Prognostic Value of Immunophenotyping in Acute Myeloid Leukemia

By Kenneth Bradstock, Jane Matthews, Elizabeth Benson, Fiona Page, James Bishop, and the Australian Leukaemia Study Group

The diagnostic and prognostic value of immunophenotyping with 18 murine monoclonal antibodies (MoAbs) to a variety of leucocyte differentiation antigens was assessed in 168 adults aged 15 to 60 years with acute myeloid leukemia (AML). Patients were entered on the multicentre Australian Leukaemia Study Group M4 protocol, and were randomized to receive either standard or high-dose Ara-C together with daunorubicin and etosode as induction chemotherapy, followed by standard consolidation and maintenance therapy. Diagnostic bone marrow aspirate (152 cases) or peripheral blood samples (16) were analyzed by indirect immunofluorescence and flow cytometry. MoAbs used were directed at myeloid (CD11b, CD13, CD14, CD15, CD33, CD41), lymphoid (CD2, CD3, CD7, CD9, CD10, CD19), or stem cell (HLA-DR, CD34, c-kit receptor) antigens, as well as the leucocyte integrins CD18 and CD49e, and the transferrin receptor CD71. Of the myeloid markers, CD13 and CD33 were the most useful diagnostically (71% and 79% of cases positive, respectively), with CD11b, CD14, and CD15 less commonly positive. A minority of cases expressed lymphoid antigens, either T cell (CD2 16%, CD3 7%, CD7 28%) or B cell (CD10 2%, CD19 7%). CD34 was detected on 42% and c-kit receptor on 48%. When patients were analyzed for response to treatment, CD2, CD9, and CD14 were significantly associated with complete remission rate: cases expressing these antigens had a poorer response than negative cases. In univariate analysis, CD11b cases had shorter periods of remission (relative risk of relapse, 2.33; P = .003) and shorter survival (relative death rate, 1.97; P = .008). In multivariate analysis, adjusting for other prognostic factors, CD9 and CD11b were significantly predictive of shorter survival. No other marker had a significant predictive effect. We conclude that myeloid MoAbs are useful in confirming the diagnosis of AML, but their prognostic value may be limited to CD11b. Lymphoid antigen expression is a consistent phenomenon in a minority of cases of AML, but appears to have little clinical significance.

© 1994 by The American Society of Hematology.

ACUTE MYELOID leukemia (AML) in adults is a heterogeneous disease, with a variable response to chemotherapy with anticancer cytotoxic agents. Using combination chemotherapy protocols, approximately 60% of patients achieve a complete remission (CR) but only a minority remain leukemia free.1-3 A variety of clinical and biologic parameters have been examined for potential value in predicting treatment response and survival; among these is immunophenotype.4-7 Leukemic myeloblasts express a variety of leucocyte differentiation antigens, which reflect commitment to the myeloid lineage as well as level of maturation.8,9 These antigenic phenotypes have proven very useful in the diagnosis of AML, but the prognostic value has remained uncertain. Initial reports suggested a relationship between patterns of myeloid lineage differentiation antigens and patient prognosis,10 but subsequent studies have produced conflicting and inconsistent results.11-17 The only markers with significant prognostic associations in more than one study have been CD13,10,15 CD14,10,14,17 and CD1512,15 with the stem cell marker CD34 also recently claimed to be associated with poor clinical outcome in AML.12,16-19

In addition, leukemic cells in AML can also express lymphoid-lineage antigens, a phenomenon variously described as lineage infidelity,20 lineage promiscuity,21,22 mixed-lineage acute leukemia,23 or hybrid or biphenotypic leukemia.24 The prognostic significance of this phenomenon in AML has also been examined, but with highly conflicting claims of poorer,22,25 better,26 or unaltered prognosis.14

In this report, we attempt to clarify the clinical value of immunophenotype in AML, within the context of a large clinical trial of chemotherapy for adult AML. Blast cells from 168 cases were analyzed with a uniform panel of monoclonal antibodies (MoAbs) directed at myeloid, lymphoid, stem cell, and other leucocyte antigens, and results correlated with treatment outcome.

