Time-to-progression after front-line fludarabine, cyclophosphamide, and rituximab chemoimmunotherapy for chronic lymphocytic leukaemia: a retrospective, multicohort study

Carmen D Herling, Kevin R Coombes*, Axel Benner, Johannes Bloehdorn, Lynn L Barron, Zachary B Abrams, Tadeusz Majewski, Jolanta E Bondaruk, Jasmin Bahlo, Kirsten Fischer, Michael Hallek, Stephan Stilgenbauer, Bogdan A Czerniak, Christopher C Oakes, Alessandra Ferrajoli, Michael J Keating, Lynne V Abruzzo*

Lancet Oncol 2019; 20: 1576–86

Published Online September 30, 2019 https://doi.org/10.1016/S1470-2045(19)30503-0

See Comment page 1478

*Contributed equally

Department I of Internal Medicine, Center for Integrated Oncology, Aachen-Bonn-Cologne-Duessel dorf, Cologne, Germany (C D Herling MD, J Bahlo PhD, K Fischer MD, Prof M Hallek MD); **Department of Biomedical Informatics** (Prof K R Coombes PhD, Z B Abrams PhD), **Department of Internal Medicine** (C C Oakes PhD), **and Department of Pathology** (Prof L V Abruzzo MD), **The Ohio State University, Columbus, OH, USA; Division of Biostatistics, German Cancer Research Center, Heidelberg, Germany** (A Benner MSc); **Internal Medicine III, University Hospital Ulm, Ulm, Germany** (J Bloehdorn MD, Prof S Stilgenbauer MD); **and Department of Hematopathology** (L L Barron BS), **Department of Pathology and Laboratory Medicine** (T Majewski PhD, J E Bondaruk PhD, Prof B A Czerniak MD), **and Department of Leukemia** (Prof A Ferrajoli MD, Prof M J Keating MD), **The University of Texas MD Anderson Cancer Center, Houston, TX, USA**

Correspondence to: Prof Lynne V Abruzzo, Department of Pathology, The Ohio State University, Columbus, OH 43202, USA **lynne.abruzzo@osumc.edu**

Summary

Background Fludarabine, cyclophosphamide, and rituximab (FCR) has become a gold-standard chemoimmunotherapy regimen for patients with chronic lymphocytic leukaemia. However, the question remains of how to treat treatment-naive patients with *IGHV*-unmutated chronic lymphocytic leukaemia. We therefore aimed to develop and validate a gene expression signature to identify which of these patients are likely to achieve durable remissions with FCR chemoimmunotherapy.

Methods We did a retrospective cohort study in two cohorts of treatment-naive patients (aged ≥18 years) with chronic lymphocytic leukaemia. The discovery and training cohort consisted of peripheral blood samples collected from patients treated at the University of Texas MD Anderson Cancer Center (Houston, TX, USA), who fulfilled the diagnostic criteria of the International Workshop on Chronic Lymphocytic Leukemia, had received at least three cycles of FCR chemoimmunotherapy, and had been treated between Oct 10, 2000, and Oct 26, 2006 (ie, the MDACC cohort). We did transcriptional profiling on samples obtained from the MDACC cohort to identify genes associated with time to progression. We did univariate Cox proportional hazards analyses and used significant genes to cluster *IGHV*-unmutated samples into two groups (intermediate prognosis and unfavourable prognosis). After using cross-validation to assess robustness, we applied the Lasso method to standardise the gene expression values to find a minimum gene signature. We validated this signature in an external cohort of treatment-naive patients with *IGHV*-unmutated chronic lymphocytic leukaemia enrolled on the CLL8 trial of the German Chronic Lymphocytic Leukaemia Study Group who were treated between July 21, 2003, and April 4, 2006 (ie, the CLL8 cohort).

Findings The MDACC cohort consisted of 101 patients and the CLL8 cohort consisted of 109 patients. Using the MDACC cohort, we identified and developed a 17-gene expression signature that distinguished *IGHV*-unmutated patients who were likely to achieve a long-term remission following front-line FCR chemoimmunotherapy from those who might benefit from alternative front-line regimens (hazard ratio 3·83, 95% CI 1·94–7·59; p<0·0001). We validated this gene signature in the CLL8 cohort; patients with an unfavourable prognosis versus those with an intermediate prognosis had a cause-specific hazard ratio of 1·90 (95% CI 1·18–3·06; p=0·008). Median time to progression was 39 months (IQR 22–69) for those with an unfavourable prognosis compared with 59 months (28–84) for those with an intermediate prognosis.

Interpretation We have developed a robust, reproducible 17-gene signature that identifies a subset of treatment-naive patients with *IGHV*-unmutated chronic lymphocytic leukaemia who might substantially benefit from treatment with FCR chemoimmunotherapy. We recommend testing the value of this gene signature in a prospective study that compares FCR treatment with newer alternative therapies as part of a randomised clinical trial.

Funding Chronic Lymphocytic Leukaemia Global Research Foundation and the National Institutes of Health/National Cancer Institute.

Introduction

Chronic lymphocytic leukaemia has a variable clinical course. Over the past decade, two clinical trials, the FCR300 study in the USA and the CLL8 study in Germany, reported that chemoimmunotherapy with fludarabine, cyclophosphamide, and rituximab (FCR) is a highly efficacious front-line therapy for chronic lymphocytic leukaemia.[1,2] FCR was the first regimen to improve progression-free survival and overall survival in patients with chronic lymphocytic leukaemia, and has become a gold-standard chemoimmunotherapy regimen in physically fit patients. Since then, follow-up studies

Research in context

Evidence before this study

We searched PubMed for all studies that identified the subset of treatment-naive patients with chronic lymphocytic leukaemia who might achieve durable remissions following chemoimmunotherapy with fludarabine, cyclophosphamide, and rituximab (FCR). We used the search terms "gene expression profiling", "CLL", "fludarabine", "FCR", and "prognosis". There were no date or language restrictions. Several studies show that a subset of relatively young, fit patients whose chronic lymphocytic leukaemia cells contain mutated *IGHV* genes and do not have high-risk cytogenetic abnormalities achieve durable remissions following FCR chemoimmunotherapy. However, we found no studies that sought to identify patients with unmutated *IGHV* genes who might also achieve durable remissions following FCR chemoimmunotherapy.

Added value of this study

Using data from two clinical trials that showed the efficacy of front-line FCR chemoimmunotherapy for the treatment of chronic lymphocytic leukaemia we developed a robust and reproducible 17-gene expression signature that distinguishes between treatment-naive patients with *IGHV*-unmutated chronic lymphocytic leukaemia who are likely to achieve long-term remissions following front-line FCR chemoimmunotherapy from those who might benefit from alternative regimens. Our results suggest that differences in the expression of genes involved in oxidative phosphorylation and purine metabolism account, at least in part, for the differences in response to front-line FCR chemoimmunotherapy in patients with *IGHV*-unmutated chronic lymphocytic leukaemia.

Implications of all the available evidence

Our study indicates that, using the 17-gene signature, it is possible to identify patients with *IGHV*-unmutated chronic lymphocytic leukaemia who are likely to achieve long-term remissions with FCR. Thus, for this subset of patients who might desire a short therapy course, or who cannot tolerate therapy with tyrosine-kinase inhibitors, FCR remains a viable option. We would recommend testing the value of the 17-gene signature in a prospective study that compares FCR treatment with alternative therapies, such as ibrutinib, as part of a randomised clinical trial.

have shown that a subset of patients achieve long-term durable remissions.[3–5] In general, these patients are quite young (under 65 years of age) and deemed relatively fit without comorbidities. Their chronic lymphocytic leukaemia cells usually have somatic mutations in *IGHV* and do not have high-risk cytogenetic abnormalities, such as del(11)(q22–23) and del(17)(p13.1). These findings suggest that FCR chemoimmunotherapy results in durable remissions in a carefully selected subset of patients.

