The Prognosis in Aplastic Anemia

By Robert E. Lynch, Darryl M. Williams, James C. Reading, and George E. Cartwright

The biphasic shape of the survival curve of 99 patients with aplastic anemia suggested that there may be at least two subgroups of patients with this disease, one with a very short survival and another with a longer survival. Patients who survived for 4 mo or less after the first clinic visit (group A) were different from the patients who survived longer (group B) with respect to their modes of onset, sex, intervals from the onset of symptoms to first clinic visit, and initial hematologic values. These differences suggested that short survival could be predicted from data available at the first contact with the physician. From these measurements, a prognostic index could be calculated which was useful in identifying the patients in group A. Although this method of prognostication needs further testing, if validated, it may prove useful in selecting patients for therapeutic trials and could explain the divergent results in previous studies of androgen treatment of aplastic anemia. When our androgen-treated subjects were compared with subjects with a similar prognostic index who had not received androgens, a beneficial effect of androgen therapy on survival could not be demonstrated.

A PLASTIC ANEMIA IS a disease with a variable course. While some patients live for only a few days after the onset of their symptoms, others may survive for years. In some instances, the survival curve has suggested that there could be two subpopulations of individuals with aplastic anemia: one with a very rapid, and another with a vastly slower, rate of dying. If such were the case, the interpretation of therapeutic trials would be most difficult, since prolongation of survival in a treated group might reflect a smaller proportion of patients with a high risk of early death, rather than the effect of therapy.

Since the refinement of techniques for homologous bone marrow transplantation has made it feasible to offer this new, but hazardous, therapy to individuals who have little chance of survival without it, the need to predict survival in a patient with aplastic anemia has increased. If patients with rapidly lethal aplasia could be separated from the others, those with a better chance of survival without treatment could be spared the cost and risks of attempted marrow engraftment.

The present study was undertaken in an effort to supply prognostic information in aplastic anemia. From a previously published retrospective analysis of a group of patients, we have attempted: (1) to identify the short survivors, (2) to contrast them with the longer survivor, (3) to devise a prognostic formula that would predict the probability of short survival, (4) to assess the validity of such a prediction by cross-validation, and (5) to reexamine the question of whether androgens are of benefit in aplastic anemia.
MATERIALS AND METHODS

Patient Population
The characteristics of the 101 patients were described in a previous report. Two patients were eliminated because they were observed for less than 2 mo before being lost to follow-up.

Definitions
The reticulocyte counts in per cent were corrected for anemia but not for stress reticulocytes by multiplying the determined value by the patient's volume of packed red cells (VPRC) divided by the mean normal VPRC for sex. The neutrophil concentration refers to the absolute number of band and polymorphonuclear neutrophils per microliter of blood. The per cent nonmyeloid cells in the marrow aspirates is meant to represent the fraction of nonhemopoietic cells present. It was computed by dividing the total cells counted minus the sum of normoblasts, granulocyte precursors, and megakaryocytes by the total cells counted. The intervals from the onset of symptoms to the first visit to our clinic, and from the first visit to our clinic to death, were calculated in whole months. The type of onset of symptoms was classified as hemorrhagic if any form of bleeding (including petechiae or ecchymoses) was the first symptom noted.

Statistical Methods
To attempt to classify patients into group A or group B (see Results) on the basis of data available at the first clinic visit, two types of discriminant analysis were performed on a UNIVAC Computer. Discriminant analysis, one type of multivariate statistical analysis, produces that linear relation of variables (discriminant function or, in this case, prognostic formula) which maximizes the variation between groups relative to the variation within groups and is designed to take into account the dependency between observations on the same individual.

The first type of discriminant analysis was a stepwise procedure using the BMDO7M program. By this technique, the variables which made significant independent contributions to the discrimination between the members of groups A and B were selected. The second type of discriminant analysis using program BMDO4M generated a prognostic formula in which an appropriate coefficient was given each selected measurement so that the numerical values of the prognostic index for the members of groups A and B were as different as possible.