MATERIALS AND METHODS

Patient population. Cases suitable for immunophenotypic evaluation were derived from the M4 study conducted by the Australian Leukaemia Study Group (ALSG).27 This study accrued 301 patients aged 15 to 60 years with newly diagnosed untreated AML at 25 participating centres throughout Australia. Patients were randomized at diagnosis to receive conventional dose therapy with the 737 protocol (100 mg/m2 Ara-C per day for 7 days by continuous intravenous infusion, 50 mg/m2 daunorubicin intravenously on days 1, 2, and 3, and 100 mg/m2 etosode intravenously on days 1 to 7, inclusive) or high-dose therapy with HIDAC-37 (consisting of 3 g/m2 ARA-C by 3-hour infusion every 12 hours on days 1, 3, 5, and 7, with daunorubicin and etosode as above), as previously described.27 Patients received up to two induction courses, and if in complete remission, were then planned to be treated with two cycles of 525 protocol (100 mg/M2 Ara-C per day for 5 days by continuous intravenous infusion, 50 mg/M2 daunorubicin intravenously on days 1 and 2, and 100 mg/M2 etosode intravenously on days 1 to 5), followed by 24 months maintenance treatment with 2 monthly cycles of ARA-C plus daunorubicin or thioguanine. The CR rate was 71% in the HIDAC-37 arm, and 74% in the 737 arm. The median survival was 1.6 years for the HIDAC-37 arm and 1.4 years for the 737 arm, with estimated 5-year survival rates of 29% and 24%, respectively. Immunophenotyping was performed on 217 of the 301 patients.
entered on the trial. In 159 cases, the material analyzed was bone marrow (BM) aspirate; in 38, peripheral blood (PB) was used; whereas in 20 cases, both PB and BM were used. To exclude cases with low proportions of leukemic blast cells that might complicate the analysis, further morphologic eligibility criteria were applied to the 217 phenotyped cases. All cases were reviewed by a central morphology committee, with a classification being assigned according to the guidelines published by the French-American-British group.28 Marrow aspirates with less than 40% blasts were excluded whereas in 20 cases, both PB and BM were used. To exclude cases of the 179 BM samples, and 22 (58%) of the 38 PB samples. The 217 phenotyped cases. All cases were reviewed by a central morphology committee, with a classification being assigned according to the guidelines published by the French-American-British group.28 Marrow aspirates with less than 40% blasts were excluded where there was an excellent correlation between the two, only ungated results are presented. Positivity for each MoAb was arbitrarily defined as 20% or more of cells positive above the isotype-matched negative control. Comparative analysis of 7 cases where both PB and BM were analyzed showed a low discrepancy rate (5/73 or 7%) between positive and negative results, thereby validating the use of PB cells in those cases where marrow aspirate was not available. Finally, to rule out the possibility that reactivity with CD2 and CD3 antibodies was caused by contamination with normal T lymphocytes, CD2+ and CD3+ cases were further examined to ensure an overlap of expression of these markers on the leukemic populations, as defined by the percentage of CD2+ or CD3+ cells being the highest value for a lineage (CD13, CD34) or immature (HLA-DR, CD34) marker exceeding 110%. This redefinition resulted in the reclassification of four CD2+ and four CD3+ cases.

**Immunofluorescence analysis.** Isolation of leukemic cells, staining with MoAbs, and immunofluorescence analysis by flow cytometry were performed according to a common protocol in the laboratories associated with the 25 participating centres. Briefly, heparinized marrow aspirate or PB samples taken at diagnosis were diluted in phosphate-buffered saline and centrifuged on Ficoll-Isopaque (Pharmacia, Uppsala, Sweden). Washed mononuclear cells were incubated with predetermined amounts of MoAbs or isotype-matched negative controls, washed, then incubated with sheep antiserum to mouse Ig conjugated with fluorescein isothiocyanate (Silenus, Melbourne, Australia). MoAbs listed in Table 1 were obtained through participants (Westmead series, K. Bradstock, Westmead Hospital, Sydney; FMC series, H. Zola, Flinders Medical Centre, Adelaide; PHM-2, W. Atkis, Prince Henry Hospital, Melbourne; CIMT, G. Pilkington, Peter MacCallum Cancer Institute, Melbourne; YB5.B8, L. Ashman, Royal Adelaide Hospital, Adelaide, Australia), or purchased commercially (B4, MY10, Coulter, Hialeah, FL; OKT-9, Johnson and Johnson, North Ryde, New South Wales, Australia). Labeled cells were examined by flow cytometry using instruments available at each institution. Results were recorded on both ungated and blast-gated populations (defined by forward- and right-angle scatter characteristics), but because there was an excellent correlation between the two, only ungated results are presented. Positivity for each MoAb was arbitrarily defined as 20% or more of cells positive above the isotype-matched negative control. Comparative analysis of 7 cases where both PB and BM were analyzed showed a low discrepancy rate (5/73 or 7%) between positive and negative results, thereby validating the use of PB cells in those cases where marrow aspirate was not available. Finally, to rule out the possibility that reactivity with CD2 and CD3 antibodies was caused by contamination with normal T lymphocytes, CD2+ and CD3+ cases were further examined to ensure an overlap of expression of these markers on the leukemic populations, as defined by the percentage of CD2+ or CD3+ cells being the highest value for a lineage (CD13, CD34) or immature (HLA-DR, CD34) marker exceeding 110%. This redefinition resulted in the reclassification of four CD2+ and four CD3+ cases.

