ELSEVIER

Journal of Clinical Epidemiology 54 (2001) 747–754

Journal of Clinical Epidemiology

Evaluating treatment strategies in chronic lymphocytic leukemia: Use of quality-adjusted survival analysis

Vincent Levy[a,b], Raphaël Porcher[a,b], Fabrice Delabarre[a], Michel Leporrier[c], Bruno Cazin[d], Sylvie Chevret[a,b,*], for the French Cooperative CLL Group

[a]*Département de Biostatistique et Informatique Médicale, Hôpital Saint Louis, 1 Avenue Claude Vellefaux, 75475 Paris, France,* [b]*U444–INSERM, Paris, France,* [c]*Service d'Hématologie Clinique, CHRU de Caen, Avenue Georges Clémenceau, 14033 Caen, France,* [d]*Service des Maladies du Sang, CHRU de Lille, Hôpital Claude Huriez, 59037 Lille, France*

Received 29 June 2000; received in revised form 1 November 2000; accepted 8 November 2000

Abstract

To assess comparatively, in terms of quality-adjusted survival, three front-line treatments in patients with stage B- or C-chronic lymphocytic leukemia (CLL). To describe better and compare the survival after randomization of patients from the CLL90 trial that randomly compared ChOP (cyclophosphamide, doxorubicin, oncovin, prednisone), CAP (cyclophosphamide, doxorubicin, prednisone) and fludarabine in advanced CLL, we performed a quality-adjusted survival analysis. This consisted of defining four clinical states (toxicity, treatment free of toxicity, no treatment nor symptoms, relapse), then summing up the average times spent in each state weighted by utility coefficients that reflect relative value according to quality of life. The resulting quality-adjusted time without symptoms or toxicity (Q-TWIST) was compared between randomized groups, and sensitivity (threshold) analyses to the choice of utility coefficients was performed. Over 73 months after randomization, the fludarabine group gained a mean of 45 days of toxicity-free survival at CAP, and 61 days over ChOP. The mean TWIST was 27.05 months with CAP, 31.5 months with ChOP and 32.95 months with fludarabine. The threshold analyses showed that, whatever the utility weights, the mean Q-TWIST was always greater with ChOP or fludarabine as compared to CAP. Fludarabine was consistently a better treatment than ChOP, except in the unlikely case of high utility weights attributed to toxicity and low utility weights attributed to treatment. Nevertheless, from a clinical point of view, differences between ChOP and fludarabine were moderate or event slight (mean difference in TWIST of 1.45 months). We conclude that patients with advanced CLL have a moderate benefit in terms of Q-TWIST when treated with fludarabine over ChOP. These two treatments are always superior to CAP.

Keywords: Quality of life; Q-TWIST; Chronic lymphocytic leukemia; Treatment*

1. Introduction

Chronic lymphocytic leukemia (CLL) is the most common leukemia in Europe and North America [1,2] and is considered to be an incurable disease. Actually, its management is palliative and disease-related symptoms directed. The different randomized clinical trials conducted in CLL have clearly demonstrated that conventional chlorambucil schedules do not prolong survival of previously untreated stage A patients [3,4]. More recently, fludarabine, a purine analogue, has been used in refractory or relapsed CLL with an impressive remission rate of 30–70% including 10–15% complete remissions (CR), defined by the absence of any lymphadenopathy or spleen enlargement and normal blood cell count. In previously untreated patients, reported remission rates lie about 80%, with CR rates up to 60% [5–8]. A randomized clinical trial was initiated by our group in 1990 in advanced stages of CLL, to assess comparatively first-line treatment with fludarabine and two anthracyclin-containing regimens, namely CAP (cyclophosphamide, adriamycine, prednisone) and ChOP (cyclophophamide, adriamycine, vincristine, prednisone). Endpoints were overall survival (OS), treatment remission rate and tolerance (toxicity of the treatment). On February 9, 1996, the observation of response and survival rates lower with CAP than with fludarabine and ChOP prompted us to close the CAP arm, while no conclusive findings concerning ChOP and fludarabine were reached, although there was a trend toward better results with fludarabine [9]. Accrual in these two arms were stopped on April 15, 1998, due to achievement of the scheduled sample size. Analysis at the reference date of January

* Corresponding author. Département de Biostatistique et Informatique Médicale, Hôpital Saint Louis, 1 avenue Claude Vellefaux, 75475, Paris, France. Tel.: 33 (1) 42 49 97 42; fax: 33 (1) 42 49 97 45.

E-mail address: chevret@dbim.jussieu.fr (S. Chevret)

0895-4356/01/$ – see front matter
PII: S0895-4356(00)00359-0

1, 1999, based on a median follow-up of 58 months, confirmed the difference in response rates, but with no subsequent difference in survival times between CAP, ChOP and fludarabine [10].

To integrate quality-of-life considerations into the comparison of treatments that are being evaluated in randomized clinical trials, quality-adjusted survival analyses have been proposed. The so-called "Q-TWiST" approach (Quality-Adjusted Time Without Symptoms or Toxicity) estimates the mean time spent in a series of ordered clinical health states that differ in terms of quality of life. These states are weighted according to relative values regarding their quality of life, and the weighted mean health state durations are then used to compare the treatments in terms of quality-of-life-adjusted survival [11–15]. Such a partitioned survival analysis was developed in the context of multiple endpoints, when one treatment shows some advantage on some endpoints but a disadvantage on others, particularly applicable to trials of cancer chemotherapy in breast cancer [13,16,17], melanoma [18] or rectal cancer [17]. This approach for comparing cancer treatments has been presented by Feldstein [19]. Besides later extensions to therapies for HIV infection [20] or multiple sclerosis [21,22], it has been also applied to hematological diseases such as multiple myeloma [23,24], non Hodgkin lymphoma [18], and acute myeloid leukemia in childhood [25].