We therefore aimed to develop and independently validate a gene expression signature to identify patients likely to achieve durable remissions with FCR chemoimmunotherapy. We hypothesised that inherent differences in the expression of protein-coding genes in the chronic lymphocytic leukaemia cells of treatment-naive patients determine the duration of response following FCR chemoimmunotherapy.

Methods

Study design and sample collection

In this retrospective cohort study, the discovery and training cohort consisted of peripheral blood samples collected from treatment-naive patients with chronic lymphocytic leukaemia at the University of Texas MD Anderson Cancer Center, Houston, TX, USA (ie, the MDACC cohort) and processed as described previously.[6–8] We included patients (aged ≥18 years) with chronic lymphocytic leukaemia that fulfilled the diagnostic criteria of the International Workshop on Chronic Lymphocytic Leukemia,[1] had received at least three cycles of FCR chemoimmunotherapy, and had been treated between Oct 10, 2000, and Oct 26, 2006. Patients were only included if data for clinical follow-up were available, and if they had provided peripheral blood samples within 6 months of start of therapy and had given written informed consent for gene expression profiling. We excluded patients who did not meet the inclusion criteria or had an RNA quality that was insufficient (RNA integrity number ≤7; appendix 1 p 1). Clinical and routine laboratory data were obtained from the medical records. The *IGHV* somatic mutation status and ZAP70 expression, measured by either flow cytometry or immunohistochemistry, were assessed on blood or bone marrow samples (appendix 1 p 4).[8–10] Common abnormalities associated with chronic lymphocytic leukaemia, ie, del(11)(q22.3) including *ATM*, del(13)(q14.3) including the *DLEU2/mir-15a/16-1* cluster, del(17)(p13.1) including *TP53*, and trisomy 12, were assessed by array-based single-nucleotide polymorphism genotyping[6,8] and grouped according to the Döhner hierarchy.[11]

See Online for appendix 1

The validation cohort consisted of treatment-naive patients with *IGHV*-unmutated chronic lymphocytic leukaemia enrolled on the CLL8 trial of the German Chronic Lymphocytic Leukaemia Study Group between July 21, 2003, and April 4, 2006 (ie, CLL8 cohort). These patients had given written informed consent for the use of their samples. Peripheral blood samples were collected at the University of Ulm (Ulm, Germany) just before study entry. Complete clinical data and prognostic markers for this cohort have been published previously.[1,3,12]

This study was approved by the institutional review boards of each hospital centre and done according to the principles expressed in the Declaration of Helsinki.

Gene expression profiling

We extracted total RNA from CD19-positive peripheral blood B cells enriched using immunomagnetic beads.[8] For the MDACC cohort, transcriptional profiling was done using the HumanHT-12 v4 Expression BeadChip Kit (Illumina, San Diego, CA, USA), according to the manufacturer's instructions. Data were background corrected, normalised, and log-transformed using the lumi R package (version 2.20.1).[13] Standard exploratory plots, hierarchical clustering, and principal component analyses were prepared to assess the intrinsic quality of individual arrays and exclude technological or processing artifacts. To account for multiple testing, we bounded the false discovery rate (FDR), which was estimated using a β-uniform mixture model.[14]

For the CLL8 cohort, transcriptional profiling using the GeneChip Human Exon 1.0 ST Array (Affymetrix, Santa Clara, CA, USA) and expression data processing were done as described previously.[15]

Statistical analysis

For both cohorts, response to treatment and progression were classified according to the National Cancer Institute working group criteria.[16] Time to progression was defined as the time from the start of therapy until disease progression; death was excluded in this definition, because in our cohort the cause of death was often not disease-related or was unknown. Using the R survival package (version 2.14-3), we fitted univariate per-gene Cox proportional hazards models to find genes that were statistically associated with time to progression as a continuous variable. Genes with a p<0·0085 (FDR of 33%) were used in hierarchical clustering with Pearson correlation and Ward's linkage to define subtypes based on gene expression profiling. The gene expression profiling and time-to-progression subtypes were verified by principal component analyses and investigated for interactions with known chronic lymphocytic leukaemia prognostic markers (Fisher's exact test and χ^2 test). The gene expression profiling and time-to-progression subtypes were included as variables in multivariable Cox proportional hazards models competing with classical predictors of clinical prognosis. We used the Akaike Information Criterion to add or subtract variables in a stepwise manner and optimise models with the strongest prognostic value.[17] We thoroughly cross-validated the clustering into subtypes and the resulting models. To identify samples with a high posterior probability of being reliably assigned to subtypes, we applied linear discriminant analysis using the five principal (ie, the first five) components as predictors. Using those reliable samples, we did preliminary feature selection based on gene-by-gene *t* tests (with p<0·0026, FDR of 15%) between the groups. We then applied the Lasso method to find a minimum-gene signature.[18]

After using the Lasso method to finalise a model on the MDACC cohort, details of the algorithm were sent to the biostatistician (AB) on the CLL8 trial. Although progression-free survival was the primary outcome in previous reports on the CLL8 cohort, these data were reanalysed to compute time to progression. Using samples from *IGHV*-unmutated FCR-treated patients from the CLL8 cohort, data from each of the genes in the final prognostic model were standardised to have a mean of zero and a variance of one. A continuous score was computed as the linear predictor of the standardised gene expression levels using the coefficients from the prognostic model as weights. This score was dichotomised into a binary predictor of low-risk (≤0) and high-risk (>0) categories. We evaluated the model's performance on the independent validation set using a proportional cause-specific hazards model of time to progression, using either the continuous score or the binary predictor. Cumulative incidence of progression was estimated using the Aalen-Johansen estimator. Prediction error curves were used to assess the predictive accuracy of the models,[19] in which prediction error was defined using Brier's score as a function of time.[20] Clinical models whose variables and coefficients were learned from the MDACC cohort were also validated on the CLL8 cohort.

Network construction, canonical pathway analyses, and functional annotations were done using Ingenuity Pathway Analysis (IPA), ToppGene, and DAVID 6.8. Separate gene files for each of the identified gene clusters were analysed using the default settings for each tool. Gene sets were evaluated by a modified Fisher's exact test with a Benjamini-Hochberg correction for multiple testing. To make it clear where each result came from, we report adjusted p values from DAVID and FDR q values from ToppGene.

We did all the statistical computations using the R statistical programming environment (version 3.4.0), including the survival package (version 2.41–3), glmnet (version 2.0–10), ClassComparison (version 3.1.4), and ClassDiscovery (version 3.3.5). The significance of time-to-event models was assessed using the p values from a score (log-rank) test based on Cox proportional hazards models.

For more on **Ingenuity Pathway Analysis** see www.ingenuity.com

For more on **ToppGene** see https://toppgene.cchmc.org/

For more on **DAVID 6.8** see http://david.abcc.ncifcrf.gov/

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, manuscript preparation, or in the decision to submit the manuscript for publication. The corresponding author had full access to all data and had final responsibility for the decision to submit for publication.