No values were available for the initial reticulocyte count in 11 patients, the initial platelet count in four patients, and the initial per cent nonmyeloid cells in the marrow in ten patients. The principal component method was used to supply these missing values, since the discriminant analysis required that each patient have a complete set of data. However, in the comparison of groups in Table 1, only actual data are compared.

### Table 1. Comparison of Characteristics of Groups A and B

<table>
<thead>
<tr>
<th>Observation</th>
<th>Group A (Survival ≤ 4 mo)</th>
<th>Group B (Survival &gt; 4 mo)</th>
<th>P(A vs. B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemorrhagic onset (%)</td>
<td>83</td>
<td>47</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Sex (% male)</td>
<td>67</td>
<td>45</td>
<td>0.06*</td>
</tr>
<tr>
<td>Mean interval from onset to first visit (mo)</td>
<td>3.6</td>
<td>14.7</td>
<td>&lt;0.001†</td>
</tr>
<tr>
<td>Mean reticulocyte count (%)</td>
<td>0.6</td>
<td>1.3</td>
<td>&lt;0.001†</td>
</tr>
<tr>
<td>Mean neutrophils (cells/cu mm)</td>
<td>550</td>
<td>1226</td>
<td>&lt;0.001†</td>
</tr>
<tr>
<td>Mean platelets (thousands/cu mm)</td>
<td>17</td>
<td>42</td>
<td>&lt;0.001†</td>
</tr>
<tr>
<td>Mean per cent nonmyeloid cells in marrow</td>
<td>69</td>
<td>42</td>
<td>&lt;0.001†</td>
</tr>
</tbody>
</table>

*Chi square.
†Wilcoxon rank sum.
RESULTS

Survival

Survival was plotted in months after the onset of symptoms. The curve was constructed by dividing the number of patients alive at each month after the onset of symptoms by the number of patients who had been observed for that period. When lost to follow-up, a patient was subtracted from the number known to be alive and from the number at risk. No correction for expected mortality from other causes was made because the mean age of the population was low (33 yr), the interval for which survival was plotted was short (5 yr), and all of the deaths were attributable to aplastic anemia.

When plotted semilogarithmically (Fig. 1), the survival curve can be approximated by two intersecting straight lines, suggesting that shortly after the onset of symptoms a rapid rate of death predominates, while a slower decline in survival obtains later. Such a finding could be produced if the over-all survival were a composite of the survivals of two groups of individuals, short survivors and longer survivors. If that were the case, the two groups might be separable simply by choosing an interval beyond which none of the short survivors lived. Most of the patients dying before that time would be short survivors, while those who live longer should all be longer survivors. To arrive at an estimate of the size of the two groups, the slopes of the second phase of the curve was extrapolated to 0 mo (Fig. 1), providing an estimate of the percentage of the total group who were longer survivors (56%). The short survivors should, by subtraction, comprise 44% of the total.

Fig. 1. Survival after the onset of symptoms of 99 patients with aplastic anemia.
Unfortunately, Fig. 1 would not suffice to define the long and short survival groups, both because the shape of the curve might reflect the patterns of referral to our clinic, and because information from which to forecast his outcome was not available on each patient at the onset of his symptoms. Both difficulties could be obviated by reploting survival by months after the first clinic visit. Although that curve (Fig. 2) showed an equally impressive early mortality, it was not so satisfyingly resolved into two components. Somewhat arbitrarily, the patients dying within 4 mo after the first visit to our clinic were designated group A (a group of short survivors). This interval was selected because (1) the number of patients dying in this time (42) approximated the number of short survivors estimated from Fig. 1; (2) a natural dividing line was suggested, since only one patient had died at 4 mo after the first visit (data not shown); and (3) 4 mo seemed a clinically relevant interval both because a patient who is likely to die sooner has a poor prognosis indeed, and because 3–4 mo is frequently said to be the minimum effective duration of androgen therapy.