To get further insight in the comparison of first-line treatment strategies in CLL, we decided to analyze the CLL 90 trial using the Q-TWiST method.

2. Materials and methods

2.1. Patients and treatments

The study design and original results from the "CLL 90" trial have been described elsewhere [10] and will be only summarized here. This multicenter randomized clinical trial was open in France, Algeria and Brazil for untreated patients, who had Binet's stage B- or C-CLL [26]. All patients were scheduled to receive six monthly courses of either ChOP, CAP or fludarabine. In case of no response after three courses of CAP or fludarabine, patients were switched to the alternate treatment for six further courses. After relapse, treatment decision was left to the judgement of each center.

A total of 938 patients were registered onto the study on April 15, 1998, 357 (38%) of whom randomized to receive ChOP, 240 (26%) to receive CAP and 341 (36%) to receive fludarabine. The analysis was made on an intent-to-treat basis. The number of patients in the CAP arm was lower than that in the two other arms because of the early closure of this arm for lower response rate and survival after the first interim analysis [9]. Median age of the patients was 62 years (25th–75th percentiles: 56–68). Seventy-one percent of the patients were males. Of the 938 randomized patients, 15 (2%) did not receive one course of either chemotherapy and were excluded from the following analyses, as well as the 17 patients (2%) lost to follow-up and the 26 (3%) patients for whom the case report form of toxic adverse events was not available. Therefore, 880 patients (91%) were analyzed, 334 (38%) in the ChOP group, 232 (26%) in the CAP group and 314 (36%) in the fludarabine group, with 612 (70%) patients enrolled as stage B and 268 (30%) as stage C.

2.2. Derivation of the Q-TWiST

To specifically handle the trial data, four clinical health states that were relevant for treatment decision making in CLL were first defined: the time spent with toxicity (TOX) due to chemotherapy for either regimen; the time period with chemotherapy but free of toxicity (CT); the time period without symptoms of disease progression and without toxicity (TWiST); and the period following relapse until death (REL). This gave the following derived Q-TWiST:

$$\text{Q-TWiST} = U_{TOX}\,\text{TOX} + U_{CT}\,\text{CT} + \text{TWiST} + U_{REL}\,\text{REL}$$

where U_{TOX}, U_{CT}, U_{REL} are utility weights to reflect the value of time relative to TWiST in the health states TOX, CT and REL, respectively, on a scale from 0 (as bad as death) to 1 (as good as possible). TOX, CT, TWiST, REL and Q-TWiST are expressed in days.

Table 1
Distribution of the number of days with different toxic adverse events obtained from a sample of 32 independent experts

Type of toxicity	WHO grade	TOX, days mean/median (25th–75th percentiles)
Nausea, vomiting	1–2	2.7/3 (2–3)
	3–4	4.9/5 (3–6.2)
Alopecia	2–4	82.4/90 (40–90)
Infection	1–2	5.2/5 (3–7)
	3–4	19.9/14.5 (8–15.5)
Cardiac toxicity	3–4	30.6/30 (29–30)
Neurologic toxicity	2	40.7/26.5 (10–41.25)
	3–4	86/60 (30–99)
Isolated fever	2–4	6.5/5 (3–8.5)
Hemorrhage	3–4	9.4/8 (5–10)
Pulmonary embolism	–	45.6/30 (21–52.5)
Thrombosis	–	27.1/17.7 (10–30)
Extravasation, venous irritation	–	12.2/10 (6.2–15)
Occlusion	–	18.6/15 (10–29)
Constipation, colitis, enteritis, diarrhea	–	9.1/7 (5–12.5)
Itching, allergy, toxidermy	–	6.3/6 (3–8)
Muscle, bone or back pain	–	8.3/7 (5–10)
Therapeutic consequences of the toxicity		
	1–3 units	3.16/2 (2–5)
RBC transfusion	4–6 units	8.6/6 (4–8)
	>6 units	16.8/12 (10–30)
	1–2 units	2.4/1 (1–3)
Platelets transfusion	3–6 units	7.6/4 (3–10)
	>6 units	16.2/5 (6–28)

TOX = estimated time with toxicity expressed in days; RBC = red blood cells.

2.3. Statistical analysis

The average lengths of TOX, CT, TWiST and REL were estimated from the trial data, with comparison of treatments in terms of average Q-TWiST [13].

For each patient, the type and WHO grade of all adverse events experienced for each monthly chemotherapy course were recorded, while detailed data regarding duration of these events were not. They were estimated each by the mean of guesses independently given by 32 hematologists of the French Society of Hematology (Table 1). As previously stated by Gelber [11,17], we computed TOX by summing all of these estimated time lengths of toxic adverse events up to a maximum of the total chemotherapy duration, as shown below. Sensitivity analyses were carried out using different estimates of the toxic duration (either the 25th, the 50th, or the 75th percentiles of this distribution.)

Total duration of chemotherapy was individually computed from the number of actual days receiving chemotherapy, time to relapse was estimated by the time to occurrence of stage B or C after randomization.

