The UK NCRI study of chlorambucil, mitoxantrone and dexamethasone (CMD) versus fludarabine, mitoxantrone and dexamethasone (FMD) for untreated advanced stage follicular lymphoma: Molecular response strongly predicts prolonged overall survival.

Mark Bishton, Simon Rule, William Wilson, Deborah Turner, Russell Patmore, Laura Clifton-Hadley, Andrew McMillan, Richard Lush, Andrew Haynes

Summary

We present long term follow up of the UK CMD versus FMD for untreated advanced, symptomatic follicular lymphoma (FL). This trial was the first to prospectively assess molecular response and the impact on outcomes for 400 patients. The median progression free survival (PFS) and overall survival (OS) for CMD were 3.6 years and 14.6 years versus 3.0 years and 15.7 years for FMD respectively. Restricted Mean Survival Time (RMST) estimates suggested no difference in PFS or OS. For the whole cohort there was a highly significant difference in survival by POD24, with a median OS of 6.7 years compared to 15.7 years for all others (RMST p<0.001). Molecular remission was achieved in 25/46 patients (54.3%) in the CMD arm and 20/41 (48.8%) in the FMD arm (p=0.6). Molecular negativity resulted in median PFS of 5.6 years versus 2.3 years for molecularly positive (log-rank p<0.001) and median OS not reached versus 12.5 years (log-rank p<0.01). No cases of progression occurred in minimal residual disease (MRD) negative patients after six years of follow up. Although there was no difference in outcomes between arms, this is the first prospective study to report that MRD negativity results in significantly improved OS.

Introduction

Improvements in therapies for FL mean median survival from diagnosis is estimated to exceed 15 years (Junlen, et al 2015, Tan, et al 2013). Attempts to identify patients with excellent outcomes to immuno-chemotherapy have focussed on PET-CT and eradication of minimal residual disease (MRD) in peripheral blood and bone marrow. Both baseline total metabolic tumour volume (TMTV) and the achievement of complete metabolic response (CMR) following induction immuno-chemotherapy, are associated with improvements in PFS and OS (Meignan, et al 2016, Trotman, et al 2014), while, to date, MRD negativity has only corresponded to improvements in PFS (Galimberti, et al 2014). Moreover, progression of disease within 24 months (POD24) of starting initial treatment has been identified as a reliable surrogate endpoint for survival, regardless of immuno-chemotherapy induction, with a 5-year OS of 50% versus 90% for POD24 vs non-POD24 groups (Casulo, et al 2015, Marcus, et al 2017, Sarkozy, et al 2019). The choice of induction therapy is based on the outcomes of three randomized controlled trials (RCTs): (i) PRIMA which demonstrated improved PFS following two years of rituximab maintenance for those responding to induction rituximab-chemotherapy (Bachy, et al 2019) (ii) STiL which demonstrated a bendamustine backbone improved PFS compared to CHOP (Rummel, et al 2013), and (iii) GALLIUM which reported that the second generation monoclonal antibody obinutuzumab at induction and maintenance resulted in a modest improvement in PFS compared to rituximab (Marcus, et al 2017). Of these, GALLIUM has the shortest follow up, and is the only one to prospectively assess the impact of both PET-CT and MRD
response. GALLIUM reported 3-year PFS of 80.0% and high levels of MRD negativity for all regimens, which was associated with significantly longer PFS, although no association with OS has been reported. Those who achieved CMR had 2.5-year PFS of 87.4% compared to 54.9% those who did not achieve CMR (Trotman, et al 2018). To date, in the context of a long natural history, no large prospective study has reported improvements in OS following the achievement of MRD negativity.

We present data from the UK CMD versus FMD for untreated advanced, symptomatic FL. This trial compared two non-CHOP based combinations shown to have activity in FL (McLaughlin, et al 1996, Nickenig, et al 2007). The trial recruited patients at a time when rituximab was not approved for use in combination with chemotherapy in the UK, and was the last UK study to not incorporate either rituximab induction or maintenance in either arm. In contrast, this study was one of the first to prospectively assess molecular response rates and their correlation with long-term outcomes.

