evidence quality on the basis of the inconsistency, per protocol analysis used by one of the studies contributing to the pooled estimate (12), and the likelihood of publication bias. Three trials (28–30) found through the systematic search had no published data and were not included in the meta-analysis. One of the three excluded trials showed lack of benefit of the transplantation strategy in an interim analysis (30). This publication bias may have biased the results toward benefit and, if included, could change the pooled estimate toward no difference or even harm. In our review, we pooled the hazard ratios of progression-free survival reported by Lenz et al. (12) with hazard ratios of EFS from the other trials. Pooling hazard ratios of progression-free survival with hazard ratios of EFS would likely increase the apparent statistical significance of the difference in EFS and for our study, bias the pooled hazard ratio in favor of the ASCT assuming similar induction failures in both treatment arms.

Although the heterogeneity in the EFS was substantial, it could be explained by the duration of follow-up. This subgroup analysis suggests that on long-term follow-up, the treatment effect disappears. Although this subgroup analysis was between trials, it was planned a priori, and likely represents a true subgroup effect (33). Furthermore, it is consistent with the natural history of this indolent lymphoma (16) and the tendency for patients to experience multiple relapses. No curative treatments are available, except for a highly selected group receiving allogeneic stem cell transplantation (34).

The serious methodological limitations (early termination and per protocol analysis) along with the imprecision and the likelihood of publication bias resulted in a low evidence quality rating for TRM. The MDS/AML outcome was also rated low for the inconsistency and the limitation of early termination of one trial. None of the a priori hypotheses to explore heterogeneity explained this inconsistency. The pooled estimate with the 95% confidence interval suggests large potential harm that may be introduced by ASCT, although it also includes the small potential harm of chemotherapy.

Because the introduction of monoclonal antibody therapy for indolent lymphoma is relatively recent, only two trials (11,30) used rituximab as part of their treatment in all patients in both treatment arms. The relevance of any conclusions regarding the

effectiveness of therapy intensification with ASCT in light of the evolution of the standard of care for FL remains limited given the survival benefit shown with adding rituximab to chemotherapy (35). For example, the median time to progression was 32 months after rituximab-containing chemotherapy vs 15 months for chemotherapy alone ($P < .001$) in a landmark clinical trial (35). In addition, maintenance rituximab improves the median progression-free survival compared with observation (51.5 months for rituximab vs 14.9 months for observation, $P < .001$), as shown in relapse-refractory settings (14). The subgroup effect of trials using rituximab in our analysis suggests that use of rituximab has a synergistic effect when combined with high-dose chemotherapy and ASCT. The interpretation of this finding is limited by the publication bias of at least one other trial (30) that used rituximab and found a non-statistically significant difference.

There are a number of limitations associated with our systematic review and meta-analysis. The lack of complete data from unpublished trials may have led to a potential publication bias. Trials with no statistically significant treatment effect or those that stopped early because of toxic effects in the ASCT arm are more likely not to be published. This tendency would bias EFS toward an apparent benefit. No formal test of publication bias was done because of the small number of studies analyzed. The other limitation was the power, especially in the OS outcome. With a pooled sample size of 701 patients, we have 80% power to detect a minimum treatment effect (HR = 0.65) using a two-sided alpha of 0.05 (36). Therefore, one cannot rule out a smaller effect size than that of a hazard ratio of 0.65. Finally, the subgroup analyses were done comparing treatment effects between trials and not within trials as those subgroup analyses were not reported for each trial.

The current systematic review and meta-analysis showed no benefit for high-dose chemotherapy and ASCT in OS for patients with previously untreated FL. Long-term follow-up is needed from clinical trials to better estimate the impact of ASCT on EFS and the risk of secondary MDS/AML. In addition, data from the unpublished trials should be made available for a more complete assessment of the treatment effect, and to guide the development of future clinical trials. Trials of ASCT in the context of current chemoimmunotherapy approaches in FL are needed to further evaluate the ability of intensification of therapy using ASCT to improve OS.

