Treatment of acute myeloid leukemia: are we making progress?

Alan K. Burnett

1 School of Medicine, Cardiff University, Cardiff, United Kingdom

Introduction

Being predominantly a disease of older people, the therapeutic strategy offered for AML is determined by assessment of the patient’s age and general fitness level. For those thought to be suitable for, or who might benefit from (not necessarily the same patients), the standard of care for over 3 decades has been the combination of daunorubicin given for 3 days combined with cytarabine (Ara-C) given over 7 days, the so-called “3 + 7” schedule. The majority of younger patients who achieve a CR will do so with one course; however, a second course will convert half of the remainder so that, overall, 70%-75% will enter CR.1 Based on the landmark Cancer and Leukemia Group-B (CALGB) study,2 the standard of care for consolidation was 4 courses of high-dose IV Ara-C 3 g/m² given every 12 hours on days 1, 3, and 5, which was demonstrated to be more effective than either a dose of 100 or 400 mg/m² by continuous infusion for 5 days followed by 4 courses of monthly maintenance. If a suitable donor is available (formerly limited to a matched sibling, but now a matched unrelated donor as well), an allogeneic transplantation will be considered as alternative consolidation.3,4 Delivering this approach to younger patients will provide a 40%-45% chance of cure. This represents a steady improvement over the past 40 years (Figure 1), although much of this can be attributed to improvements in supportive care, much of which has been learned from the transplantation experience. There are, however, several unresolved issues around this “conventional” approach, such as alternative anthracyclines or nucleoside analogs in induction, dose, and number of courses of consolidation, and which patients should receive an allograft in first remission.

With a few subgroups as exceptions, such as younger patients with more favorable genetic disease, improvement in the treatment of acute myeloid leukemia has been slow. There is a possibility that improving the quality of remission can reduce the risk of relapse. Escalation of daunorubicin dose, addition of Ab-directed chemotherapy, and alternative nucleoside analogs in induction may displace the longstanding standard of “3 + 7” daunorubicin + cytarabine (Ara-C) as induction, and several prognostic factors are emerging that enable a more personalized approach to postinduction treatment, in particular, which patients should be offered allogeneic transplantation in first remission. In addition to providing prognostic information, molecular characterization provides potential therapeutic targets and, in some cases, an opportunity to more precisely monitor residual disease. With few exceptions, the predictive value of prognostic factors (ie, what therapy to adopt) has yet to be established. A major challenge is the treatment of older patients with acute myeloid leukemia (AML), who represent the majority of patients with this disease. Only about half of older AML patients will enter complete remission (CR) with conventional chemotherapy and, of these, most will relapse within 2 years. Little impact has been made on these dismal outcomes over the past 3 decades, and new treatments and approaches to trial design are required. Another population of concern is older AML patients who are not considered to be fit for an intensive approach based on concerns about their ability to withstand the consequences of treatment. This group is not easy to define objectively, but age represents a useful surrogate because it is associated with more chemoresistant disease and medical comorbidity. Older patients represent a therapeutic challenge, but several new treatments may offer some potential to improve their situation.

Options for improvement

Several alternatives to 3 + 7 have been assessed in multiple studies, with little convincing evidence that they will change the standard of care. In particular, there is no consistent evidence to support the addition of a conventional therapeutic as a third drug or escalation of the dose of Ara-C.3,4 However, there is recent interest in daunorubicin dose escalation or intensification and the role of Ab-directed chemotherapy as a way of delivering more treatment without significant extra toxicity. It may well be possible to improve the quality of remission, which would be reflected in a reduction of relapse risk, rather than just the remission rate. A substantial number of prognostic factors (Table 1) have been identified, each of which in many cases can be shown to independently indicate a favorable or unfavorable average outcome and thus offer the potential to direct therapy appropriately. Not all of these factors are molecular or
genetic, and traditional factors such as initial response, presenting WBC count, and secondary disease, still have value. However, it is hoped that the substantial molecular heterogeneity of AML currently being revealed will bear therapeutic fruit by providing “druggable” targets, rather than simply adding to the already extensive list of prognostic factors.