Results

The MDACC cohort consisted of 101 patients with chronic lymphocytic leukaemia (figure 1); survivors were followed up for a median of 146·5 months (IQR 137·7–162·3) after treatment (table 1). Established markers of chronic lymphocytic leukaemia prognosis, including *IGHV*

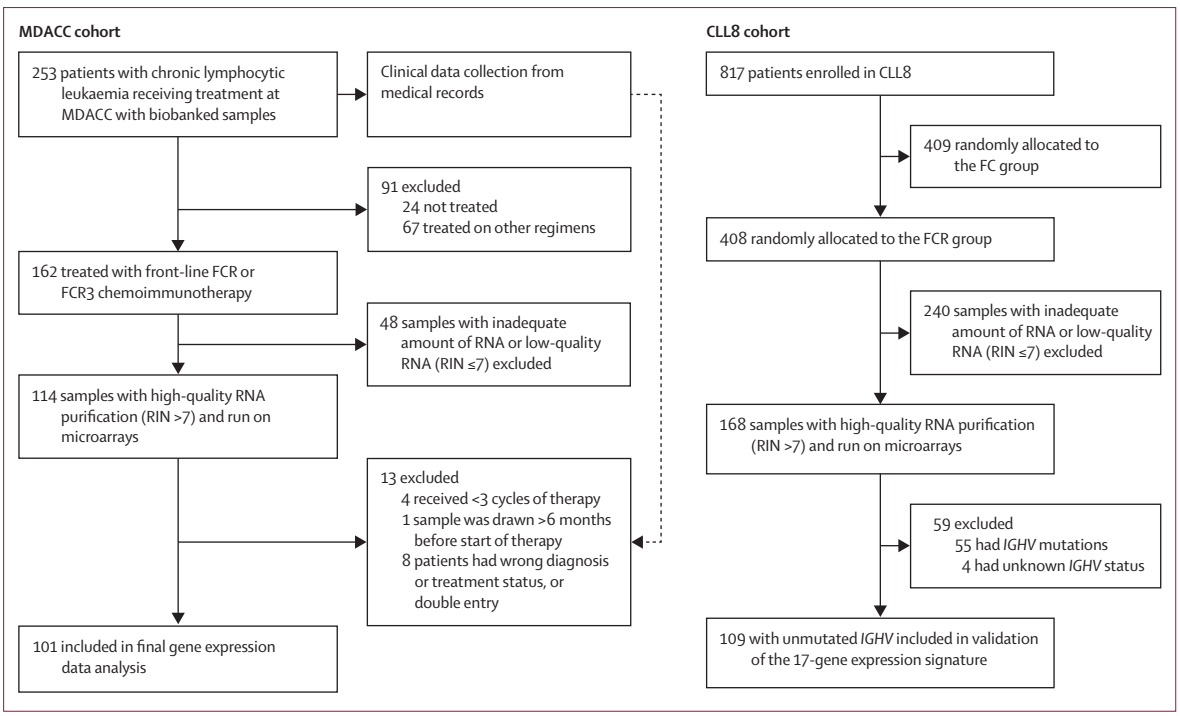

Figure 1: **CONSORT diagram for the MDACC discovery cohort and the CLL8 validation cohort**
CLL=chronic lymphocytic leukaemia. FC=fludarabine and cyclophosphamide. FCR=fludarabine, cyclophosphamide, and rituximab. FCR3=fludarabine, cyclophosphamide, and multiple-dose rituximab. MDACC=MD Anderson Cancer Center. RIN=RNA integrity number.

somatic mutation status, ZAP70 expression, and serum β-2-microglobulin concentration were represented at expected frequencies (table 1; appendix 1 pp 2, 3). All patients received at least three cycles of front-line therapy with either standard FCR (n=61)[2,22] or rituximab-intensified FCR (three instead of one rituximab dose per cycle; n=40).[23] 79 (78%) of 101 patients completed six cycles of chemoimmunotherapy and 98 (97%) responded to treatment. 84 (83%) had complete remissions, 14 (14%) had partial remissions, two (2%) had no response, and one (1%) had an unknown status. At the final follow-up, 49 (49%) of 101 patients had progressed, 42 (42%) had died, and 63 (62%) had either progressed or died. The median overall survival from the start of FCR therapy was not reached at the end of follow-up. The median time to progression was 92·4 months (IQR 44·1–not reached [NR]) and the median progression-free survival was 66·6 months (40·5–NR).

The CLL8 cohort consisted of 109 patients with available gene expression profiling data (figure 1). When compared with the *IGHV*-unmutated subset of the MDACC cohort, fewer patients in the CLL8 cohort were ZAP70 positive, but more patients had high CD38 expression and unfavourable cytogenetics (table 1; appendix 1 pp 2, 3). Median follow up was 73·4 months (IQR 68·2–82·9), median overall survival was 80·8 months (58·3–NR), median time to progression was 46·5 months (23·1–83·0), and median progression-free survival was 42·4 months (21·4–72·0). At the final follow-up, 77 (71%) of 109 patients had progressed, 46 (42%) had died, and 83 (76%) had either progressed or died.

Fitting univariate Cox proportional hazards models to gene expression profiling data for the MDACC cohort at an FDR of 0·33, we identified 1136 probes as significantly associated with time to progression (p≤0·0085; appendix 1 p 11 and appendix 2). Hierarchical clustering using these probes divided patients into three patient subsets, driven by the gene expression patterns in four subsets (figure 2A). Log-rank test and Cox proportional hazards analysis showed that the three patient subsets had significantly different time to progression (log-rank $p=5·9\times10^{-10}$; figure 2B). Compared with patients in the favourable subset, patients assigned to the intermediate prognosis subset (hazard ratio [HR] 4·64, 95% CI 1·06–20·18), and those assigned to the unfavourable prognosis subset (18·36, 4·37–77·18) had a shorter time to progression.

See Online for appendix 2

Pathway analyses with IPA, ToppGene, and DAVID using the 1136 genes yielded similar results (FDR of 0·05; appendix 1 p 12 and appendix 3). The purple cluster contains 424 genes that are highly expressed in unfavourable cases with shorter time to progression (figure 2). This cluster is significantly enriched for genes associated with metabolic pathways, including oxidative phosphorylation and ribonucleoside metabolism. ToppGene analysis showed that the most significant Gene Ontology (GO) categories are oxidoreductase

See Online for appendix 3

	MDACC cohort		CLL8 cohort (*IGHV*-unmutated samples; n=109)
	All samples (n=101)	*IGHV*-unmutated samples (n=66)	
Age at diagnosis, years	56 (49–62)	56 (49–63)	60 (53–65)
Sex			
Male	76 (75%)	52 (79%)	86 (79%)
Female	25 (25%)	14 (21%)	23 (21%)
Rai stage			
0–2	77 (76%)	52 (79%)	63 (69%)
≥3	24 (24%)	14 (21%)	28 (31%)
Not available*	0	0	18
IGHV status			
Mutated	35 (35%)	0	0
Unmutated	66 (65%)	66 (100%)	109 (100%)
β-2-microglobulin			
≤4 mg/L	68 (67%)	42 (64%)	85 (79%)
>4 mg/L	33 (33%)	24 (36%)	23 (21%)
Not available*	0	0	1
White blood cells			
$\leq 150 \times 10^9$ cells per L	83 (82%)	50 (76%)	83 (77%)
$>150 \times 10^9$ cells per L	18 (18%)	16 (24%)	25 (23%)
Not available*	0	0	1
ZAP70 expression			
Positive	49 (56%)	43 (74%)	31 (48%)
Negative	39 (44%)	15 (26%)	34 (52%)
Not available*	13	8	44
CD38 expression			
<30%	75 (74%)	45 (68%)	59 (58%)
≥30%	26 (26%)	21 (32%)	42 (42%)
Not available*	0	0	8
Cytogenetics			
del(17)(p)	1 (1%)	1 (2%)	13 (12%)
del(11)(q)	17 (17%)	16 (24%)	42 (39%)
Trisomy 12	17 (17%)	10 (15%)	3 (%)
FISH normal	31 (31%)	24 (36%)	22 (20%)
del(13)(q)	35 (35%)	15 (23%)	29 (27%)
Time-to-event parameter (months)			
Diagnosis to sample	21·0 (7·4–44·2)	ND	NA
Diagnosis to FCR treatment	21·0 (8·3–45·6)	ND	NA
Sample to FCR treatment	0·0 (0·0–0·1)	ND	NA
FCR treatment to final follow-up†	146·5 (137·7–162·3)	ND	73·4 (66·8–82·3)

Data are median (IQR), n (%), or n. Percentages might not sum to 100% due to rounding. MDACC=MD Anderson Cancer Center. CLL=chronic lymphocytic leukaemia. FISH=fluorescence in-situ hybridisation. FCR=fludarabine, cyclophosphamide, and rituximab. ND=analysis not done. NA=not applicable. *Not included in the calculation of the proportions. †Based on the method of Schemper and Smith[21] and the number of patients still alive at last follow-up (n=59 in MDACC cohort; and n=63 in CLL8 cohort).