Restricted mean time spent in each health state was then estimated separately for each treatment group by the difference between the areas under the non parametric Kaplan–Meier [27] estimates, with data truncation at median survival time (73 months). Non parametric bootstrapping [28], a numerical re-sampling method, was then carried out to obtain reliable estimates of the standard errors of the restricted mean durations of TOX, CT, TWiST and REL [13].

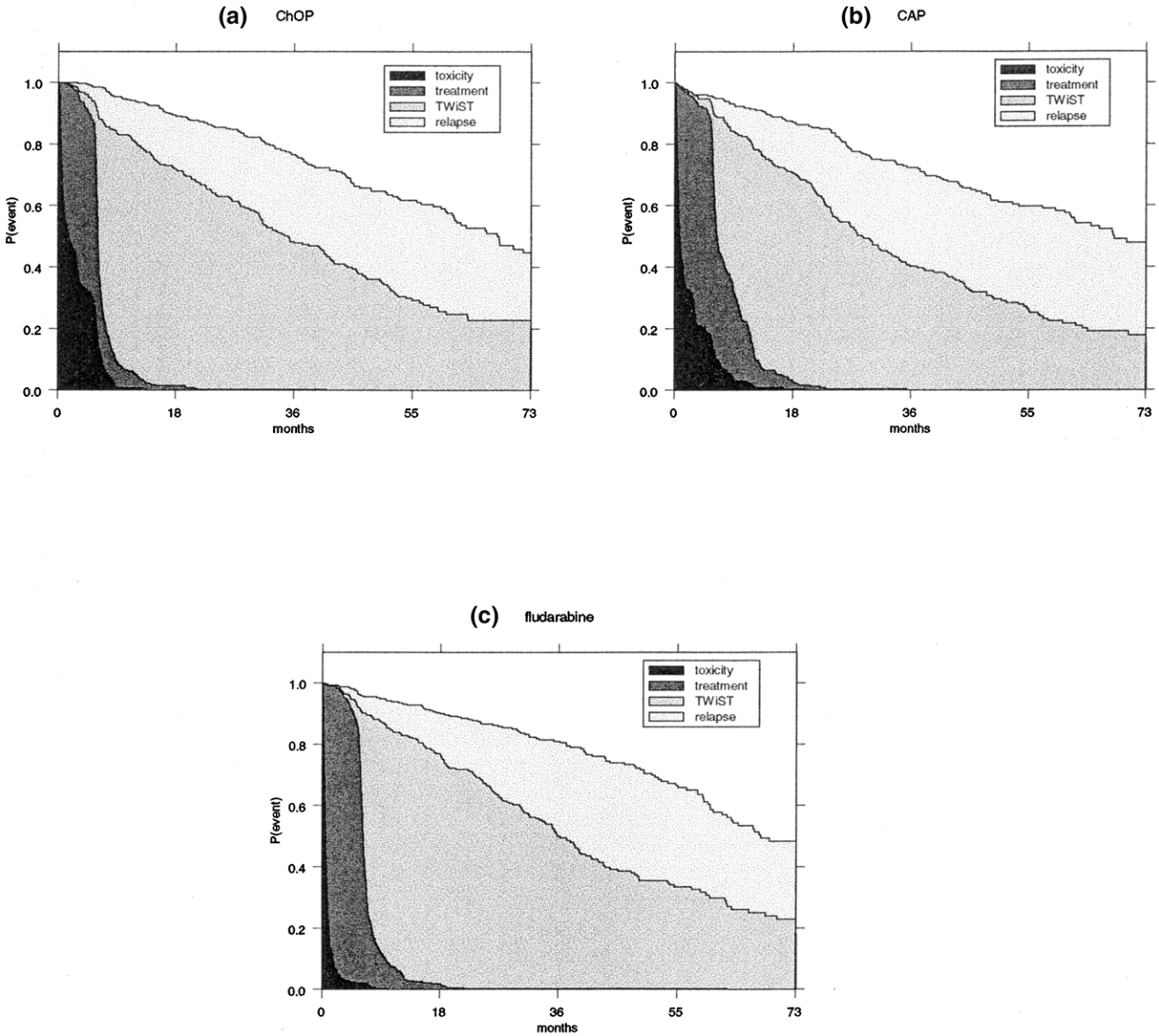

Fig. 1. Partitioned survival plots for (a) ChOP, (b) CAP and (c) fludarabine. Survival curves are plotted for OS, PFS (Progression Free Survival), treatment and toxicity. Areas between the curves represent the mean times spent in the health states.

The quality-of-life-adjusted model was constructed by introducing utility weights, as previously described [18]. Because no patient-level data were available to estimate the utility weights, results are presented as threshold analyses over all possible values for the three utility weights, using 3D figures [13]. Finally, the Q-TWiST gain function was used to illustrate the evolution of the treatment comparison over the time interval [15].

Statistical analysis was performed on SAS (SAS Inc., Cary, NC) and S-Plus software packages.

3. Results

A total of 880 patients were enrolled in this study. Median survival was 73 months (81 months in stage B- and 60 months in stage C-CLL).

Fig. 1 displays the partitioning of survival time into the four states according to treatment group, illustrating the time spent in each over the 73 months after randomization. The mean amounts of time spent in each health state are reported in Table 2 with bootstrap standard errors.

When considering the whole cohort, the fludarabine group gained a mean of 45 days of toxicity-free survival as compared to CAP, and 61 days as compared to ChOP. Differences in average TWiST were slight, 27.05 months with CAP, 31.5 months with ChOP and 32.95 months with fludarabine (Table 2).