Methods

Patients and entry criteria

This study was a prospective, randomised, open-label, multi-centre phase III trial enrolling patients between 18 and 70 years of age with untreated advanced stage III-IV follicular lymphoma with physician-determined indication to treat on the basis of B-symptoms, bone marrow failure, bulky or progressive disease or compression syndromes. CT imaging was used for staging and response assessment. Performance status was assessed using the WHO scale. Exclusion criteria included left ejection fraction <45% by echocardiogram, serious concomitant disease affecting lifespan, impairment of liver or renal function (alkaline phosphatase/bilirubin or creatinine more than 2.5 times upper limit of normal) unrelated to lymphoma, previous malignancy except cervical carcinoma in situ or squamous cell skin cancer, known HIV infection, central nervous system involvement by lymphoma, pregnant or lactating women and a history of, or currently, active haemolytic anaemia with a positive direct anti-globulin test.

Randomisation and treatment protocol

After informed consent was obtained, patients were randomised centrally and stratified prospectively using the International Prognostic Index Score. The CMD combination comprised chlorambucil 10mg orally on days 1-10, mitoxantrone 12mg/m² administered intravenously on day 1 and dexamethasone 20mg administered daily orally on days 1-5. The FMD combination consisted of fludarabine 25mg/m² administered intravenously daily on days 1-3 with mitoxantrone and dexamethasone administered as per the CMD combination. Treatment cycles were to be repeated every 28 days for a maximum of 8 cycles. No interim blood counts were mandated. Secondary G-CSF prophylaxis was permitted in patients with delayed haematological recovery or who were admitted for neutropenic sepsis. Treatment cessation at six cycles was permitted in those patients receiving FMD experiencing increasing delays in recovery of their blood counts. All patients receiving fludarabine were given irradiated blood products until at least six months after completing therapy to reduce the risk of transfusion-related graft-versus-host disease. All patients were given cotrimoxazole or pentamidine nebulisers until at least six months after completion of therapy. Anti-
emetics and other prophylactic antibiotics were allowed as per local protocols. Patients with bulky disease were given adequate hydration prior to their first dose of fludarabine to reduce the risk of tumour lysis syndrome.

Evaluation and Response Criteria

In both arms of the study, response assessment by CT was performed after four cycles of therapy. If there was no response or disease progression patients were withdrawn from the study and treated at the local investigator’s discretion. All patients showing a response were to continue to a maximum of eight cycles after which the disease response was formally re-assessed by CT. Response was defined according to the International Working Group criteria (Cheson, et al 1999). Patients who provided end of treatment responses despite early withdrawal were also included in response analyses.

The primary end point for the trial was PFS. Secondary end points were OS, TTF, complete and partial remission rates, molecular response rates and the impact of molecular responses on both progression free and overall survival. TTF was defined as the interval between the date of randomisation and whichever occurred first of treatment discontinuation for any reason, including toxicity or progression, relapse or death due to lymphoma. Patients stopping treatment due to progression, lymphoma-related death or toxicity were classified as failures. PFS was defined as the time from randomisation to whichever occurred first of first relapse, progression or death. Patients still alive and progression free or lost to follow up were censored at the last date on which they were known to be alive. Duration of OS was defined as the time from randomisation to death from any cause. Patients still alive or lost to follow up were censored at the last date on which they were known to be alive. Patients were categorised as either POD24 or non-POD24 based on whether or not they progressed within 24 months of randomisation. Patients who died without progression or were lost to follow up within 24 months were excluded from POD24 analyses. OS from a risk-defining event was calculated as survival time from progression of disease for patients with POD24 or from 2 years post-randomisation for non-POD24 patients.

The frequency and severity of adverse events was recorded according to WHO criteria. A serious adverse event (SAE) was defined as any undesirable experience occurring to a patient, whether or not related to the investigational drugs and which resulted in death, immediate risk of death at the time the observation was made, hospitalisation or prolonged hospital stay, persistent or significant disability or incapacity or a congenital anomaly or birth defect. All SAEs were reported within 24 hours of observation to the central trial’s office.

Molecular Assessment

Diagnostic peripheral blood (PB) and bone marrow (BM) samples were to be sent to a central laboratory and screened to detect a t(14;18) translocation by nested PCR (expected sensitivity 1:10^5) and, if negative, a single round multiplex PCR for mcr/IgH JH and for other breakpoints within the bcl-2 locus as reported by the BIOMED-II consortium (Langerak, et al 2012, van Dongen, et al 2003). The expected sensitivity of these single round PCR analyses was 1:10^4. Patients positive at
randomisation were re-analysed at the end of treatment using the same primer sets in both marrow and peripheral blood if provided. A molecular response was defined as no evidence of Bcl2 gene rearrangement on testing bone marrow/peripheral blood samples by PCR in patients who were PCR positive at presentation.