Table 1: **Patient characteristics**

activity (GO:0016491; FDR $q=1{\cdot}01\times10^{-8}$), glycosyl compound, nucleoside, and ribonucleoside metabolic processes (GO:1901657, GO:0009116, GO:0009119; all with FDR $q=1{\cdot}22\times10^{-11}$); and the mitochondrion (GO:0005739; FDR $q=9{\cdot}79\times10^{-16}$). The most significant pathways are metabolic pathways (KEGG:132956; FDR $q=3{\cdot}04\times10^{-13}$) and oxidative phosphorylation (KEGG:82942; FDR $q=4{\cdot}31\times10^{-7}$). DAVID analysis yielded similar results (data not shown).

The cyan cluster contains 401 genes that are highly expressed in favourable or intermediate cases with longer time to progression. ToppGene analysis showed that the most significant GO categories are ATP binding (GO:0005524; FDR $q=2{\cdot}38\times10^{-3}$) and purine ribonucleoside triphosphate binding (GO:0035639; FDR $q=2{\cdot}38\times10^{-3}$). DAVID analysis showed that in the most significant annotation cluster (E=3·79), the most significant GO categories are nucleic acid binding (GO:0003676; $p=6{\cdot}20\times10^{-4}$), DNA-templated transcription (GO:0006351; $p=5{\cdot}28\times10^{-6}$), and the nucleus (GO:0005634; $p=2{\cdot}17\times10^{-8}$). Significant UniProt keywords ($p=4{\cdot}78\times10^{-7}$) and sequence features ($p=1{\cdot}47\times10^{-5}$) show enrichment for zinc-finger transcription factors. Expression of both the red (n=241) and green (n=70) groups of genes are variably expressed within the gene expression profiling and time-to-progression subtypes (figure 2), and neither group shows significant enrichment with GO categories (appendix 1 p 12 and appendix 3).

Multivariate analysis of clinical variables in the MDACC cohort (n=100, excluding the del(17)(p) sample) using the Akaike Information Criterion (appendix 1 p 13) selected age at diagnosis, *IGHV* mutation status, serum β-2-microglobulin, and cytogenetics as independent markers of time to progression (table 2). These clinical variables were included in additional analyses for model development or validation (appendix 1 pp 5–6). When we added the categories based on gene expression profiling of the patients (favourable, intermediate, or unfavourable prognosis) as a competing variable, age at diagnosis, serum β-2-microglobulin, cytogenetics, and the gene expression profiling subtypes were retained as significant predictors of time to progression (table 2). When we repeated this analysis on the *IGHV*-unmutated subset, the only difference was that *IGHV* status could no longer be included as a factor. Thus, in both the full dataset and the *IGHV*-unmutated subset, gene expression profiling subtype was a significant predictor of early versus late progression after FCR, independent of clinical variables. We also did exploratory univariate analyses (appendix 1 pp 14, 15).

We developed a 17-gene expression signature to distinguish patients with an unmutated *IGHV* status with an intermediate outcome versus those with an unfavourable outcome (appendix 1 pp 16–20). Clustering cannot, by itself, predict outcomes on new patient samples. Because the cluster with the best outcome was

strongly associated with known markers of good outcome (*IGHV*-mutated status, ZAP70 negative), we focused on developing a prognostic model with fewer genes to separate those with an unmutated *IGHV* status with an intermediate prognosis from those with an unfavourable prognosis. Cross-validation suggested that the border between the two groups was imprecise (appendix 1 pp 7–10). To increase our ability to distinguish them, we identified reliable (ie, easily classified) samples and built our model using these samples (n=54; 24 intermediate and 30 unfavourable). We identified 726 differentially expressed genes (FDR 15%). Of these 726 genes, 177 (24%) were in the set associated with time to progression used for clustering. After omitting duplicate probes, pseudogenes, and genes of uncertain function, 139 candidate predictors remained (appendix 1 p 18). Using the Lasso method, we selected a minimal set of genes separating the two prognostic groups; the final optimal model contained 17 genes (table 3). Using the predicted classes from the 17-gene model, we successfully separated patients with an *IGHV*-unmutated status into cases with intermediate and unfavourable risk of progression after front-line FCR therapy (figure 3A). The cause-specific hazard ratio was 3·83 (95% CI 1·94–7·59; p<0·0001). We validated this gene signature in the CLL8 cohort, in which patients classified as high-risk (ie, unfavourable prognosis) had a cause-specific hazard ratio of 1·90 (95% CI 1·18–3·06; p=0·008; median time to progression of 39 months (IQR 22–69) compared with low-risk (ie, intermediate prognosis) patients (median time to progression of 59 months; IQR 28–84). When we applied the prognostic model to the so-called unreliable samples that were not used for training to evaluate the generalisability of the model, none of the favourable outcome samples were assigned to the unfavourable group (as part of an informal sensitivity analysis using the MDACC cohort before trying to validate the results on the CLL8 cohort). Among *IGHV*-mutated samples, only four (11%) of 35 were assigned to the unfavourable group (appendix 1 pp 19, 20).

We validated the 17-gene expression signature using 109 patients with an *IGHV*-unmutated status from the CLL8 trial. The score is given as a linear combination of standardised gene expression levels using model coefficients, which were learned from the training set (table 3), as weights (standardised coefficients). Using a proportional hazards cause-specific Cox regression model, the estimated cause-specific hazard ratio for the continuous score was 1·41 (95% CI 1·13–1·76). For