Challengers to standard of care

**Anthracycline dose**

Whether daunorubicin should remain as standard of care has been questioned in randomized comparisons with idarubicin or mitoxantrone in several randomized trials, without convincing evidence for substantial differences in survival. Four recent studies evaluated different ways of intensifying daunorubicin administration. The E1900 trial, which was undertaken by the Eastern Cooperative Oncology Group (ECOG) in 657 younger patients, compared a daunorubicin dose of 90 mg/m² versus 45 mg/m² in a 3 + 7 schedule for the first induction course. The higher dose achieved a significantly higher overall remission rate, with more patients in CR after the first course and a better overall survival (23.7 vs 15.7 months). A similar approach was taken by the Dutch-Belgian Hemato-Oncology Cooperative Group and the Swiss Group for Clinical Cancer Research (HOVON-SAKK) in older patients. In that study, the overall remission rate was similarly improved, but overall survival was not, except in the case of patients in the 60- to 65-year subgroup (38% vs 23%). These data received further support from a Korean trial in younger patients showing that the higher dose of daunorubicin (90 mg/m² vs 45 mg/m² by continuous infusion over 3 days) also resulted in a superior remission rate and overall survival.

There has been discussion for a long time about dose equivalence between daunorubicin and alternatives such as idarubicin. A study from the Acute Leukemia French Association (ALFA) group showed that daunorubicin 80 mg/m² on days 1-3 versus idarubicin 12 mg/m² for either 3 or 4 consecutive days in patients 50-70 years of age produced an equivalent overall survival. A recent Japanese trial in 1057 younger patients tested intensification, not by dose escalation, but by administration of 50 mg/d for 5 days compared with the traditional 3 days (ie, a total first-course exposure of 250 vs 150 mg against idarubicin for 3 days). The remission rates and survival were similar. Whether these studies justify a higher daunorubicin dose as the standard of care is not clear. All studies confirm, at least with the follow-up available, that cardiotoxicity is not a concern in adults. The benefits appear to differ in the cytogenetic risk subgroups, being less clear in the favorable and....
patients. In a second, more recent study,28 the addition of cladribine demonstrated, possibly because there was an insufficient number of patients, the benefit of being included in the study. The addition of cladribine significantly increased the proportion of patients who achieved CR with one consolidation course, but an improvement in survival was not seen.16 Several trials have been reported where augmented daunorubicin was created that gained regulatory approval for the treatment of older patients in relapse who were considered unsuitable for intensive therapy.16 Several trials have been reported recently; however, failure to show benefit and some concern about an excess of induction death in the SWOG-106 trial, which augmented daunorubicin + Ara-C with a single dose of GO (6 mg/m²)17 and which was presented for regulatory approval, led to withdrawal of the drug from the US market. Three other European trials involving nearly 3000 randomized patients have shown a survival benefit when combined, at lower daily dose levels, with induction chemotherapy. In the United Kingdom Medical Research Council/National Cancer Research Institute (9MRC/NCRI) trials in younger and older patients, 3 mg/m² was safe and effective,18,19 whereas the French ALFA trial used 3 mg/m² in a fractionated schedule.20 All showed an improved survival in favorable and intermediate-risk, but not in poor-risk patients. The Italian Group for Haematological Diseases in Adults (GIMEMA) AML17 trial compared pretreatment with GO at a dose of 6 mg/m² before initiating chemotherapy. However, when used in this way, there was no overall benefit.21 In addition, GO is a very effective agent in acute promyelocytic leukemia and provided a useful option for relapsed disease,22 so perhaps it will reemerge as a therapeutic option.23

**Ab-directed chemotherapy**

Augmentation of standard chemotherapy in relapsed or refractory patients with naked CD33-directed Ab therapy failed.15 However, when the Ab was conjugated to the powerful intercalating agent calicheamicin, an effective agent, gemtuzumab ozogamicin (GO; Mylotarg), was created that gained regulatory approval for the treatment of older patients in relapse who were considered unsuitable for intensive therapy.16 Several trials have been reported recently; however, failure to show benefit and some concern about an excess of induction death in the SWOG-106 trial, which augmented daunorubicin + Ara-C with a single dose of GO (6 mg/m²)17 and which was presented for regulatory approval, led to withdrawal of the drug from the US market. Three other European trials involving nearly 3000 randomized patients have shown a survival benefit when combined, at lower daily dose levels, with induction chemotherapy. In the United Kingdom Medical Research Council/National Cancer Research Institute (9MRC/NCRI) trials in younger and older patients, 3 mg/m² was safe and effective,18,19 whereas the French ALFA trial used 3 mg/m² in a fractionated schedule.20 All showed an improved survival in favorable and intermediate-risk, but not in poor-risk patients. The Italian Group for Haematological Diseases in Adults (GIMEMA) AML17 trial compared pretreatment with GO at a dose of 6 mg/m² before initiating chemotherapy. However, when used in this way, there was no overall benefit.21 In addition, GO is a very effective agent in acute promyelocytic leukemia and provided a useful option for relapsed disease,22 so perhaps it will reemerge as a therapeutic option.23