Table 3 illustrates sensitivity to the estimation of toxicity durations (either mean, quartiles or median of the distribution from 32 experts) for the whole sample and according to base-line stage. Whatever the estimate, the ordering of the TOX durations of the three randomized groups was not modified, and differences between the three groups almost constant, so that further analyses only dealt with the mean individual estimates.

Separate results for stage B and C patients are shown in Table 2. Overall, as well as in either stage, the mean benefit of ChOP or fludarabine over CAP in terms of TWiST was about 6 months while the benefit of fludarabine over ChOP was much lower (from 0.7 to 1.5 months).

The mean Q-TWiST was computed and compared between the randomized groups for all possible combinations of utility weight values using a threshold utility analysis. For clarity, only the comparison between ChOP and fludarabine is illustrated in Fig. 2. The comparisons between CAP and ChOP, or CAP and fludarabine, showed that, whatever the utility coefficients, the mean Q-TWiST is always greater with ChOP or fludarabine, respectively. When comparing fludarabine to ChOP, fludarabine is always better in terms of average Q-TWiST, whatever the choice of the utility weights, except in the case of a high utility weight given to the toxicity and a low weight given to the treatment. Moreover, for a narrow range of utility coefficients (see Fig. 2), the average Q-TWiST was significantly increased in the fludarabine arm.

Figure 2b and c display the same results for stage B patients (70% of the whole sample) and stage C patients (30% of the sample), respectively. The results between CAP and the two others arms did not differ from those of the whole cohort. When comparing ChOP to fludarabine, the latter was most of the time [in about 90% of the combinations in stage B (Fig. 2b), and 75% in stage C (Fig. 2c)] superior to ChOP but the difference was never found to be statistically significant.

Fig. 3 illustrates the comparison of mean Q-TWiST between ChOP and fludarabine, as it unfolds over the follow-up. The solid line illustrates the mean overall survival comparison, while the region between the dotted lines illustrates the range of difference in mean Q-TWiST as the utility

Table 2
Components of the mean Q-TWiST within the first 73 months after randomization for the whole sample and according to Binet's stage (B or C)[a]

	ChOP mean (SE) in days	CAP mean (SE) in days	Fludarabine mean (SE) in days
Whole cohort			
TOX	88 (5.2)	71 (6.7)	26 (2.7)
CT	117 (7.2)	178 (10.0)	190 (7.1)
TWiST	962 (44.4)	825 (48.4)	1005 (47.3)
REL	507 (44.3)	550 (44.5)	506 (45.8)
Stage B-CLL			
TOX	91 (6.9)	72 (7.9)	24 (3.4)
CT	120 (9.1)	174 (11.2)	199 (9.7)
TWiST	1076 (55.5)	921 (56.7)	1093 (55.8)
REL	457 (52.0)	548 (53.6)	488 (52.2)
Stage C-CLL			
TOX	81 (8.2)	67 (14.2)	28 (3.9)
CT	111 (10.9)	187 (21.6)	170 (8.3)
TWiST	732 (71.7)	554 (88.7)	777 (83.1)
REL	604 (78.4)	565 (92.9)	557 (87.0)

[a]Each column gives the estimated mean duration (with bootstrap standard error) over the follow-up (expressed in days).

TOX = time with toxicity; CT = time with treatment for CLL; TWiST = time without symptoms or toxicity; REL = time with relapse.

Table 3
Sensitivity to the computation of toxicity durations (either mean, quartiles or median of the distributions from 32 experts) for overall sample of the CLL trial and according to Binet's stage[a]

Toxicity	ChOP	CAP	Fludarabine
Whole cohort			
Mean	87.8 (5.2)	70.9 (6.7)	25.8 (2.7)
25th percentile	67.0 (4.9)	46.0 (5.0)	14.7 (1.8)
Median	86.8 (5.5)	71.5 (7.2)	23.6 (2.5)
75th percentile	92.2 (5.7)	76.8 (7.2)	31.8 (2.9)
Stage B-CLL			
Mean	91.4 (6.9)	72.5 (7.9)	24.3 (3.4)
25th percentile	69.8 (6.0)	47.6 (5.5)	14.1 (2.5)
Median	91.3 (6.8)	73.6 (8.2)	22.6 (3.2)
75th percentile	95.1 (7.3)	77.8 (8.3)	29.5 (3.4)
Stage C-CLL			
Mean	81.4 (8.2)	66.9 (14.2)	28.3 (3.9)
25th percentile	62.0 (7.8)	45.6 (10.5)	15.8 (2.1)
Median	78.8 (8.3)	66.1 (15.0)	25.5 (4.0)
75th percentile	87.3 (8.1)	74.4 (14.6)	36.1 (4.3)

[a]Each column gives the estimated restricted mean TOX duration (bootstrap standard error) over the follow-up (expressed in days).