Sample Size and Statistical Analysis

The 2-year PFS for the CMD regimen was expected to be 58%. Using a two-sided log-rank test of survival curves with a 5% significance level and at least 90% power, 500 patients were required to detect a 15% improvement in PFS for the FMD arm. It was anticipated that 250 patients would be informative for the molecular study. Prior studies of FMD suggested molecular remission rates in the order of 50%, whilst it was anticipated there would be few molecular remissions following CMD. On this basis, this number was therefore considered sufficient to assess the impact of molecular remission on patient outcome following therapy. OS, PFS and TTF curves were constructed using Kaplan Meier methods, with differences calculated using Cox proportional hazards models and evaluated using log rank tests. Analyses of PFS and OS were on an intention to treat basis while TTF, response and toxicity were only considered in patients known to have started treatment. In cases where the proportional hazards assumption appeared to be violated, the difference in RMST has been reported instead of the hazard ratio (HR). The time point used for calculating the difference in RMST has been taken as the largest observed time-to-event of either treatment arm.

Study conduct

This study was conducted according to the modified Declaration of Helsinki. The study was approved by the Northern and Yorkshire REC (MREC/99/3/63) and the MHRA (EudraCT 2005-003931-40) and the trial was managed by the Cancer Research United Kingdom and University College London Cancer Trials Centre. Informed consent was obtained from all patients before randomisation.

An Independent Data Monitoring Committee (IDMC) was appointed to assess data for the primary end point and toxicity at six-month intervals. If at any time a statistically significant difference was found (p=0.001) between the two arms, the IDMC was empowered to stop the study.

Results

Patient characteristics

Between May 2000 and April 2006 400 patients were enrolled into the trial. The IDMC approved this reduced recruitment total in March 2005 because of potential significant difference in PFS between the two arms. 11 CMD and nine FMD patients were deemed ineligible and were omitted from all analyses (figure 1). The median age of evaluable patients was 55 years (28-70). One hundred and thirty (34%) were 60 or more years of age. Central histological review confirmed that all had follicular lymphoma of grades 1 to 3A. Patient characteristics are summarised in table 1.
Patient Outcomes

325 patients (167 CMD; 158 FMD) were able to provide at least one response assessment following treatment. 12 patients (5 CMD, 7 FMD) progressed or died before any response assessment and were considered non-responders along with 1 FMD patient who was withdrawn early and only achieved SD. Two patients (1 CMD, 1 FMD) were withdrawn early but still achieved responses of CR and PR respectively. Best response over the full course of treatment gave an ORR of 94.2% for CMD versus 90.9% for FMD with CR rates of 41.9% and 40% respectively, with no significant difference between the two arms (table 2). Nine patients (CMD n=4, FMD n=5) had responses that were less than a PR after four cycles but continued to receive treatment and went on to have a second response assessment with all but two achieving at least a PR. In the CMD arm 158 patients (83.6%) had six or more cycles of therapy compared to 141 patients (73.8%) in the FMD arm. Median intervals between cycles were 28 days in both arms.

With a median follow-up period of 11.4 years the median PFS and OS for CMD were 3.6 years and 14.6 years versus 3.0 years and 15.7 years for FMD respectively. There was evidence to suggest the proportional hazards assumption was violated for both PFS (p<0.01) and OS (p=0.02), as following an apparent initial benefit for both PFS and OS for CMD, this trend was reversed after approximately eight years of follow up. RMST estimates suggested those receiving CMD had a non-significant +0.27 year (-0.98 to 1.51) difference in PFS over 15.2 years (p=0.7) and a non-significant -0.03 year (-1.29 to 1.23) difference in OS over 15.7 years compared to FMD (p>0.9) (figures 2A and 2B). Patients receiving CMD had a trend towards superior TTF with a median of 2.9 years compared to 1.6 years in the FMD arm. TTF curves appeared proportional with an estimated 22% improvement in TTF for patients receiving CMD compared to FMD (HR = 0.78 [0.60-1.01], log-rank p = 0.06) (figure 2C). Forest plots confirmed the hazard ratio for both PFS and OS changing from being in favour of CMD to FMD over time (figure 1S A and B). Multiple factors such as toxicity, MRD rates and death without relapse were considered, but there was no single reason explaining why the survival curves cross.