**Alternative nucleoside analogs in induction**

Escalation of the Ara-C dose has been unconvincing, but there has been recent investigation into alternative nucleoside analogs such as cladribine,24 fludarabine,25 and clofarabine.26,27 The Polish Adult Leukaemia Group (PALG) compared cladribine with cytarabine in 2 randomized trials. In the first trial,24 the inclusion of cladribine increased the proportion of patients who achieved CR with one course significantly, but an improvement in survival was not demonstrated, possibly because there was an insufficient number of patients. In a second, more recent study,28 the addition of cladribine or fludarabine to 3 + 7 cytarabine was compared with 3 + 7 alone in a 3-arm study (daunorubicin/Ara-C [DA]/cladribine vs DA/ fludarabine vs DA alone) in patients less than 60 years of age. The inclusion of cladribine, but not fludarabine, improved the remission rate and overall survival. In a subgroup analysis, the addition of either nucleoside appeared to be beneficial in patients with an adverse karyotype. The MRC AML15 trial compared FLAG-Ida (fludarabine/Ara-C-CSF/idarubicin) with DA or DAE (DA + etoposide) in induction in younger patients. There was no difference in remission rate or overall survival, but FLAG-Ida reduced the risk of relapse significantly.25 However, this reduction in risk was offset by the consequences of increased myelosuppression, which also limited compliance with postinduction therapy. The overall survival of the selected patients who did receive the planned consolidation was 67%. The United Kingdom NCRI AML16 trial compared clofarabine at a lower dose level of 20 mg/m² against Ara-C, both in combination with standard-dose daunorubicin in older patients. This was based on encouraging data when clofarabine was used as monotherapy, particularly in patients with adverse cytogenetics. However, there was no difference in response or overall survival in any risk subgroup.

**Postinduction therapy**

The prognosis for patients who have completed induction course one, even though they have not achieved CR, can be estimated by several molecular and genetic factors (discussed below) which are increasingly being considered when deciding postremission strategy. Important information can be derived at this point in relation to treatment response to the first course. If the patient has achieved at least a partial remission (ie, BM blast reduction to <15%),29 or blast eradication on a day 15 BM or is in a CR with full count recovery, he or she has more chemosensitive disease and a lower risk of subsequent relapse from remission. There remains uncertainty about the value of complete BM response with incomplete count recovery, the so-called “CRp/Cri,” which was a response criterion brought to life in the assessment of the Mylotarg responses. Although there may be caveats about when in relation to the BM sampling the recovery definition was made and the level of the platelet count and transfusion dependence, it is generally concluded that CRp/Cri is associated with an inferior survival compared with confirmed CR.30,31 The ability to move forward from a morphological definition to one based on cytogenetic or molecular or immunophenotypic assessment of minimal residual disease may prove to be more useful.