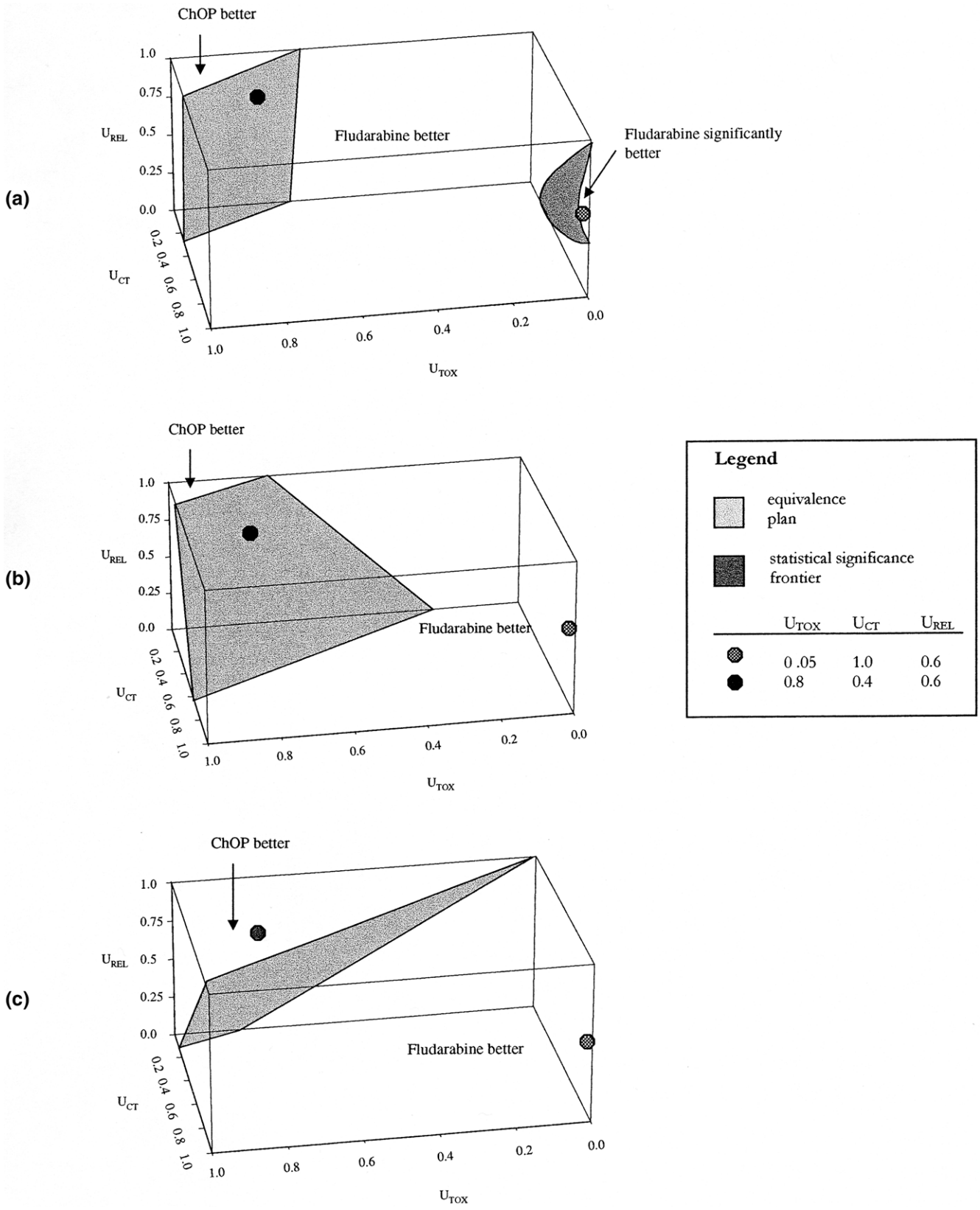

Fig. 2. Threshold utility analysis comparing ChOP to fludarabine for the whole population (a), stage B patients (b) and stage C patients (c). U_{TOX}, U_{CT}, U_{REL} are utility weights that reflect the value of time relative to TWiST in the health states TOX, CT and REL, respectively, on a scale from 0 (as bad as death) to 1 (as good as possible). In the whole cohort (a) three regions are defined, delineated by either equivalence plan in term of Q-TWiST between ChOP and fludarabine or statistical significance section of the fludarabine benefit. In both stages, either stage B (b) or stage C (c), only two regions are defined according to an equivalence plan (no combination of U_{TOX}, U_{CT} and U_{REL} allows to reach statistically significant benefit of fludarabine with ChOP). For example, when U_{TOX} = 0.05, U_{CT} = 1.0 and U_{REL} = 0.6, ignoring the stage (a), the resulting benefit of fludarabine in Q-TWiST is statistically significant. This benefit failed to reach statistical significance when considering separately either stage B (b) or stage C (c) patients. When U_{CT} = 0.4, U_{TOX} = 0.8 and U_{REL} = 0.6, there is an overall nonsignificant superiority of fludarabine over ChOP (a). This is also observed in stage B (b), but in stage C (c), ChOP becomes superior to fludarabine.

weight values vary between 0 and 1. The initial dip below 0 of the solid line is due to the negative impact of the fludarabine treatment on mean OS during the first 28 months. Despite the fact that the average OS benefit slightly increases over the time in the fludarabine arm, reaching approximately 50 days after 66 months, the gain in main Q-TWiST never shifts upward entirely above 0, even at 73 months. This demonstrates that it is not possible to find a time at which, whatever the values of the utility coefficients, the Q-TWiST will always be in favor of one arm of treatment. The two other comparisons (ChOP versus CAP, fludarabine versus CAP) showed a systematic advantage, in terms of OS, for the fludarabine or the ChOP arms versus CAP (data not shown).

4. Discussion

Treatment of advanced stage CLL remains a challenging problem for clinicians. Despite constant progress in the supportive care and the introduction of purine analogues, results in terms of survival are still discouraging, and CLL remains an incurable disease [4,6,29–31]. As a matter of fact, the original report of the CLL clinical trial did not indicate a significant improvement in OS for fludarabine as compared to ChOP or CAP for the treatment of stage B- and C-CLL patients [10]. In this context, evaluation of quality of life appears a central issue.