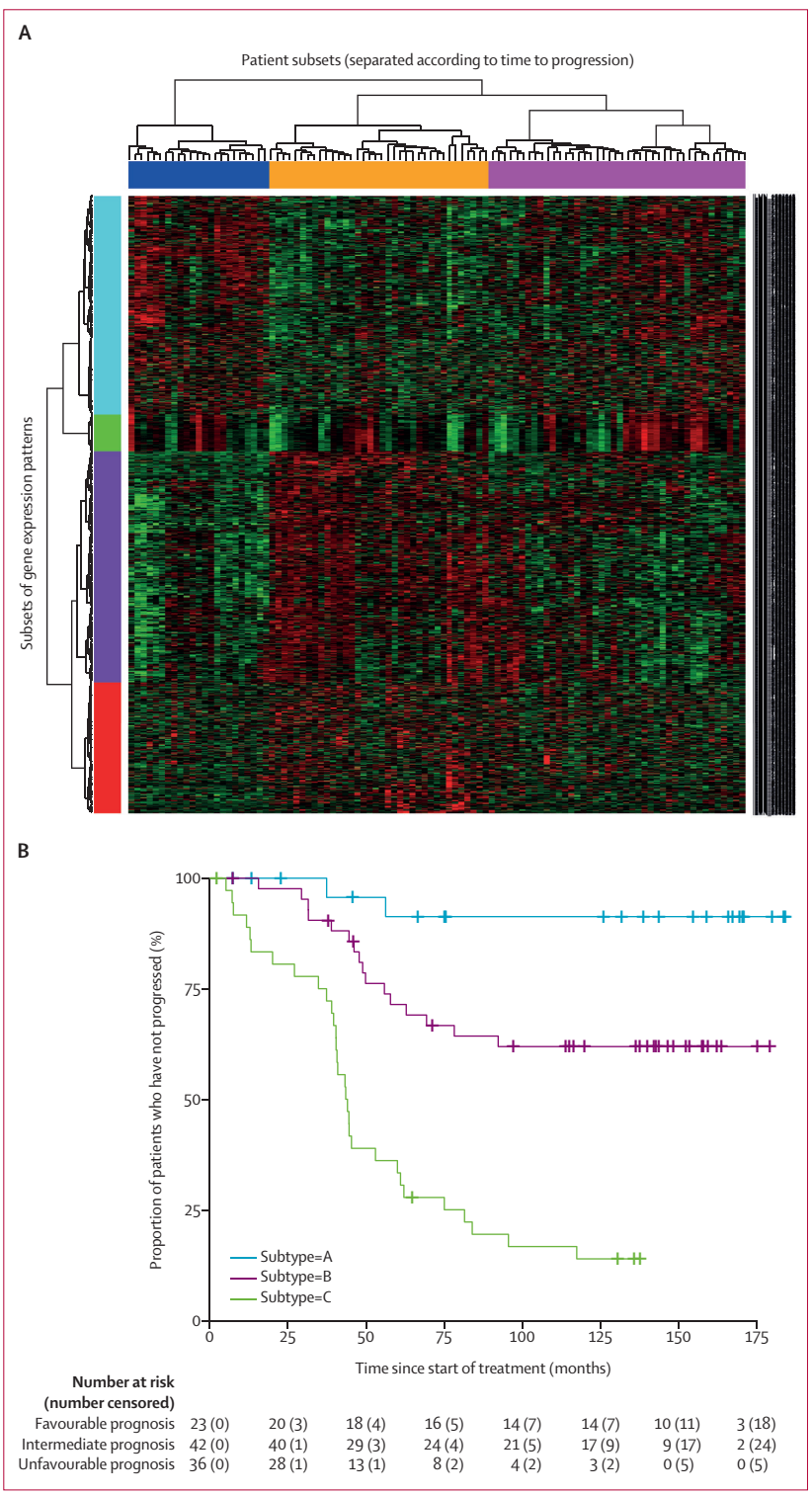

Figure 2: **Hierarchical gene clustering identifies gene sets that divided patients into prognostic groups**

(A) Heat map illustrating hierarchically clustered gene sets that separate three patient subgroups with different clinical outcomes (time to progression) after front-line fludarabine, cyclophosphamide, and rituximab therapy. 1136 genes associated with time to progression in univariate analysis were included in the supervised clustering process. Each row depicts a single gene and each column depicts a single patient with chronic lymphocytic leukaemia (n=101) included in this microarray study. The y-axis shows the gene sets clustering together (cyan: 401 probes; green: 70 probes; purple: 424 probes; red: 241 probes). The x-axis shows the patient subsets. (B) Kaplan-Meier survival functions for time to progression after frontline FCR therapy in the three prognostic subsets segregated by gene expression profiling and hierarchical clustering.

	MDACC cohort				CLL8 cohort (*IGHV*-unmutated samples; n=109)	
	All samples (n=100)		*IGHV*-unmutated samples (n=65)			
	HR (95% CI)	p value	HR (95% CI)	p value	HR (95% CI)	p value
Clinical only						
Demographics						
Age at diagnosis*	0·97 (0·94–1·00)	0·0247	0·97 (0·94–1·00)	0·0311	0·97 (0·95–1·01)	0·047
Unmutated *IGHV* status	4·51 (1·99–10·21)	0·0003	NA	NA	NA	NA
β-2-microglobulin >4 mg/L	3·00 (1·59–5·67)	0·0007	2·75 (1·38–5·46)	0·0039	0·68 (0·38–1·21)	0·402
Cytogenetics						
FISH normal	1·22 (0·42–3·56)	0·7182	1·00 (0·33–3·00)	0·9958	0·92 (0·52–1·65)	0·297
del(17)(p)	..	..	..	..	4·16 (2·16–8·02)	<0·0001
del(11)(q)	1·23 (0·41–3·69)	0·7096	1·10 (0·37–3·27)	0·8707	1·31 (0·83–2·05)	0·046
del(13)(q)	3·18 (1·14–8·87)	0·0270	2·73 (0·94–7·95)	0·0658	0·75 (0·48–1·17)	0·572
Clinical plus gene expression profiling subtype						
Demographics						
Age at diagnosis*	0·94 (0·91–0·98)	0·0010	0·94 (0·91–0·98)	0·0033	ND	ND
β-2-microglobulin >4 mg/L	3·54 (1·75–7·17)	0·0004	3·24 (1·42–6·90)	0·0023	ND	ND
Cytogenetics						
FISH normal	0·72 (0·24–2·21)	0·5667	0·47 (0·14–1·61)	0·2305	ND	ND
del(11)(q)	1·22 (0·40–3·69)	0·7304	0·96 (0·32–2·92)	0·9484	ND	ND
del(13)(q)	2·77 (0·98–7·89)	0·0557	1·80 (0·23–5·48)	0·2982	ND	ND
Prognosis subtype†						
Intermediate	7·09 (1·58–31·82)	0·0105	Ref	Ref	ND	ND
Unfavourable	28·47 (6·48–125·05)	<0·0001	4·30 (2·01–9·18)	0·0002	ND	ND
Clinical plus 17-gene expression signature						
Demographics						
Age at diagnosis*	0·96 (0·93–0·99)	0·0159	0·97 (0·94–1·00)	0·0357	0·97 (0·95–1·00)	0·060
Unmutated *IGHV* status	2·93 (1·26–6·83)	0·0129	NA	NA	NA	NA
β-2-microglobulin >4 mg/L	2·59 (1·30–5·15)	0·0067	1·85 (0·95–3·62)	0·0722	0·62 (0·33–1·18)	0·149
Cytogenetics						
FISH normal	0·94 (0·32–2·76)	0·9046	NS	NS	1·31 (0·50–3·44)	0·589
del(17)(p)	..	..	..	..	8·23 (3·26–20·8)	<0·0001
del(11)(q)	1·25 (0·41–3·79)	0·6980	NS	NS	1·80 (0·96–3·39)	0·069
del(13)(q)	2·53 (0·89–7·17)	0·0805	NS	NS	1·28 (0·68–2·39)	0·442
High 17-gene expression signature	2·88 (1·51–5·52)	0·0014	3·40 (1·70–6·82)	0·0006	2·05 (1·24–3·40)	0·0052

MDACC=MD Anderson Cancer Center. CLL=chronic lymphocytic leukaemia. HR=hazard ratio. NA=not applicable. ND=analysis not done. FISH=fluorescence in-situ hybridisation. Ref=reference. NS=not selected (via Akaike Information Criterion). *Treated as a continuous variable. Thus, the HR is the change associated with 1-year increase in age. †In comparison with all samples, the favourable prognosis subtype is used as the baseline. In comparison within unmutated samples, the intermediate prognosis subtype is used as the baseline.