Seven patients died without progression within the first two years post-randomisation and a further 14 had less than two years of follow up and so were excluded from POD24 analyses. For the remaining 359 patients there was a highly significant difference in survival by POD24. Patients with POD24 had a median OS from a risk-defining event of 3.9 years compared to 13.7 years for those without POD24 (RMST difference -4.15 years [-5.32 to -2.99] over 13.7 years, p<0.001) (figure 3A). By treatment arm, RMST analysis showed that for CMD the POD24 patients had a -3.10 year (-4.73 to -1.46) difference in survival over 13.2 years compared to those without POD24 (p<0.001), while for FMD patients there was a -5.17 (-6.76 to -3.58) year difference over 13.7 years (p<0.001). In this analysis there was strong evidence that the proportional hazards assumption was not met (p<0.001). POD24 was associated with worse outcomes regardless of treatment arm and RMST analysis for interaction was non-significant, although there was a trend for those patients with POD24 FMD to have the poorest outcomes (p=0.09) (figure 3B). There was no difference in OS from the point of progression between the groups as a whole, with a median OS of 9.5 years for CMD and 8.7 years for FMD (log-rank p=0.7; data not shown).
Only four patients were reported to have had disease transformation at first progression (CMD n=1 and FMD n=3). Later transformation data was not captured.

Molecular Studies

A total of 92 patients (47 CMD, 45 FMD) had samples allowing molecular response assessment, however five were ineligible (CMD n=1; FMD n=4) and excluded from analyses. The baseline clinical characteristics of the patients in each arm remained well balanced (data not shown). Molecular remission was achieved in 25/46 patients (54.3%) in the CMD arm and 20/41 (48.8%) in the FMD arm (p=0.6). The achievement of molecular negativity was associated with median PFS of 5.6 years versus 2.3 years for those who remained molecularly positive (log-rank p<0.001) and median OS not yet reached versus 12.5 years (log-rank p<0.01) (figure 4A and 4B). For those patients who achieved MRD negativity, no cases of disease progression were reported after six years of follow up. There was no evidence of an interaction between molecular response and treatment when considering PFS as the outcome measure (p=0.6) (figure 4C). At a median follow up of over 12 years, only one of the twenty patients who achieved molecular remission in the FMD arm has died (at 15.7 years). Survival outcomes for FMD patients who remained MRD positive were notably worse than those who achieved MRD negativity, while CMD patients had similar survival times regardless of MRD status. However, numbers were small and a test for interaction in OS did not meet significance (p=0.09) (figure 4D). MRD positivity was strongly associated with early disease progression; only six (CMD n=2, FMD n=4) out of 45 MRD negative patients progressed within 24 months of randomisation while almost half (19/42) of the MRD positive patients had POD24 (p=0.001). Interim analysis on 70 patients comparing the nested PCR and single-round multiplex PCR techniques on bone marrow samples did not show any significant difference in PFS curves, suggesting single round PCR analysis to be adequate in predicting outcome. Analysis comparing bone marrow with peripheral blood MRD rates and clinical outcomes was not performed.

Adverse Events

Haematological toxicities were minor with very little grade 3-4 haematological adverse events in either arm. Grade 3-4 non-haematological adverse events were rare with the exception of infection, which occurred in 14.1% (CMD) and 11.3% (FMD) of patients (table 3). A total of 71 patients treated with CMD and 66 patients with FMD have died. Of these 39 (54.9%) and 32 (48.5%) were considered related to lymphoma, with 13 deaths (CMD n=5; FMD n=8) considered treatment related. Notably, after eight or more years post randomisation, the proportional number of deaths due to lymphoma has markedly decreased (CMD 6/18; FMD 1/10). Rates of secondary MDS/AML appeared low with four cases reported for CMD and three for FMD.

Discussion

The CMD versus FMD study has only previously been preliminary reported in abstract form and at the time had a median follow-up of 31 months and suggested CMD was superior with regards to PFS, with a trend to improved OS despite a modest increase in haematological toxicity. The median PFS of 3.0-3.6 years in this final report remains comparable to that seen with rituximab/obinutuzumab-CVP (Marcus, et al. 2017, Marcus, et al. 2008) although a median OS of
14.6-15.7 years corroborates recent epidemiological data suggesting the natural history of FL has been extended by improvements in immuno-chemotherapy, and, likely, supportive care. The extended follow up shows the curves to have reversed after approximately eight years, primarily due to fewer late progression events in the FMD cohort. RMST estimates suggested no significant difference in either PFS or OS. Currently the only validated surrogate biomarkers for survival in FL difference are POD24 following first line immuno-chemotherapy (Casulo, et al 2015, Marcus, et al 2017, Sarkozy, et al 2019) and CMR post induction on the basis of a recent meta-analysis of three large clinical trials, albeit with a modest absolute benefit (Trotman, et al 2018, Trotman, et al 2014). The FLIPI as a prognostic marker for newly diagnosed FL was established half way through the recruitment period and the routinely collected data did not allow for its calculation for the majority of patients. Our study also predated the incorporation of PET-CT into response assessment in FL, however, we confirm POD24 to be a prognostic factor for OS in patients who have received chemotherapy without antibody induction. Moreover, although not significant, there is a trend for non-POD24 FMD patients to have the best outcomes, while POD24 FMD patients have the worst. Given the issue of stem cell toxicities reported with fludarabine (Gill, et al 2010, Morgan, et al 2004, Strati, et al 2013) the differences may occur due to patients with FMD POD24 being more difficult to salvage due to haematological toxicity with second line therapy as well as problems with stem cell mobilisation, both scenarios that are less of a problem if relapse occurs later.