Our analysis evaluated the potential gains in terms of quality-adjusted survival using the Q-TWiST approach. We assumed that the quality of life spent with toxicity, chemotherapy, or disease progression was reduced as compared to that of time without disease progression and treatment toxicity. Usually, Q-TWiST is defined from three clinical states (TOX, TWiST and REL) assorted by utility weights (U_{TOX} and U_{REL}), the utility weight of TWiST being unity by definition. In CLL, as for most chronic lymphoproliferative disorders, a frequent status of the patients is to be in relapse, without any toxicity or treatment. This additional state was taken into account in two previous publications analyzing quality-adjusted time without symptoms or toxicity in hematological malignancies (myeloma and follicular non-Hodgkin's lymphoma [18,23]. Therefore, we distinguished the duration of which can vary from 1 to 3 months, allowing the patient to attribute different utility weights to each state (U_{TOX} and U_{CT}, respectively).

Some previous Q-TWiST analyses [18,32] have only considered severe toxicities in time with toxicity, while duration of TOX was fixed (e.g., treatment cycle duration). We performed our analysis by including lower grade of toxicity for symptoms such as nausea or vomiting, infection, isolated fever and neurological adverse events given they have a potential impact on quality of life. Because the use of a patient diary to capture side effects was not scheduled in the protocol, the various durations of toxic adverse events was estimated from a sample of 32 independent experts. The results show the robustness of this evaluation with a narrow variability (Table 1). The median TOX duration in the whole cohort represented between 1.2% and 2% of the follow-up, indicating that, even if effective measures of the individual TOX duration had been provided, the subsequent impact on the Q-TWiST would have been low (Table 3).

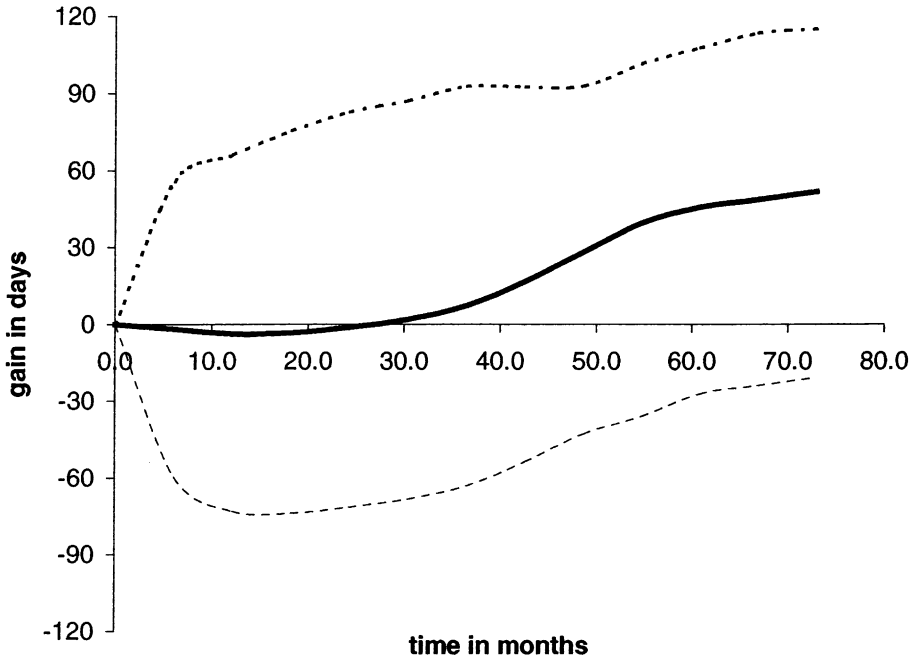

Fig. 3. Q-TWiST treatment comparison between ChOP and fludarabine over the course of follow-up. Solid line indicates the overall survival comparison. The region between the dotted lines indicates the range of the Q-TWiST treatment comparison as the three utility coefficients U_{TOX}, U_{CT} and U_{REL} vary from 0 to 1. Positive values indicate a benefit for fludarabine arm, while negative values indicate benefit for the ChOP arm.

However, a prospective evaluation of the effective duration should be incorporated in any future protocol, in order to reflect as best as possible the impact of the toxicity on the quality of life of individuals. Moreover, the impact of the toxicity on the quality of life should be evaluated not only by using the WHO classification and the duration of this toxicity, but also by collecting the patient's own perception regarding his/her quality of life.

The optimal first-line treatment in CLL is controversial. Results of the CLL 90 trial clearly demonstrated that fludarabine does not prolong survival as compared to ChOP or CAP in stage B- or C-CLL patients [9,10]. The present analysis confirmed the difference between CAP and the two other randomized groups in both mean TWiST and mean Q-TWiST whatever utility weights. However, differences in mean TWiST between ChOP and fludarabine were not significant ($P = 0.53$). On the contrary fludarabine increased significantly mean Q-TWiST as compared to ChOP for a set of utility weights, typically when $U_{TOX} < 0.1$, $U_{CT} > 0.9$ and U_{REL} between 0.3 and 1 (see Fig. 2a).

For example $U_{TOX} = 0.05$, $U_{CT} = 1$ and $U_{REL} = 0.6$ can be considered as a possible choice for a patient mostly fearing toxic effects. In this case fludarabine will significantly improve the mean Q-TWiST of 3.6 months ($P = 0.04$), which can be considered as a clinically relevant benefit. Nevertheless, for most combinations of utility weights, no benefit of either treatment could be pointed out.