Table 2: **Multivariate models to predict time to progression**

the binary score, the cause-specific hazard ratio (high [intermediate] risk *vs* low [unfavourable] risk) was 1·90 (95% CI 1·18–3·06) with a significant segregation between the two groups (p=0·008, figure 3B). For multivariate analysis by cause-specific Cox regression, age at diagnosis, β-2-microglobulin of more than 4 mg/L, and cytogenetics were considered as additional covariates available in the CLL8 cohort. The effect of the continuous score (p=0·0032) and the binary risk score (p=0·0061) remained significant.

Discussion

We have developed and independently validated a robust and reproducible 17-gene signature that distinguishes treatment-naive patients with *IGHV*-unmutated chronic lymphocytic leukaemia who are likely to achieve durable

	Symbol	Standardised coefficient*	Gene name	Entrez gene identification number	Gene cluster†
Rank 1	*OSBPL5*	+0·633	Oxysterol-binding protein like-5	114 879	Purple
Rank 2	*MSI2*	+0·234	Musashi RNA-binding protein 2	124 540	Purple
Rank 3	*KSR2*	+0·219	Kinase suppressor of RAS2	283 455	Purple
Rank 4	*NME1*	+0·206	NME/NM23 nucleoside diphosphate kinase 1	4830	Purple
Rank 5	*SLC35A4*	+0·199	Solute carrier family 35 member A4	113 829	Purple
Rank 6	*TXN*	+0·188	Thioredoxin	7295	Purple
Rank 7	*LAG3*	+0·187	Lymphocyte activating 3	3902	Red
Rank 8	*ZNHIT1*	+0·162	Zinc finger HIT-type containing 1	10 467	Purple
Rank 9	*PDE8A*	–0·159	Phosphodiesterase 8A	5151	Cyan
Rank 10	*RGS10*	+0·150	Regulator of G-protein signalling 10	6001	Purple
Rank 11	*TSPO*	+0·145	Translocator protein	706	Purple
Rank 12	*CRLF3*	–0·129	Cytokine receptor-like factor 3	51 379	Cyan
Rank 13	*DCAF12*	+0·058	DDB1 and CUL4-associated factor 12	25 853	Purple
Rank 14	*ADSL*	+0·040	Adenylosuccinate lyase	158	Purple
Rank 15	*AQP1*	–0·037	Aquaporin 1 (Colton blood group)	358	Cyan
Rank 16	*GRN*	+0·025	Granulin	2896	Purple
Rank 17	*TTC38*	+0·018	Tetratricopeptide repeat domain 38	55 020	Purple

Rank is the order in which terms were introduced to the model; lower rank genes are more relevant than higher rank genes. *Weight given to the standardised gene expression (subtracting the mean and dividing by the SD of all *IGHV*-unmutated samples). The full model also uses a constant intercept term of 0·178. †Refers to the subsets of gene expression patterns as shown in figure 2.

Table 3: **Final 17-gene expression signature model to predict prognosis (time to progression) in patients with chronic lymphocytic leukaemia and an unmutated *IGHV* status**

remissions following front-line FCR chemoimmunotherapy from those who might benefit from alternative regimens. Differences in the expression of genes involved in oxidative phosphorylation and purine metabolism seem to account, at least in part, for differences in therapy response in patients with *IGHV*-unmutated chronic lymphocytic leukaemia.

Our final prognostic model uses only 17 of 1136 genes that were related to time to progression in univariate analyses. It is well known that it is difficult to reproduce lists of prognostic genes selected from different datasets.[24,25] The reasons for this instability include individual false-positive genes, tumour heterogeneity, small sample sizes, and highly correlated genes.[26,27] To address this limitation, we validated multiple steps in the model construction and, more importantly, validated the model in an independent patient cohort treated in a separate clinical trial. However, the CLL8 cohort differed slightly from the MDACC cohort—ie, the CLL8 cohort had more patients with advanced stage disease and unfavourable cytogenetics than did the MDACC cohort. The MDACC cohort contained only one patient with del(17)(p), which we excluded from the survival models. *TP53* mutation status and minimal residual disease testing were unavailable because the samples were acquired before these analyses was done routinely. Despite small patient numbers in both cohorts, we believe that our validation process ensures that the gene signatures obtained are robust. A primary feature of the model is that 13 of the 17 genes come from the cluster collection of genes shown in purple in figure 2; all 13 have positive coefficients, indicating that increased expression corresponds to increased risk of progression. By contrast, three of the 17 genes come from the cluster collection of genes in cyan in figure 2 and have negative coefficients, indicating that increased expression corresponds to decreased risk.

To understand the biological factors explaining the model, we applied gene set analysis in the context of the larger set of univariately important genes. Although not apparent from the 17 genes used to construct the model, gene set analysis shows that the cluster collection of 13 genes (shown in purple in figure 2) whose upregulation is associated with a poor outcome, is enriched for genes involved in metabolism and oxidative phosphorylation, including genes encoding proteins in complexes I, III, IV, and V, which participate in electron transport across the inner mitochondrial membrane (appendix 1 p 21). Recent studies have shown that, unlike many other highly proliferative lymphoid neoplasms and other cancers, chronic lymphocytic leukaemia cells do not rely upon aerobic glycolysis to generate energy, the Warburg effect. Rather, their major energy source is oxidative phosphorylation. Enhanced oxidative phosphorylation in chronic lymphocytic leukaemia cells is associated with poor prognostic