Factors That Correlate With Survival

Having defined the group-A patients as those individuals dying within 4 mo of the first visit (and recognizing that this group might contain a few patients who belong in group B) and the group B patients as those surviving longer than 4 mo after the first visit, we compared the two groups with respect to sex, mode of onset, initial blood counts, intervals between the onset of symptoms and the first visit to our clinic, and per-cent nonmyeloid cells in the initial marrow aspirates (Table I). These particular measurements were selected because they
were the ones identified by the stepwise discriminant analysis (BMOD7M) as those that helped distinguish between the members of group A and group B, and because it was anticipated that they might correlate with the severity of the marrow insult.

Factors which were not found to discriminate between the members of groups A and B were etiology (chloramphenicol versus other etiologies), age, therapy (androgens versus no androgens), first symptoms due to anemia versus other types of first symptoms, and first symptoms due to infection versus other types of first symptoms. The hemoglobin concentration was not tested because the effect of prior blood transfusion could not be satisfactorily determined from the records.

One would expect that the patients with a shorter survival (group A) might have lesser concentrations of reticulocytes, neutrophils, and platelets, as well as a greater percentage of nonmyeloid cells in the marrow aspirates, than the patients in group B. It was expected that the interval from the onset of symptoms to the first clinic visit would be shorter in group A than in group B because patients with profound aplasia and rapid clinical deterioration might seek medical care sooner and go through the referral process more quickly than those with a gradual onset of symptoms (and presumably less severe marrow damage). Sex ratios were compared merely because sex helped to discriminate between the two groups.

Most of the differences between groups A and B (Table 1) are of the type that one would expect if the members of group A had experienced a more severe marrow insult. In addition, they suggest that it may be possible to estimate the probability of a patient’s belonging to either survival group on the basis of information available at the first clinic visit. The reason for the greater proportion of males in group A is unclear.

The Prognostic Formula

An attempt to categorize the patients into the actual survival groups from the results of determinations made at the first clinic visit was made by means of discriminant analysis (BMDO4M). The following prognostic formula produced optimal separation of groups A and B under the conditions chosen:

\[
C = -0.01796(B) + 0.01272(S) - 0.00008(OFV) - 0.00359(R) - 0.00002(N) - 0.00018(P) + 0.00046(NM)
\]

where \(C\) is the prognostic index (unitless). \(B\) refers to onset with bleeding; it was graded 0 if bleeding was present at the onset of symptoms and graded 1 if bleeding was absent. \(S\) (sex) was graded 1 if the patient was female, 2 if the patient was male. \(OFV\) is the interval between onset of symptoms and first clinic visit in months. \(R\) is the corrected initial reticulocyte count (per cent). \(N\) is the initial neutrophil concentration (cells/cu mm). \(P\) is the initial platelet concentration (thousands/cu mm). \(NM\) is the per cent nonmyeloid cells in the initial marrow aspirate.
By calculating a prognostic index (C) for every patient on whom the results of more than four of the initial measurements are available (97 individuals), the results in group A could be compared with results in group B (Fig. 3). Although there is overlap, the patients in group A tend to cluster at a higher value of C than those in group B. By inspection, values of C were selected (0.041, 0.033, and 0.0), as indicated by the horizontal dashed lines in Fig. 3, which define groups composed of differing proportions of patients from groups A and B. A summary of Fig. 3 appears in Table 2, in which the vertical columns represent the categories defined by the selected values of C. Of the patients who were classified by prognostic index as A or probably A, the probability of actually being a member of group A ($P_A$) was 30/33 or 0.91. The corresponding probability of actually belonging to group B ($P_B$) was 52/64 or 0.81. Of the members of group A, 30 of 42 were detected (71% sensitivity), and of group B, 52 of 55 were detected (95% sensitivity). Different results would be obtained by choosing other values of C. For example, all 25 of the patients with values greater than 0.041 were actually from group A, but the gain in specificity is obtained at a loss of sensitivity from 71% to 60%.