When considering separately stage B- and stage C-CLL patients, no combination of utility weights gave a statistically significant benefit of either arm. However, this could be due to lower sample sizes, even if mean differences of Q-TWiST were, most of the time, slightly lower than when considering the whole cohort (see Table 2).

As any modeling, this approach considered toxic events jointly, simplifying the underlying potential differences between these events. Nevertheless, we tried to somewhat homogenize these events in mostly considering grade 2–4 toxic events, taking into account their differences in defining different time durations according to experts (either mean, median or quartiles) without consistently modifying the results. Moreover, sensitivity analysis allowed us to assess the influence of alopecia. To investigate whether or not the gain in mean quality-adjusted survival reached by fludarabine was influenced by the lack of alopecia in this group, which is the main difference in toxicity between fludarabine and anthracycline-containing regimens, we performed a threshold utility analysis of the Q-TWiST excluding this event in the computation of TOX. Even in this situation, the mean Q-TWiST was always in favor of fludarabine over ChOP. The mean Q-TWiST treatment comparison of ChOP versus fludarabine over the course of follow-up has also been performed in this situation. The average gain in Q-TWiST shifts upward entirely above 0 at 40 months of follow-up, indicating that 40 months after randomization, whatever the choice of utility weights, the mean Q-TWiST is always greater in the fludarabine arm.

The threshold utility analysis represents a tool to evaluate clinical trial data with respect to patient's own perception regarding his/her quality of life and it can assist the clinical decision regarding treatment choice. Measures of patient's own perceptions of quality of life should be developed and performed prospectively to be incorporated in the analysis. Based on the results of this analysis, we conclude that patients with stage B- and C-CLL may have a moderate benefit in terms of quality-adjusted time without symptoms or toxicity when treated by fludarabine versus ChOP or CAP. Further studies are needed to evaluate properly the effective time with toxicity, and this should be done in future prospective randomized trials evaluating therapeutic interventions in CLL.

Acknowledgments

We acknowledge all the participants of the CLL French Cooperative Group and Schering SA France for making this research possible, as well as the 32 members of the Société Française d'Hématologie who estimated the duration of the toxicities. We also thank Dominique Meffre and Cécile Dumont–Paris for their helpful assistance in the data management of the CLL trial. Supported by a grant of the Association pour la Recherche sur le Cancer (ARC) N°:6531.

References

[1] O'Brien S, del Giglio A, Keating M. Advances in the biology and treatment of B-cell chronic lymphocytic leukemia. Blood 1995;85:307–18.

[2] Rozman C, Montserrat E. Chronic lymphocytic leukemia [published erratum appears in N Engl J Med 1995 Nov 30;333(22):1515]. N Engl J Med 1995;333:1052–7.

[3] Dighiero G, Maloum K, Desablens B, Cazin B, Navarro M, Leblay R, Leporrier M, Jaubert J, Lepeu G, Dreyfus B, Binet JL, Travade P. Chlorambucil in indolent chronic lymphocytic leukemia. French Cooperative Group on Chronic Lymphocytic Leukemia. N Engl J Med 1998;338:1506–14.

[4] CLL Trialists' Collaborative Group. Chemotherapeutic options in chronic lymphocytic leukemia: a meta-analysis of the randomized trials. J Natl Cancer Inst 1999;91:861–8.

[5] Johnson S, Smith AG, Loffler H, Osby E, Juliusson G, Emmerich B, Wyld PJ, Hiddemann W. Multicentre prospective randomised trial of fludarabine versus cyclophosphamide, doxorubicin, and prednisone (CAP) for treatment of advanced-stage chronic lymphocytic leukaemia. The French Cooperative Group on CLL. Lancet 1996;347:1432–8.

[6] Keating MJ, O'Brien S, Lerner S, Koller C, Beran M, Robertson LE, Freireich EJ, Estey E, Kantarjian H. Long-term follow-up of patients with chronic lymphocytic leukemia (CLL) receiving fludarabine regimens as initial therapy. Blood 1998;92:1165–71.

[7] Clavio M, Miglino M, Spriano M, Pietrasanta D, Vallebella E, Celesti L, Canepa L, Pierri I, Cavaliere M, Ballerini F, Beltrami G, Rossi E, Vimercati R, Bruni R, Conglu M, Nati S, Damasio E, Santini G, Gobbi M. First line fludarabine treatment of symptomatic chronic lymphoproliferative diseases: clinical results and molecular analysis of minimal residual disease. Eur J Haematol 1998;61:197–203.

[8] Richardson DS, Johnson SA, Hopkins JA, Howe D, Phillips MJ. Absence of minimal residual disease detectable by FACS, Southern blot or PCR in patients with chronic lymphocytic leukaemia treated with fludarabine. Acta Oncol 1994;33:627–30.

[9] Leporrier M, Chevret S, Cazin B, Feugier P, Rapp MJ, Boudjerra N, Autrand B, Dreyfus B, Dessablens B, Dighiero G, Travade P,

Chastang C, Binet JL. Randomized comparison of fludarabine, CAP and ChOP in 695 previously untreated stage B and C chronic lymphocytic leukemia (CLL). Early stopping of the CAP accrual. Blood 1997;90:529A.

[10] Leporrier M, Chevret S, Cazin B, Boujderra N, Maloum K, Feugier P, Desablens B, Rapp MJ, Jaubert J, Autrand C, Divine M, Dreyfus B. Randomized comparison of fludarabine, CAP and ChOP in 938 previously untreated stage B and C- Chronic lymphocytic leukemia. submitted 2000.