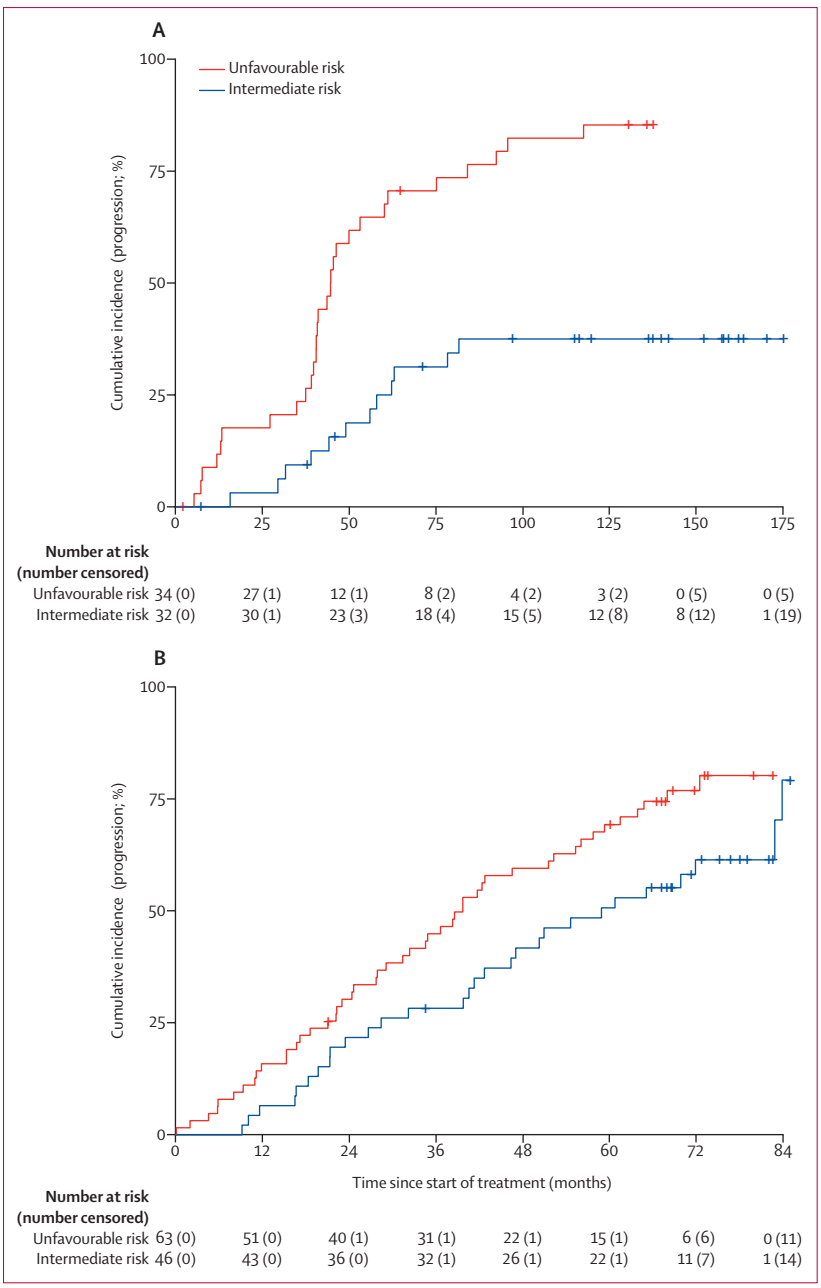

Figure 3: **Aalen-Johansen estimator of cumulative incidence of progression using the final 17-gene expression signature model**

Time to progression was defined as from the start of fludarabine, cyclophosphamide, and rituximab therapy to progression in patients with chronic lymphocytic leukaemia and an *IGHV*-unmutated status; and as a function of predicted risk class based on 17 genes in the final prognostic model. (A) Cumulative incidence plot for progression in 66 *IGHV*-unmutated cases in the MDACC cohort. 5-year cumulative incidence was 0·25 (95% CI 0·10–0·40) for intermediate risk and 0·65 (0·49–0·81) for unfavourable risk. 10-year cumulative incidence was 0·38 (0·21–0·54) for intermediate risk and 0·83 (0·73–0·97) for unfavourable risk. (B) Cumulative incidence plot for progression in 109 *IGHV*-unmutated samples in the CLL8 cohort. 5-year cumulative incidence was 0·51 (IQR 0·36–0·65) for intermediate risk and 0·69 (0·58–0·81) for unfavourable risk. 10-year cumulative incidence was not available. MDACC=MD Anderson Cancer Center.

features including unmutated *IGHV* genes, ZAP70 positivity, high Rai stage, increased serum β-2-microglobulin, and fludarabine resistance.[28,29]

The cluster collection of genes shown in purple in figure 2 is also enriched for genes involved in purine metabolism. Purines, essential components of nucleic acids, also provide energy and cofactors required for cell survival and proliferation.[30] Purine nucleotides are synthesised through two different pathways, the purine salvage pathway and the de-novo purine biosynthesis pathway. One of the most highly expressed genes in this cluster collection is adenylosuccinate lyase, which participates in both pathways. In the purine salvage pathway, this enzyme catalyses the formation of adenylate from adenylosuccinate in the conversion of inosine monophosphate into adenine nucleotides. In the de-novo pathway, adenylosuccinate lyase catalyses the cleavage of succinyl groups to yield fumarate. Other highly expressed genes in this group include adenosine deaminase and deoxyuridine triphosphate nucleotidohydrolase. These analyses suggest that differences in metabolic activity of chronic lymphocytic leukaemia cells distinguish patients with different outcomes after FCR therapy.

As expected, the vast majority of patients identified as having a favourable prognosis were those with mutated *IGHV*; about two-thirds showed del(13)(q) as the sole abnormality. However, about a third of intermediate or poor prognosis cases also had mutated *IGHV*, and one-third of these showed del(13)(q). Thus, mutated *IGHV* genes and del(13)(q) did not ensure a good prognosis. Other than the clinical outcome and the results of the 17-gene signature, we could not identify any clinical or laboratory features that would allow us to distinguish these cases from those with a good prognosis. A recent study showed that a subset of *IGHV*-mutated cases with isolated del(13)(q) show karyotypic complexity on stimulated chromosome banding analysis and have a poor prognosis.[31] We are unable to explore this possibility because our cohorts predate routine stimulated chromosome banding analysis. In our multivariate models, del(13)(q) appears to have a deleterious effect. This artifact might be due to an interaction between *IGHV* mutation status and cytogenetics that we could not investigate because of the small sample size.

Tyrosine-kinase inhibitors—such as ibrutinib, a BTK inhibitor, and idelalisib, a PI3Kδ inhibitor—have revolutionised the treatment of patients with del(17)(p) and with relapsed, refractory chronic lymphocytic leukaemia. However, treatment with tyrosine-kinase inhibitors is associated with clinically significant toxicities.[32] Patients treated with ibrutinib are at increased risk for atrial fibrillation and haemorrhage; a small number of patients have developed ventricular arrhythmias or died suddenly.[32] Patients treated with idelalisib are at increased risk for transaminitis, colitis, and pneumonitis.[32,33] About 10% of patients will discontinue therapy because of side-effects. Finally, it remains to be determined whether the disease will relapse if therapy is discontinued.

A major advantage of FCR chemoimmunotherapy with intent to cure is that the regimen is relatively brief, every

4 weeks for six cycles, and inexpensive. But treatment with FCR also has substantial limitations.[34] Up to 5% of patients treated with front-line FCR will develop a therapy-related myeloid neoplasm.[22,35] Patients older than 65 years or those with comorbidities often cannot tolerate a full course of therapy, and are at increased risk to develop myelosuppression and opportunistic infections.[34] However, for patients who desire a short therapy course, or who cannot tolerate tyrosine-kinase inhibitor therapy, FCR remains a viable option. We would recommend testing the value of the 17-gene signature in a prospective study that compares FCR treatment with alternative therapies, such as ibrutinib, as part of a randomised clinical trial.

Contributors

CDH, KRC, and LVA designed the study. CDH, KRC, AB, and JBl collected, analysed, and interpreted data for this study. JBl, MH, and SS designed and monitored data collection for the CLL8 trial. MJK designed and monitored data collection for the FCR300 trial, and AF has continued to monitor data collection. CDH, KRC, AB, JBl, ZBA, TM, JEB, JBa, KF, BAC, and CCO collected and analysed data. CDH, JBl, LLB, TM, JEB, and BAC did experiments and interpreted results.

Declaration of interests

CDH reports grants from Hoffmann-La Roche during the conduct of the study, and research funding and travel support from Roche in the context of other clinical trials. KRC reports grants from the National Institutes of Health/National Cancer Institute (NIH/NCI) during the conduct of the study; and grants from the NIH/NCI and NIH/National Library of Medicine outside the submitted work. JBa reports honoraria and travel support from Roche during the conduct of the study. KF reports non-financial support from Roche during the conduct of the study, and personal fees from AbbVie outside the submitted work. MH reports grants or other support, or both, from Roche, AbbVie, Gilead, Janssen, and Celgene outside the submitted work. SS reports grants, personal fees, and non-financial support from Hoffmann-La Roche during the conduct of the study; and grants, personal fees, and non-financial support from AbbVie, Amgen, AstraZeneca, Celgene, Gilead, GSK, Hoffmann-La Roche, Janssen, Novartis, Pharmacyclics, and Sunesis outside the submitted work. MJK reports grants from AbbVie during the course of the study. LVA reports grants from the Chronic Lymphocytic Leukemia Global Research Foundation and NIH/NCI during the conduct of the study. All other authors declare no competing interests.