Of more direct relevance to current practice is the long-term outcome data following the achievement of molecular negativity. Both CMD and FMD confirmed high MRD rates can be seen despite the lack of CD20 directed therapy (Montoto, et al 2008). Although only 87 of all eligible subjects had complete sets of molecular data, these groups remained well balanced in terms of their baseline characteristics and likelihood of MRD negativity. As might be expected MRD eradication more than doubled PFS, but notably no case of disease progression was reported in molecularly negative patients after six years of follow up. Moreover, this is the first prospective study to suggest that achieving MRD negativity results in significantly improved OS. This is likely predominantly due to prolonged disease control following initial therapy, but could also potentially be a surrogate marker for greater chemo-sensitivity and duration of response at the point of salvage therapy. One potential confounding factor is that, as the study ran from 2000-2006, those MRD positive patients who relapsed earlier may have been less likely to be salvaged with rituximab containing regimens, although data to confirm this was not available. Abstract data presented from the GALLIUM study demonstrated MRD negativity at end of induction resulted in longer PFS, regardless of the antibody used (Pott et al, 2016). The Italian FOLL05 study compared three rituximab containing induction regimens, R-CHOP, R-CVP and R-FM, all without maintenance rituximab, reporting 3-year PFS of 66% for MRD negative cases versus 41% for MRD positive at 12 months post induction (Galimberti, et al 2014). The 2-year PFS rate of 86% for MRD negative patients in our study is similar to both FOLL05,
and GALLIUM despite the lack of antibody at either induction or maintenance, suggesting that PFS outcomes are excellent regardless of how MRD negativity is achieved.

Registry data for FL suggests such excellent long term survival that it is likely that future trial design may begin to focus on identifying the small group of patients who progress early on treatment with significantly poorer outcomes, and focus on reducing toxicity for other patients. It is likely that CMR, MRD negativity and POD24 rates are likely to become standard endpoints for future clinical studies, in addition to the expanding use of cell free DNA (Delfau-Larue, et al 2018). As the first prospective study to show MRD negativity following induction chemotherapy for follicular lymphoma results in significantly improved OS, our data emphasises the value of extended follow up and the use of OS as a valid endpoint.

Acknowledgements

MB wrote the paper and analysed the data, SR ran the study and enrolled patients, WW analysed the data and helped write the paper, DT ran the study and enrolled patients, RP ran the study and enrolled patients, LCH ran the study, AM ran the study and enrolled patients, RL performed the molecular analysis, AH designed and ran the study, and enrolled patients. The trial was managed by the Cancer Research United Kingdom and University College London Cancer Trials Centre. Funding for running the study was received from Schering. The authors would like to thank Paul Smith, Wendi Qian and Pavlina Mesiri from University College London Cancer Trials Centre for their help in running the study and performing earlier analyses, as well as the patients and their families who took part in the study and the investigators and research staff at all participating centres.
Figures

Figure 1. Trial flow diagram

Figure 2 Patient Outcomes by Treatment Arm
A. Time to Treatment failure
B. Progression free Survival
C. Overall Survival

Figure 3. Outcome by POD24
A. Analysis of PFS for patients with POD24 by treatment arm
B. Analysis of OS for POD24 by treatment arm

Figure 4. Outcomes by Molecular Response
A. Progression free survival by molecular response
B. Overall Survival by molecular response
C. Progression free survival by molecular response and treatment arm
D. Overall survival by molecular response and treatment arm

Tables
Table 1. Patient Baseline Characteristics
Table 2. Best Overall Response
Table 3. Adverse Events

Supplementary Figure 1.
A. Forest plot of progression free survival over time
B. Forest plot of overall survival over time
References


of a randomized comparison by the German Low-Grade Lymphoma Study Group. Ann Oncol, 18, 136-142.


suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. 
*Leukemia*, **17**, 2257-2317.