[11] Gelber RD, Goldhirsch A. A new endpoint for the assessment of adjuvant therapy in postmenopausal women with operable breast cancer. J Clin Oncol 1986;4:1772–9.

[12] Goldhirsch A, Gelber RD, Simes RJ, Glasziou P, Coates AS. Costs and benefits of adjuvant therapy in breast cancer: a quality-adjusted survival analysis. J Clin Oncol 1989;7:36–44.

[13] Glasziou PP, Simes RJ, Gelber RD. Quality adjusted survival analysis. Stat Med 1990;9:1259–76.

[14] Gelber RD, Goldhirsch A, Castiglione M. The duration of a life of quality should become the focus of "quality-of-life" studies. J Clin Oncol 1989;7:542–4.

[15] Gelber RD, Cole BF, Gelber S, Goldhirsch A. Comparing treatments using quality-adjusted survival: the Q-TWiST method. Am Stat 1995; 49:161–9.

[16] Cole BF, Gelber RD, Goldhirsch A. A quality-adjusted survival meta-analysis of adjuvant chemotherapy for premenopausal breast cancer. International Breast Cancer Study Group. Stat Med 1995;14:1771–84.

[17] Gelber RD, Goldhirsch A, Cole BF, Wieand HS, Schroeder G, Krook JE. A quality-adjusted time without symptoms or toxicity (Q-TWiST) analysis of adjuvant radiation therapy and chemotherapy for resectable rectal cancer. J Natl Cancer Inst 1996;88:1039–45.

[18] Cole BF, Solal-Celigny P, Gelber RD, Lepage E, Gisselbrecht C, Reyes F, Sebban C, Sugano D, Tendler C, Goldhirsch A. Quality-of-life-adjusted survival analysis of interferon alfa-2b treatment for advanced follicular lymphoma: an aid to clinical decision making. J Clin Oncol 1998;16:2339–44.

[19] Feldstein ML. Quality-of-life-adjusted survival for comparing cancer treatments. A commentary on TWiST and Q-TWiST. Cancer 1991; 67:851–4.

[20] Lenderking WR, Gelber RD, Cotton DJ, Cole BF, Golhirsch A, Volberding PA, Testa MA. Evaluation of the quality of life associated with zidovudine treatment in asymptomatic human immunodeficiency virus infection. The AIDS Clinical Trials Group. N Engl J Med 1994;330:738–43.

[21] Schwartz CE, Cole BF, Gelber RD. Measuring patient-centered outcomes in neurologic disease. Extending the Q-TWiST method. Arch Neurol 1995;52:754–62.

[22] Schwartz CE, Coulthard-Morris L, Cole B, Vollmer T. The quality-of-life effects of interferon beta-1b in multiple sclerosis. An extended Q-TWiST analysis. Arch Neurol 1997;54:1475–80.

[23] Zee B, Cole B, Li T, Browman G, James K, Johnston D, Sugano D, Pater J. Quality-adjusted time without symptoms or toxicity analysis of interferon maintenance in multiple myeloma. J Clin Oncol 1998; 16:2834–9.

[24] Fermand JP, Ravaud P, Chevret S, Divine M, Leblond V, Belanger C, Macro M, Pertuiset E, Dreyfus F, Mariette X, Boccacio C, Brouet JC. High-dose therapy and autologous peripheral blood stem cell transplantation in multiple myeloma: up-front or rescue treatment? Results of a multicenter sequential randomized clinical trial. Blood 1998;92:3131–6.

[25] Parsons SK, Gelber S, Cole BF, Ravindranath Y, Ogden A, Yeager AM, Chang M, Shuster J, Weinstein HJ, Gelber RD. Quality-adjusted survival after treatment for acute myeloid leukemia in childhood: a Q-TWiST analysis of the pediatric oncology group study 8821. J Clin Oncol 1999;17:2144–52.

[26] Binet JL, Auquier A, Dighiero G, Chastang C, Piguet H, Goasguen J, Vaugier G, Potron G, Colona P, Oberling F, Thomas M, Tchernia G, Jacquillat C, Boivin P, Lesty C, Duault MT, Monconduit M, Belabbes S, Gremy F. A new prognostic classification of chronic lymphocytic leukemia derived from a multivariate survival analysis. Cancer 1981;48:198–206.

[27] Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958;54:457–81.

[28] Efron B, Tibshirani RJ. An introduction to the bootstrap. vol. 57. Monographs on statistics and applied probability. New York: Chapman and Hall, 1993.

[29] Oscier D. Chronic lymphocytic leukaemia. Br J Haematol 1999;105:1–3.

[30] Cheson BD, Bennett JM, Grever M, Kay N, Keating MJ, O'Brien S, Rai KR. National Cancer Institute-sponsored Working Group guidelines for chronic lymphocytic leukemia: revised guidelines for diagnosis and treatment. Blood 1996;87:4990–7.

[31] Sorensen JM, Vena DA, Fallavollita A, Chun HG, Cheson BD. Treatment of refractory chronic lymphocytic leukemia with fludarabine phosphate via the group C protocol mechanism of the National Cancer Institute: five-year follow-up report. J Clin Oncol 1997;15:458–65.

[32] Cole BF, Gelber RD, Kirkwood JM, Goldhirsch A, Barylak E, Borden E. Quality-of-life-adjusted survival analysis of interferon alfa-2b adjuvant treatment of high-risk resected cutaneous melanoma: an Eastern Cooperative Oncology Group study. J Clin Oncol 1996;14: 2666–73.