Acknowledgments

This work is supported by grants from the Chronic Lymphocytic Leukaemia Global Research Foundation and the National Institutes of Health/National Cancer Institute (R01 CA182905 and P30 CA016058).

References

1 Hallek M, Fischer K, Fingerle-Rowson G, et al. Addition of rituximab to fludarabine and cyclophosphamide in patients with chronic lymphocytic leukaemia: a randomised, open-label, phase 3 trial. *Lancet* 2010; **376:** 1164–74.

2 Keating MJ, O'Brien S, Albitar M, et al. Early results of a chemoimmunotherapy regimen of fludarabine, cyclophosphamide, and rituximab as initial therapy for chronic lymphocytic leukemia. *J Clin Oncol* 2005; **23:** 4079–88.

3 Fischer K, Bahlo J, Fink AM, et al. Long-term remissions after FCR chemoimmunotherapy in previously untreated patients with CLL: updated results of the CLL8 trial. *Blood* 2016; **127:** 208–15.

4 Thompson PA, Tam CS, O'Brien SM, et al. Fludarabine, cyclophosphamide, and rituximab treatment achieves long-term disease-free survival in *IGHV*-mutated chronic lymphocytic leukemia. *Blood* 2016; **127:** 303–09.

5 Eichhorst B, Fink AM, Bahlo J, et al. First-line chemoimmunotherapy with bendamustine and rituximab versus fludarabine, cyclophosphamide, and rituximab in patients with advanced chronic lymphocytic leukaemia (CLL10): an international, open-label, randomised, phase 3, non-inferiority trial. *Lancet Oncol* 2016; **17:** 928–42.

6 Duzkale H, Schweighofer CD, Coombes KR, et al. LDOC1 mRNA is differentially expressed in chronic lymphocytic leukemia and predicts overall survival in untreated patients. *Blood* 2011; **117:** 4076–84.

7 McCarthy H, Wierda WG, Barron LL, et al. High expression of activation-induced cytidine deaminase (AID) and splice variants is a distinctive feature of poor-prognosis chronic lymphocytic leukemia. *Blood* 2003; **101:** 4903–08.

8 Schweighofer CD, Coombes KR, Majewski T, et al. Genomic variation by whole-genome SNP mapping arrays predicts time-to-event outcome in patients with chronic lymphocytic leukemia: a comparison of CLL and HapMap genotypes. *J Mol Diagn* 2013; **15:** 196–209.

9 Admirand JH, Knoblock RJ, Coombes KR, et al. Immunohistochemical detection of ZAP70 in chronic lymphocytic leukemia predicts immunoglobulin heavy chain gene mutation status and time to progression. *Mod Pathol* 2010; **23:** 1518–23.

10 Rassenti LZ, Huynh L, Toy TL, et al. ZAP-70 compared with immunoglobulin heavy-chain gene mutation status as a predictor of disease progression in chronic lymphocytic leukemia. *N Engl J Med* 2004; **351:** 893–901.

11 Zenz T, Dohner H, Stilgenbauer S. Genetics and risk-stratified approach to therapy in chronic lymphocytic leukemia. *Best Pract Res Clin Haematol* 2007; **20:** 439–53.

12 Stilgenbauer S, Schnaiter A, Paschka P, et al. Gene mutations and treatment outcome in chronic lymphocytic leukemia: results from the CLL8 trial. *Blood* 2014; **123:** 3247–54.

13 Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. *Bioinformatics* 2008; **24:** 1547–48.

14 Pounds S, Morris SW. Estimating the occurrence of false positives and false negatives in microarray studies by approximating and partitioning the empirical distribution of p-values. *Bioinformatics* 2003; **19:** 1236–42.

15 Vasyutina E, Boucas JM, Bloehdorn J, et al. The regulatory interaction of *EVI1* with the *TCL1A* oncogene impacts cell survival and clinical outcome in CLL. *Leukemia* 2015; **29:** 2003–14.

16 Cheson BD, Bennett JM, Grever M, et al. National Cancer Institute-sponsored Working Group guidelines for chronic lymphocytic leukemia: revised guidelines for diagnosis and treatment. *Blood* 1996; **87:** 4990–97.

17 Burnham KAD. Model selection and multimodel inference: a practical information-theoretic approach. New York, NY: Springer-Verlag, 2002.

18 Tibshirani R. The Lasso method for variable selection in the Cox model. *Stat Med* 1997; **16:** 385–95.

19 Graf E, Schmoor C, Sauerbrei W, Schumacher M. Assessment and comparison of prognostic classification schemes for survival data. *Stat Med* 1999; **18:** 2529–45.

20 Brier G. Verification of forecasts expressed in terms of probability. *Mon Weather Rev* 1950; **78:** 1–3.

21 Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. *Control Clin Trials* 1996; **17:** 343–46.

22 Tam CS, O'Brien S, Wierda W, et al. Long-term results of the fludarabine, cyclophosphamide, and rituximab regimen as initial therapy of chronic lymphocytic leukemia. *Blood* 2008; **112:** 975–80.

23 Short NJ, Keating MJ, Wierda WG, et al. Fludarabine, cyclophosphamide, and multiple-dose rituximab as frontline therapy for chronic lymphocytic leukemia. *Cancer* 2015; **121:** 3869–76.

24 Ein-Dor L, Kela I, Getz G, Givol D, Domany E. Outcome signature genes in breast cancer: is there a unique set? *Bioinformatics* 2005; **21:** 171–78.

25 Ntzani EE, Ioannidis JP. Predictive ability of DNA microarrays for cancer outcomes and correlates: an empirical assessment. *Lancet* 2003; **362:** 1439–44.

26 Michiels S, Koscielny S, Hill C. Interpretation of microarray data in cancer. *Br J Cancer* 2007; **96:** 1155–58.

27 Xu JZ, Wong CW. Hunting for robust gene signature from cancer profiling data: sources of variability, different interpretations, and recent methodological developments. *Cancer Lett* 2010; **296:** 9–16.

28 Sharma A, Janocha AJ, Hill BT, Smith MR, Erzurum SC, Almasan A. Targeting mTORC1-mediated metabolic addiction overcomes fludarabine resistance in malignant B cells. *Mol Cancer Res* 2014; **12:** 1205–15.

29 Vangapandu HV, Ayres ML, Bristow CA, et al. The stromal microenvironment modulates mitochondrial oxidative phosphorylation in chronic lymphocytic leukemia cells. *Neoplasia* 2017; **19:** 762–71.

30 Yin J, Ren W, Huang X, Deng J, Li T, Yin Y. Potential mechanisms connecting purine metabolism and cancer therapy. *Front Immunol* 2018; **9:** 1697.

31 Baliakas P, Jeromin S, Iskas M, et al. Cytogenetic complexity in chronic lymphocytic leukemia: definitions, associations, and clinical impact. *Blood* 2019; **133:** 1205–16.

32 Mauro FR, Caputo MD, Rosati S, et al. Balancing efficacy and toxicity of targeted agents currently used for the treatment of patients with chronic lymphocytic leukemia. *Expert Rev Hematol* 2018; **11:** 601–11.

33 Brown JR, Kay NE. Chemoimmunotherapy is not dead yet in chronic lymphocytic leukemia. *J Clin Oncol* 2017; **35:** 2989–92.

34 Skarbnik AP, Faderl S. The role of combined fludarabine, cyclophosphamide and rituximab chemoimmunotherapy in chronic lymphocytic leukemia: current evidence and controversies. *Ther Adv Hematol* 2017; **8:** 99–105.

35 Tam CS, Seymour JF, Prince HM, et al. Treatment-related myelodysplasia following fludarabine combination chemotherapy. *Haematologica* 2006; **91:** 1546–50.