Cross-validation

In Table 2, the ability of the formula to categorize patients was evaluated using the same patients from whose data the prognostic formula was generated. It seemed possible that the formula was accurate for the particular group of

<table>
<thead>
<tr>
<th>Actual Groups</th>
<th>A</th>
<th>Probably A</th>
<th>Probably B</th>
<th>B</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (Survival ≤ 4 mo)</td>
<td>25</td>
<td>5</td>
<td>12</td>
<td>0</td>
<td>42</td>
</tr>
<tr>
<td>B (Survival &gt; 4 mo)</td>
<td>0</td>
<td>3</td>
<td>37</td>
<td>15</td>
<td>55</td>
</tr>
</tbody>
</table>

$P_A = 0.91$, $P_B = 0.81$
Table 3. Cross-validation. A. Results of Classification by Second Prognostic Formula of Same
74 Patients From Which it was Derived

<table>
<thead>
<tr>
<th>Actual Groups</th>
<th>Classification From Second Prognostic Index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
</tr>
<tr>
<td>A (Survival ≤ 4 mo)</td>
<td>17</td>
</tr>
<tr>
<td>B (Survival &gt; 4 mo)</td>
<td>0</td>
</tr>
</tbody>
</table>

\[ P_A = 0.88 \]
\[ P_B = 0.82 \]

patients which had been used to generate it, but not for other groups. To test this possibility, the entire series was randomly subdivided into groups of 74 and 23 patients. The 23 patients were set aside for testing the predictive value of a new prognostic formula that was now generated solely from data on the other 74. Values of \( C \) from this second prognostic formula were selected that arranged the 74 patients into the categories shown in Table 3. Having established these new criteria, a new prognostic index was calculated for each of the 23 remaining patients. These results are shown in Table 4. The probability of belonging to group A, if so classified by this new prognostic index, was 0.77, and the corresponding probability for group B \( (P_B) \) was 0.9. The sensitivity for group A was 8/9 or 89% and the sensitivity for group B was 9/12 or 75%.

Changes in the Prognostic Index in Individual Patients

If the numerical value of the prognostic index were highly dependent on the phase of the illness when the data for its calculation were obtained, our results might not be expected to apply to other groups of patients with aplastic anemia. To examine that possibility, prognostic indexes were calculated for 26 patients on whom complete sets of data were available at two or more points during each of their illnesses. The median interval between the prognostic indexes was 5.7 mo. The results have been summarized in Table 5, from which the correlation between the prognostic classification determined from the prognostic index at the first clinic visit and the prognostic classification determined from the highest subsequent prognostic index can be judged. For 19 of the 26 patients, the prognostic classification would have remained the same. Two patients would have been assigned greater probabilities of short survival and five individuals lesser probabilities of short survival than they were given at the first clinic visit. In no patient did the prognostic index increase in value so greatly.

Table 4. Cross-validation. B. Results of Classification by Second Prognostic Formula of
Randomly Selected Subgroup not used for its Derivation

<table>
<thead>
<tr>
<th>Actual Groups</th>
<th>Classification From Second Prognostic Index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
</tr>
<tr>
<td>A (Survival ≤ 4 mo)</td>
<td>3</td>
</tr>
<tr>
<td>B (Survival &gt; 4 mo)</td>
<td>0</td>
</tr>
</tbody>
</table>

\[ P_A = 0.77 \]
\[ P_B = 0.9 \]
Table 5. Comparison of Prognostic Indexes Calculated at the First Clinic Visit With Those Calculated From Data Later in the Illness in 26 Patients