**Statistical methods.** The data were analyzed with a close-out date of February 1, 1993 to prevent bias in the follow-up of patients. The status of all patients with immunophenotyping done was known at this date. All data were censored at the close-out date. The median duration of follow-up was 2.6 (range, 1.5 to 5.4) years for patients who were lost to follow-up or censored at the close-out date. Overall survival was measured from the date of randomization. All deaths were included as failures. Duration of CR was measured from the date of achieving CR to the date of relapse for patients who relapsed, or from the date of last contact or close-out for patients who did not
relapse. Patients who died in CR had their response duration censored at the date of death.

Data were analyzed using BMDP statistical software. Survival curves were estimated using the Kaplan-Meier product-limit method and compared using the Cox proportional hazards model. Complete response rates were compared using Fisher’s exact test. The prognosis of the 168 cases selected for analysis was compared with that of the 133 excluded cases. There was no evidence of selection bias. The CR rates for the 2 groups were 73% and 71%, respectively (P = .7), and there was no significant difference in the duration of CR (P = .11) or in overall survival distribution (P = .87). Prognostic factors were identified using stepwise regression methods for all eligible patients in the trial using a logistic regression model for the probability of achieving CR and the Cox proportional hazards model for the duration of CR and overall survival. A screening level of significance of .01 was chosen. All major on-study factors were tested, together with an indicator variable for the treatment arm. Factors were treated as continuous and/or indicator variables with breakpoints determined by univariate analyses and a review of the relevant literature. The only significant factor identified for increasing the probability of achieving CR was age less than 40. The significant factors increasing the duration of response were treatment with the HIDAC-37 arm and an on-study peripheral blast count less than 10 × 10^9/L. The significant factors increasing the duration of survival were age less than 40, an on-study white blood cell count less than 100 × 10^9/L, and female sex. The significance of an individual antigen adjusting for the relevant prognostic factors was calculated on the basis of the difference in the values of the likelihood functions before and after including the antigen in a model containing these factors. Two-tailed P values have been reported throughout the text with no adjustment made for multiple comparisons. Results with a significance level (P value) less than .05 have been presented as being statistically significant. However with 18 antigens studied, the screening level of significance for a single endpoint (eg, survival duration) should be set at .0028 (0.05/18) to ensure that the overall probability of a type 1 error is less than .05. Ninety-five percent confidence intervals (CIs) have been reported for the main results.

RESULTS

Reactivity pattern of antibodies with AML. The reactivity of the panel of MoAbs with cases of AML is presented in Table 1. Among the myeloid lineage antibodies, CD13 and CD33 were most often positive (71% and 79% of cases positive, respectively). CD11b reacted with 40% of cases, whereas CD14 and CD15 were positive on less than 20%. The hematopoietic progenitor cell markers CD34 and c-kit receptor were detected on 42% and 48%, respectively, whereas HLA-DR was present on 70%. The platelet glycoprotein CD41 was found on only a small percentage (12%).

A minority of cases showed expression of lymphoid lineage markers. Apart from the high proportion of cases reactive with CD9 antibody (a marker known to be present on both normal platelets and B lymphocytes), 2% to 28% of cases were positive with MoAbs to T- or B-cell antigens. CD7 was the T-cell marker most frequently detected (28%), with 16% of cases CD2^+ and 7% CD3^+, whereas of the B-cell markers, 7% of cases were CD19^+ and 2% CD10^+.