Whereas high-dose Ara-C at a dose of 3 g/m² is the standard of care compared with an intermediate 400 mg/m² dose, no further refinement of dose or the optimal number of courses has been described. The MRC15 trial compared a 1.5 g/m² dose against the 3 g/m² dose and found no difference in major outcome end points. In addition,
there was no benefit in giving more than 2 courses in the context of having given 2 induction courses. However, compared with an amsacrine/etoposide/mitoxantrone–based consolidation approach, both high-dose Ara-C dose levels were inferior in patients with poor-risk cytogenetics.25 There is continuous debate about the role of stem cell transplantation in first remission. Large studies from collaborative groups show no significant overall survival benefit of myeloablative transplantation if the recipient’s age exceeds 35-40 years,3,4 based on a donor versus no donor method of analysis. In meta-analyses based on a heterogeneous collection of studies with risk defined in a nonuniform way (but usually cytogenetically based), the availability of a donor was associated with an improved relapse-free survival and, in some cases, overall survival for both intermediate- and poor-risk patients. There is an increasing awareness that a donor versus no donor analysis as a surrogate for a randomized comparison must be viewed with care. For example, in some studies, it is assumed that if a sibling donor is identified, a transplantation has actually been undertaken and that it has been undertaken in first remission, but these assumptions may be incorrect. For example, patients who do not have a sibling donor, but do have an unrelated donor, will be in the “no donor” group and, in the analysis, the transplantation may or may not have been done in first remission. An alternative is the Mantel-Byar method,32 in which all patients start on the “chemo” arm and when the transplantation is delivered, they populate the “transplantation” arm. To compensate for the selection of patients who survive long enough and are in good condition, a time adjustment on the “chemo” curve can be made. While not an ideal alternative, having a tendency to favor transplantation, this sort of analysis can allow the inclusion of nonsibling donors, only counts those in whom the transplantation was given, and can be applied to assess the impact of transplantation in CR1 only. An additional issue is the appropriate end point. If event-free or relapse-free survival is the chosen end point, in most studies showing a benefit from transplantation, patients on chemotherapy who are salvaged after relapse are excluded. For example, in a recent analysis of 1271 patients in the United Kingdom database who relapsed without being transplanted in CR1, 32%, 17%, and 7% of favorable-, intermediate-, and poor-risk patients, respectively, survived.33 Therefore, a Mantel-Byar analysis with overall survival as the end point appears to be the appropriate methodology. This makes the benefit for intermediate risk based on cytogenetics less clear. However, by more precisely defining risk using age, cytogenetics, response to induction course 1, presenting WBC count, and de novo or secondary disease, approximately 20%-25% of intermediate-risk patients have a higher than average risk and, based on a Mantel-Byar analysis, will benefit from a myeloablative transplantation in CR1.34 Nonmyeloablative conditioning (reduced intensity conditioning) enables older patients to access stem cell transplantation and the associated GVL effect. This is clearly feasible, but there are few comparative studies clarifying the overall benefit or describing which patient subgroups benefit.

Prognostic factors and clinical application
Cytogenetics has traditionally been used in younger patients to influence the choice of postinduction therapy because it effectively predicts the average relapse risk. The usual therapeutic option under consideration is who should, or should not, be put forward for allotransplantation. Patients in the core binding factor (CBF) subgroup have a favorable outlook and are not considered to need a first CR transplantation; these patients may have their risk reduced if they receive GO in induction. There is an increased risk of relapse for the 20%-30% of patients who have a cKIT mutation, and the molecular targets are capable of predicting impending relapse by quantitative RT-PCR with approximately 3 months’ notice based on BM transcript load in the postinduction remission BM. Experience with cKIT inhibitors is accumulating. Bi-allelic CEBPa and NPM1 mutations without an associated FLT3 mutation have a similar prognosis to the CBF leukemias, so they do not require a transplantation approach. Conversely, patients with adverse cytogenetics (including chromosome 5 and 7 abnormalities, inv3, and complex and monosomal karyotypes) will relapse rapidly. Morphological response to induction course 1 (blast reduction in BM to < 15%) is an independent adverse feature even if remission is achieved subsequently. Because there is no satisfactory conventional chemotherapeutic approach for such patients, allograft is recommended, but will only cure 30%-40% of patients, and some will relapse before the transplantation can be delivered. A large volume of molecular data have emerged showing a better or poorer prognosis with the main aim being to determine the intermediate-risk group.35 Of the NPM1 mutants, which occur in 50% of normal karyotypes, approximately 40% are associated with a FLT3 mutation that balances out the favorable effect of NPM1, although the allelic ratio of the FLT3 mutation may influence this. The impact of FLT3 mutations is complex and is influenced by the association with NPM1 whether it occurs alone and by the allelic ratio. It is likely that the prognostic impact of each mutation will be modified by its association with other mutations.

It is important to be aware that “prognostic” does not mean “predictive,” in that the latter means that a validated therapeutic option has been identified. In most cases, there is a lack of robust data pointing to a particular validated therapeutic option. Allogeneic transplantation is frequently considered for FLT3 mutations, although the results may be disappointing and the evidence base for benefit is not settled,36,37 particularly when the allelic ratio and association (or not) with NPM1 is considered. The association of a cKIT mutation with the CBF group, TET2, DMT3a, IDH1/2, and ASXL-I have all emerged recently, and a consistent prognostic picture has yet to emerge. The picture becomes even more complicated by the prognostic implication of mutations occurring alone or in combination.34 It will require large, extensively characterized clinical databases to clarify this information.

The importance of the molecular heterogeneity could be in identifying targets for therapeutics or for minimal residual disease monitoring. Most attention has so far focused on FLT3 as a target, but there are no randomized data to show a benefit of FLT3 inhibition, although such information is imminent. An important component of this approach is to have associated pharmacodynamic markers. There is preliminary information that it is only when the target is inhibited that the risk of relapse is reduced.38 All-trans-retinoic acid has a well-established role in acute promyelocytic leukemia, and has been evaluated in nonacute promyelocytic leukemia patients based on the concept that it can reduce the half-life on BCL2, but the results are contradictory. Whether the NPM1-mutated subset in particular benefit, as has been claimed, is open to question.