<table>
<thead>
<tr>
<th>Classification From Prognostic Index at First Clinic Visit</th>
<th>Classification From Highest Subsequent Prognostic Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (C &gt; 0.041)</td>
<td>A Probably A</td>
</tr>
<tr>
<td>Probably A</td>
<td>1</td>
</tr>
<tr>
<td>(0.041 &gt; C &gt; 0.033)</td>
<td>A</td>
</tr>
<tr>
<td>Probably B</td>
<td>1 12 4</td>
</tr>
<tr>
<td>(0.033 &gt; C &gt; 0.000)</td>
<td>B</td>
</tr>
<tr>
<td>B (0.000 &gt; C)</td>
<td>1 6</td>
</tr>
</tbody>
</table>

that he would have been reclassified as an A. These data suggest that the prognostic index in this subsample, composed primarily of longer survivors, does not change markedly in the course of the disease.

Effect of Androgens on Survival

To evaluate the effects of androgens on survival, patients were divided into three groups: those who had been treated with no androgens, those treated with oxymetholone, and those treated with androgens other than oxymetholone. The details of therapy in these patients have been published. An attempt was made to pair each patient with a patient in another treatment group so that the values of C matched within 0.001. The pairs of groups thus compiled were treated in different ways but were of equal size and had almost identical prognostic indexes. Sixteen pairings could be made in the oxymetholone versus no androgen comparison and 18 pairings in the other androgen versus no androgen comparison. The mean prognostic indexes in the oxymetholone versus no androgen comparison were 0.02448 versus 0.02445 and in the other androgen versus no androgen comparison, 0.03018 versus 0.03003. The survivals of the matched groups are shown in Figs. 4 and 5. Because of the small numbers of
patients, survival was plotted only as long as all were still followed. In both figures, the survival of the androgen-treated patients is not significantly different from the survival of the patients who had received no androgens ($P \geq 0.05$, Wilcoxon signed rank). When patients treated with oxymetholone and with other androgens were combined into a single group and their survival compared with matched subjects not treated with androgens, a similar relationship between survival curves was observed; but again, the difference in survival was not statistically significant ($p = 0.46$, summary Mantel-Haenszel procedure, reference 12).

**DISCUSSION**

The hypothesis that the entire group of patients with aplastic anemia is composed of at least two subpopulations with different mortality rates is supported by the data presented. The patients whose deaths accounted for the initial, steep decline in survival were different from the others in ways that may reflect the severity of the marrow insult. It may seem tempting to speculate that groups A and B represent patients with two different diseases. Although our data do not permit a rejection of that hypothesis, we believe it to be unnecessary if the relation between severity of marrow aplasia and survival is itself biphasic. That is, survival may decline slightly as the severity of aplasia increases until a threshold is reached beyond which survival drops precipitously.

That the patients who die early differ from the others has an important bearing on the interpretation of reports of androgen therapy in aplastic anemia. It is often stated that patients who do not receive androgens for at least 3 mo have not been treated long enough to respond. In many published reports they are excluded as inadequate trials. In these studies, the survival of the patients who, having lived for 3 mo, are considered to have had adequate therapy is superior to the survival of untreated patients in other series from which no patients were excluded. Another interpretation of data of this sort is that, by excluding the patients who die early, only the longer survivors (such as our group B patients) who have a much more favorable prognosis whatever the treatment, remain.
A number of studies have addressed themselves, at least partly, to prognostic factors in aplastic anemia. In most, including ours, some effort has been made to assess the severity of marrow aplasia from the degree of thrombocytopenia, neutropenia, and reticulocytopenia. In our previous report the initial reticulocyte count was lower and the initial per-cent nonmyeloid cells in the marrow was greater in the patients who died than in those who survived, but the initial platelet and neutrophil counts were not significantly different. In other studies, mortality has correlated with the severity of neutropenia, thrombocytopenia, or reticulocytopenia but seldom with all three. Why, then, did our two groups differ in each of these respects, and in several others, when the correlations in other studies have been inconsistent? The answer may lie in the way prognostic factors were sought. Often, as in our previous paper, the blood counts in patients who ultimately died were contrasted with the counts in patients who survived. When this approach is taken, no interval is specified between the predictive event (the blood count, for example) and the predicted event (mortality); one may be asking one observation to predict another that occurs a long time afterwards. In this study, the predictive and predicted events were separated by a maximum interval of 4 mo.