References

- Anderson JR, Armitage JO, Weisenburger DD. Epidemiology of the non-Hodgkin's lymphomas: distributions of the major subtypes differ by geographic locations. Non-Hodgkin's Lymphoma Classification Project. *Ann Oncol*. 1998;9(7):717-720.
- Morton LM, Wang SS, Devesa SS, et al. Lymphoma incidence patterns by WHO subtype in the United States, 1992-2001. *Blood*. 2006;107(1):265-276.
- Gribben JG. How I treat indolent lymphoma. *Blood*. 2007;109(11):4617-4626.
- Ezdinli EZ, Harrington DP, Kucuk O, et al. The effect of intensive intermittent maintenance therapy in advanced low-grade non-Hodgkin's lymphoma. *Cancer*. 1987;60(2):156-160.
- Jones SE, Grozea PN, Metz EN, et al. Superiority of adriamycin-containing combination chemotherapy in the treatment of diffuse lymphoma: a Southwest Oncology Group study. *Cancer*. 1979;43(2):417-425.
- Nickenig C, Dreyling M, Hoster E, et al. Combined cyclophosphamide, vincristine, doxorubicin, and prednisone (CHOP) improves response rates but not survival and has lower hematologic toxicity compared with combined mitoxantrone, chlorambucil, and prednisone (MCP) in follicular and mantle cell lymphomas: results of a prospective randomized trial of the German Low-Grade Lymphoma Study Group. *Cancer*. 2006;107(5):1014-1022.
- Rohatiner AZ, Gregory WM, Peterson B, et al. Meta-analysis to evaluate the role of interferon in follicular lymphoma. *J Clin Oncol*. 2005;23(10):2215-2223.
- Marcus R, Imrie K, Solal-Celigny P, et al. Phase III study of R-CVP compared with cyclophosphamide, vincristine, and prednisone alone in patients with previously untreated advanced follicular lymphoma. *J Clin Oncol*. 2008;26(28):4579-4586.
- Hochster H, Weller E, Gascogne RD, et al. Maintenance rituximab after cyclophosphamide, vincristine, and prednisone prolongs progression-free survival in advanced indolent lymphoma: results of the randomized phase III ECOG1496 Study. *J Clin Oncol*. 2009;27(10):1607-1614.
- Deconinck E, Foussard C, Milpied N, et al. High-dose therapy followed by autologous purged stem-cell transplantation and doxorubicin-based chemotherapy in patients with advanced follicular lymphoma: a randomized multicenter study by GOELAMS. *Blood*. 2005;105(10):3817-3823.
- Ladetto M, De Marco F, Benedetti F, et al. Prospective, multicenter randomized GITMO/IIL trial comparing intensive (R-HDS) versus conventional (CHOP-R) chemoimmunotherapy in high-risk follicular lymphoma at diagnosis: the superior disease control of R-HDS does not translate into an overall survival advantage. *Blood*. 2008;111(8):4004-4013.
- Lenz G, Dreyling M, Schiegnitz E, et al. Myeloablative radiochemotherapy followed by autologous stem cell transplantation in first remission prolongs progression-free survival in follicular lymphoma: results of a prospective, randomized trial of the German Low-Grade Lymphoma Study Group. *Blood*. 2004;104(9):2667-2674.
- Sebban C, Mounier N, Brousse N, et al. Standard chemotherapy with interferon compared with CHOP followed by high-dose therapy with autologous stem cell transplantation in untreated patients with advanced follicular lymphoma: the GELF-94 randomized study from the Groupe d'Etude des Lymphomes de l'Adulte (GELA). *Blood*. 2006;108(8):2540-2544.
- van Oers MH, Klasa R, Marcus RE, et al. Rituximab maintenance improves clinical outcome of relapsed/resistant follicular non-Hodgkin lymphoma in patients both with and without rituximab during induction: results of a prospective randomized phase 3 intergroup trial. *Blood*. 2006;108(10):3295-3301.
- Friedberg JW. Treatment of follicular non-Hodgkin's lymphoma: the old and the new. *Semin Hematol*. 2008;45(3) (Suppl. 2):S2-S6.
- Illidge T, Chan C. How have outcomes for patients with follicular lymphoma changed with the addition of monoclonal antibodies? *Leuk Lymphoma*. 2008;49(7):1263-1273.
- Vitolo U, Ferreri AJ, Montoto S. Follicular lymphomas. *Crit Rev Oncol Hematol*. 2008;66(3):248-261.
- Higgins J, Altman D. Chapter 8: assessing risk of bias in included studies. In: Higgins J, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0. The Cochrane Collaboration, 2011. www.cochrane-handbook.org.
- Cheson BD, Horning SJ, Coiffier B, et al. Report of an international workshop to standardize response criteria for non-Hodgkin's lymphomas. NCI Sponsored International Working Group. *J Clin Oncol*. 1999;17(4):1244.
- Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. *Stat Med*. 1998;17(24):2815-2834.
- Tierney JF, Stewart LA, Ghersi D, et al. Practical methods for incorporating summary time-to-event data into meta-analysis. *Trials*. 2007;8:16.
- Cohen JA. Coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20(1):37-46.
- Deeks J, Higgins J, Altman D. Chapter 9: analysing data undertaking meta-analyses. In: Higgins J, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. West Sussex, UK: John Wiley & Sons Ltd.; 2008:243-296.