The challenge of the older patient
Collaborative group studies have an undistinguished record in improving the outcome of older patients. Although age is an independent adverse prognostic factor, it also serves as a surrogate for other negative factors such as comorbidities and biologically more resistant disease. An illustration of the problem is the serial outcome for patients treated with conventional chemotherapy in United Kingdom trials over the past 40 years (Figure 2). Up to 60% of patients enter remission depending on the inclusion criteria, but
most will relapse. There has been little success in altering this with consolidation and, at best, maintenance has a modest effect. Recent evidence suggests that, at least for some patients, there may be a daunorubicin dose effect, and 2 of 3 European randomized trials suggest a benefit for augmenting induction with GO. A small minority of selected patients may benefit from a reduced intensity allograft. Cytogenetics is also a powerful prognostic factor in this patient group, with a smaller proportion of favorable-risk and a higher proportion of poor-risk subtypes. There is less molecular information and, because of the overall poorer prognosis, the impact is less clear. For example, FLT3 mutations still confer a negative impact, whereas NPM1 mutations are less predictive of an improved outcome. In addition, there are a substantial proportion of older patients who do not receive conventional chemotherapy. These patients are sometimes referred to “the older unfit” patients and have traditionally been offered supportive care. How these patients are defined, and whether they would be better off with conventional chemotherapy, is one of the present dilemmas. Population-based information from Sweden showed that the outcome was superior in geographic regions where intensive treatment was given compared with regions where it was not. However, there could be other variables governing the choice of treatment approach. In an effort to clarify this issue, a United Kingdom trial (AML14) intended to randomize patients for whom there was uncertainty to an intensive or nonintensive approach. However, in a recruitment of >1600 patients, only 8 were randomized. Multifactorial risk scores that can more accurately characterize risk have been developed in older patients treated either intensively or nonintensively, but this does not clarify which approach is better for which patient. There is also a category of patients who may be fit, but are thought not likely to benefit from an intensive approach, such as older patients with secondary disease or adverse cytogenetics, who are candidates for experimental treatments.

Several new drugs have been tested in this patient group in recent years, and although capable of improving response, none has achieved regulatory approval, but this continues to be an active area of study. One of the challenges is how to improve matters in a timely manner. To that end, the traditional large randomized trial may not be appropriate, so new designs are emerging, one of which is a “pick a winner” concept in which several novel treatments are compared with standard of care (low-dose Ara-C in the United Kingdom). The aim is to screen for options that could make a clinically important difference defined on a preliminary assessment against response rate, and to only continue within the trial the assessment of those that show promise; for example, double the remission rate as an early surrogate for survival. New treatments can be introduced as the trial progresses, but contemporaneous randomization remains an essential requirement. Whereas carrying some risk of progressing a treatment that will not improve survival, there is reasonable reliability in eliminating treatments that fail in the randomized setting.

Prospects for the future

It could be argued that there is limited further mileage in conventional chemotherapeutic agents. The burgeoning molecular information could be valuable in several ways apart from shedding light on leukemic mechanisms and biology. Like others, AML doctors always hope for the “AML Glivec,” but with such heterogeneity, this is unlikely except possibly in well-defined subsets. The ability to “personalize” treatment based on more accurate estimate of an individual patient’s risk has a numbers of hurdles to jump. First, new prognostic subsets must be based on statistically robust data that are applicable to contemporary treatment. They have to be validated on a different dataset. Some may be unstable as new treatments emerge. For example, it is not clear how prognostic a cKIT mutation is in the CBF subset that has received GO in induction. Minimal residual disease assessment, by molecular or immunophenotypic methods, is now capable of reliably predicting for each individual an accurate risk assessment. However, the implication is that therapeutic intervention at the time of minimal residual disease detection is a superior intervention at morphological relapse and this has yet to be established.

Disclosure

Conflict-of-interest disclosure: The author declares no competing financial interests. Off-label drug use: GO for the treatment of AML.

Correspondence

Alan K. Burnett, School of Medicine, Cardiff University, Heath Park, Cardiff CF14 4YU, United Kingdom; Phone: 44-2920-742375; Fax: 44-2920-744655; e-mail: BurnettAK@Cardiff.ac.uk.

References