Of the other markers tested, the beta-1 integrin CD49e and the beta-2 integrin CD18 were shown on most cases (84% and 72%). The transferrin receptor CD71 was also found on 57% of cases.

Relationship between immunophenotype and treatment outcome. The association between phenotype and response to initial induction therapy is shown in Table 1. Only CD9, CD14, and CD2 showed a significant association with CR rate (P = .03, .03, and .04, respectively, in the univariate analysis). CD14 and CD9 remained significant in the multivariate analysis adjusting for age (P = .02 and .01, respectively), but CD2 was no longer significant (P = .11). No other marker was predictive for achieving CR. In view of the recent report that CD34^+ cases of AML respond poorly to conventional dose therapy, but have a better response to protocols containing high-dose Ara-C, we examined the group of CD34^+ cases for a difference in response on the 737 versus HIDAC-37 protocols. No difference in CR rate was seen; 19 (73%) of 26 cases treated with 737 achieved CR, whereas 19 (79%) of 24 cases treated with HIDAC-37 achieved CR (P = .7).

CD11b was the only marker that was significantly associated with shorter duration of CR (P = .003 on univariate analysis, P = .002 adjusting for the significant on-study factors). CD11b was also significantly associated with shorter overall survival duration (Table 1, Fig 1), both on univariate analysis (P = .006) and multivariate analysis (P = .003). The only other markers that were significantly predictive of shorter overall survival duration in univariate or multivariate analysis were CD9 (P = .04 on univariate analysis, P = .014 on multivariate analysis) and CD10 (P = .011 on univariate analysis, P = .002 on multivariate analysis), however, only two patients were CD10^+. No other marker was predictive of survival duration.

DISCUSSION

In this large multivariate study, the diagnostic value of MoAbs directed at myeloid and progenitor cell-related antigens was similar to that described in other studies.

CD13 and CD33 most frequently identified cases of AML, with CD11b, CD14, and CD15 less diagnostically useful. Considerable variability in the reactivity of CD14 and CD15 MoAbs with AML has been noted by other investigators, partly because of the known epitopic heterogeneity of these antigens. The reactivity of the CD14 and CD15 antibodies used in our study was at the lower end of the reported range, and may have limited any clinical significance associated with these markers. The incidence of positivity with CD34 and c-kit antibodies is similar to that previously published.

Our study also confirmed the phenomenon of lymphoid antigen expression on a minority of cases, within the range noted by other investigators (see review by Drexler et al).

The prognostic significance of immunophenotyping in AML has been a controversial issue, with a number of conflicting reports. In our study, only five antibodies were significantly (P < .05) associated with prognosis. In both univariate and multivariate analysis, CD2, CD9, and CD14 expression were predictive for CR rate. CD11b was the only antibody that showed a significant association with both remission duration and patient survival. CD9, CD10, and CD11b were all significantly associated with survival duration after adjusting for significant factors. The reported relationship between myeloid markers and prognosis in AML has been inconsistent. Reactivity with CD13 MoAbs has
been associated with a significantly low rate of CR in two studies, but not in others. There was no suggestion of an association in our study: 71% of CD13 cases achieved CR compared with 74% of CD13 cases. CD14 cases have been reported to have a lower CR rate in two series as well as our own, and poorer survival in two other studies. However in one of these studies, cases were divided on the basis of a second phenotype (CD14+/HLA-DR or CD14 / CD16), whereas in the other study, survival was only significantly different when a 50% threshold was used for positivity. With this threshold, CD14 was not significantly associated with survival in our series (P = 0.14 in univariate analysis and P = 0.38 in multivariate analysis). None of the other series reported a correlation of CD14 with prognosis. CD11b has been associated with a significantly shorter duration of remission and with a shorter survival in one study; survival rates at 1 year were reported to be 54% (SE 6%) for CD11b cases and 41% (SE 5%) for CD11b cases. The comparable rates in our study were 75% (SE 5%) and 54% (SE 7%), respectively. The marker with perhaps the strongest reported negative effect on CR is CD34, However, we could not confirm this finding, nor document a difference in CR rate for CD34 patients treated on protocols of different intensity, as noted by others. The expression of CD15 has been associated with an improved survival in one study, and longer remission duration in another. Although the proportion of CD15 patients was low, we did not observe any difference in treatment outcome in CD15 cases.