The patients who lived for 4 mo or less appeared to be identifiable from the results of simple determinations performed at the first clinic visit with the aid of the prognostic index. Although supported by the results of cross-validation, this means of estimating prognosis must be tested in other groups of patients. If shown to be valid, the prognostic index might permit the clinician to estimate how likely it is that his patient with aplastic anemia will survive for at least 4 mo. A patient who is almost certain to be a short survivor on the basis of the prognostic index would seem a reasonable candidate for attempted marrow transplantation. Similarly, in designing prospective therapeutic trials, the prognostic index would enable one to assess whether groups of patients treated in different ways were similar with regard to several of the factors that influence survival.

Several objections may be made to the types of data from which the prognostic index is computed. The interval between the onset of symptoms and first clinic visit lacks precision because it depends on the accuracy of the patient's recall and on the complex factors that determine when a patient is referred. However, the relative weight afforded this variable is slight, as reflected by its coefficient of 0.0008 in the formula.

In addition, some of the variables are interdependent. For example, hemorrhagic symptoms must surely be related to severe thrombocytopenia. Nonetheless, the results of the multidiscriminant analysis showed that each variable included in the formula was able, independently of the others, to help distinguish between the members of the two prognostic groups.

At least two explanations may be offered for the complexity of the formula and for the inclusion of apparently interdependent variables. If the major determinant of survival is the extent of marrow hypoplasia, most of the hematologic data from which the prognostic index is calculated reflect it very indirectly. The neutrophil concentration, for instance, is affected by factors other than changes in production rate. Therefore, reliance on fewer determinations...
might increase the chance of error. On the other hand, although two variables, such as platelet count and hemorrhagic symptoms, may be somewhat interdependent, they are not entirely, so that a patient who bleeds with a platelet count of 40,000/cu mm may have a shorter survival than another patient who does not bleed with the same count.

As an example of the potential application of a prognostic index to the evaluation of clinical trials, the effects of androgen therapy on survival were considered. The major difficulty in analyzing the results of uncontrolled trials of androgen therapy in aplastic anemia is finding an untreated group which is similar enough to the treated group to permit comparisons of survival. Since our series contained patients who had not received androgens and since each patient's risk of short survival could be estimated from his prognostic index, it appeared that a patient who had received androgens could be matched with another patient who had not, purely on the basis of their having the same prognostic index. If enough matches could be made, one could compile two groups of patients of equal size and of apparently equal prognosis, but who had been treated in different ways. The beneficial effects of treatment could then be assessed by comparing their survivals. The survival curves obtained (Figs. 4 and 5) did not show any beneficial influence of therapy with either oxymetholone or with other androgens on survival.

The interpretation of these data is, of course, very difficult. They do not answer the question of whether androgens are beneficial in aplastic anemia. First, the groups in which survival was contrasted were small. Second, even though the intent was to compare survivals of prognostically similar groups, there may well have been other ways in which the groups were different that accounted for the results. Third, the doses and durations of treatment were not standardized. Fourth, survival may not be an adequate way to judge whether the patients were benefited. Finally, we have seen patients in this series in whom androgens clearly produced benefit. In these patients, blood counts rose and symptoms decreased when androgens were given, and relapse occurred when androgens were withdrawn. In two patients, this sequence occurred repeatedly. The values of the prognostic index in these two individuals, 0.00199 and -0.00788, suggest that they were very likely to be longer survivors even without therapy. It seems entirely possible that the androgen responders have a separate disease.

REFERENCES

The prognosis in aplastic anemia

RE Lynch, DM Williams, JC Reading and GE Cartwright