24. Hermans J, Krol AD, van Groningen K, et al. International Prognostic Index for aggressive non-Hodgkin's lymphoma is valid for all malignancy grades. *Blood*. 1995;86(4):1460–1463.
25. Solal-Celigny P, Roy P, Colombat P, et al. Follicular lymphoma international prognostic index. *Blood*. 2004;104(5):1258–1265.
26. Deeks J, Altman D, Bradburn M. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: Egger M, Davey Smith G, Altman D, eds. *Systematic Reviews in Health Care: Meta-analysis in Context*. 2nd ed. London, UK: BMJ Publication Group; 2001:285–312.
27. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924–926.
28. NCI-V89-0192. Phase III Comparison of Consolidation with Extended Field Radiotherapy vs Autologous Bone Marrow Therapy for Patients with Advanced Low-Grade Non-Hodgkin's Lymphoma in First CR Following ProMACE-MOPP. <http://www.cancer.gov/clinicaltrials/search/view?cdrid=76136&version=HealthProfessional&protocolsearchid=7801137>. Accessed October 21, 2011.
29. NCT00003152. Marrow-Ablative Chemo-Radiotherapy and Autologous Stem Cell Transplantation Followed by Interferon-Alpha Maintenance Treatment Versus Interferon-Alpha Maintenance Treatment Alone After a Chemotherapy-Induced Remission in Patients With Stages III or IV Follicular Non-Hodgkin's Lymphoma. A Prospective, Randomized, Phase III Clinical Trial. www.Clinicaltrials.gov. Accessed March 8, 2009.
30. Meckenstock G, Kobbe G, Kronenwett R, et al. First line treatment of advanced follicular lymphoma comparing fludarabine/mitoxantrone and high-dose chemotherapy, each combined with CD20-antibody rituximab. In: Anonymous, ed. *Onkologie*; 2005:153.
31. Gyan E, Foussard C, Bertrand P, et al. High-dose therapy followed by autologous purged stem cell transplantation and doxorubicin-based chemotherapy in patients with advanced follicular lymphoma: a randomized multicenter study by the GOELAMS with final results after a median follow-up of 9 years. *Blood*. 2009;113(5):995–1001.
32. Lenz G, Unterhalt M, Haferlach T, et al. Significant increase of secondary myelodysplasia and acute myeloid leukemia after myeloablative radiochemotherapy followed by autologous stem cell transplantation in indolent lymphoma patients: Results of a prospective randomized study for the GLSG. *Blood*. 2003;102(11):986a–987a.
33. Sun X, Briel M, Busse JW, et al. Subgroup Analysis of Trials Is Rarely Easy (SATIRE): a study protocol for a systematic review to characterize the analysis, reporting, and claim of subgroup effects in randomized trials. *Trials*. 2009;10(1):101.
34. van Besien K. Allogeneic stem cell transplantation in follicular lymphoma: recent progress and controversy. *Hematology Am Soc Hematol Educ Program*. 2009;610–618.
35. Marcus R, Imrie K, Belch A, et al. CVP chemotherapy plus rituximab compared with CVP as first-line treatment for advanced follicular lymphoma. *Blood*. 2005;105(4):1417–1423.
36. Freedman LS. Tables of the number of patients required in clinical trials using the logrank test. *Stat Med*. 1982;1(2):121–129.

Funding

This work was done with no funding.

Notes

We thank Esther Atkinson, PMH (Toronto, ON, Canada) for helping with planning literature search. We thank the following authors who contributed with missing data on “Results” and “Methods”: Dr C. Sebban (Department of Hematology of Centre Le on Berard, Lyon, France), Dr E. Deconinck (Franche-Comte University and Department of Hematology, Jean Minjot Hospital, Besanc, France), and Dr M. Ladetto (Divisione Universitaria di Ematologia, Cattedra di Ematologia, Torino, Italy).

The authors were responsible for the study design, and analysis and interpretation of the data, and preparation of the article. M. Al Khabori and J. R. de Almeida were responsible for data acquisition.

Affiliations of authors: Department of Medical Oncology and Hematology, Princess Margaret Hospital, University of Toronto, Toronto, ON, Canada (MAK, JK, MC); Department of Hematology, Sultan Qaboos University Hospital, Al-Khod, Oman (MAK); Department of Otolaryngology, Head and Neck Surgery, Sunnybrook Hospital, University of Toronto, Toronto, ON, Canada (JRdA); CLARITY Research Group, Department of Medicine, McMaster University, Hamilton, ON, Canada (GHG).