The clinical relevance of lineage infidelity, or lymphoid antigen expression, in AML has been highly controversial. Although most patients could be salvaged with alternative protocols, initial reports in pediatric AML suggested that cases expressing CD2 and CD7 antigens were biologically different to other forms of AML, and had a poor response to standard induction chemotherapy protocols. Similar findings were reported by another group. The inferior CR rate observed in CD2 cases in our study is in keeping with these results. In contrast, Ball and colleagues described a high remission rate and improved remission duration and survival in adult AML patients expressing CD2 and/or CD19 antigens, suggesting an improved outcome for these patients. With regard to the other lymphoid-associated markers in our series, CD9 was associated with lower remission rate and poorer survival. CD9 cases were reported to have a significantly lower response rate in one of two series, in which this marker was studied, but no reduction in overall survival has been reported before. No significant correlations of CD16 with prognosis have been reported. In our series, only two cases were CD10. No prognostic significance of lymphoid antigen expression in adult AML was found in five other studies.

Although the results from various studies have appeared conflicting, it is possible that some antibodies are not being reported as being significant because of the relatively small sample sizes involved in the studies. For example, if a particular antigen is positive in 50% of cases studied and positivity is associated with a 50% increase in the death rate, a total sample size of 192 patients (all followed until death) would be required to have a probability (power) of 0.8 of detecting an overall survival difference between positive and negative cases using a log-rank test with a two-tailed significance level of 0.05. If only half the patients will have died at the time of a planned analysis, twice as many patients will be
needed. If the percentage of positive cases differs from 50%, even more cases will be required. On the other hand, when many different antigens are being studied and no adjustments for multiple comparisons are made, the risk of a type I (false-positive) error is greatly increased. Because many of the studies provide summary statistics only for the significant results, it is difficult to determine if there has been a consistent pattern in the results for individual antigens. More detailed reporting of results is required to enable the outcomes from various studies to be compared.

In conclusion, CD11b was the only myeloid marker with major predictive value in this large adult AML series, and it is likely that other prognostic variables such as indicators of tumor burden and cytogenetics are of greater importance. The prognostic significance of CD2 expression on leukemic myeloblasts requires further examination.

ACKNOWLEDGMENT

We are grateful to the following scientists for providing immunophenotyping data: Anna Flokis, Prince of Wales Hospital, Sydney; Sue Francis, Royal Prince Alfred Hospital, Sydney; Desley Scott, Princess Alexandra Hospital, Brisbane; Russell Collins, Royal Brisbane Hospital; Katherine McGrath, Royal Melbourne Hospital; Glenn Pilkington, Peter MacCallum Hospital, Melbourne, Australia. Our thanks to Judith Maybin for preparation of the manuscript, and to the individuals who contributed antibodies for the study. The following clinicians participating in the ALSGM study contributed clinical data: M. Piccock, Woden Valley Hospital, Australian Capital Territory; J. Gallo, Lidcombe Hospital, New South Wales (NSW); H. Kronenberg, P. Vincent, and G. Young, Royal Prince Queensland (QLD); R. Cobcroft and Whiteside, Alfred Hospital, Victoria; I. Cooper, J. Ding, E. Januszewicz, and M. Wolf, Peter MacCallum Cancer Institute, Victoria; A. Enno and M. McLochlin, Prince of Wales Hospital, NSW; A. Manoharan, St. George Hospital, NSW; D. Lasser, A. Maupas, J. Gentilhomme O., Ehramas A., Fiere D.: Surface marker expression in adult acute myeloid leukaemia: Correlations with initial characteristics, morphology and response to therapy. Br J Haematol 72:161, 1989

REFERENCES

IMMUNOPHENOTYPING IN AML

Prognostic value of immunophenotyping in acute myeloid leukemia. 
Australian Leukaemia Study Group

K Bradstock, J Matthews, E Benson, F Page and J Bishop

Updated information and services can be found at:
http://www.bloodjournal.org/content/84/4/1220.full.html
Articles on similar topics can be found in the following Blood collections

Information about reproducing this article in parts or in its entirety may be found online at:
http://www.bloodjournal.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at:
http://www.bloodjournal.org/site/misc/rights.xhtml#reprints

Information about subscriptions and ASH membership may be found online at:
http://www.bloodjournal.org/site/subscriptions/index.xhtml