Leukemia (2000) 14, 476–479

www.nature.com/leu

Treatment of relapsed and refractory acute myelogenous leukemia

EH Estey

University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Box 61, Houston, TX 77030, USA

Evidence suggests that the salvage therapy utilized for relapsed and refractory acute myelogenous leukemia (AML) should differ based on the duration of a patient's complete remission (CR), the principal predictor of outcome. While standard regimens have produced higher CR rates than investigational regimens, these rates have not translated into improved survival in patients with initial remission durations of 1 year. Accordingly, there is no need to give standard regimens to these patients who rather should receive investigational therapy once relapse is discovered. In contrast, in patients with initial remission durations of 1–2 years, standard regimens do increase survival compared to investigational regimens. A somewhat artificial distinction has been placed between phase I and phase II studies. The agents to be studied in phase II trials are many, but the patients are limited, so we need to be more innovative in our trial designs. One such proposal, utilizing a Bayesian selection design which calls for randomizing a small number of patients among several investigational treatments, will be discussed. *Leukemia* (2000) **14,** 476–479.
Keywords: salvage therapy; AML

Introduction

In medicine, prescription of therapy presupposes knowledge of prognosis. In relapsed or refractory acute myelogenous leukemia (AML), the prognosis in question is that expected following administration of standard therapy, such as cytosine arabinoside (ara-C), in standard or high dose, or mitoxantrone plus etoposide, to patients given similar treatment at initial diagnosis. It is well known that the principal predictor of outcome following such standard salvage treatment is the duration of the initial complete remission (CR).[1]

It is perhaps less well-known that as duration of first CR decreases, CR rate, or disease-free survival, following ara-C-containing salvage therapy decreases continuously. Thus, an initial CR of 6 months is associated with a better outcome than an initial CR of 5 months, which in turn is associated with a better outcome than an initial CR of 4 months, etc. Accordingly, it makes little sense to divide patients into better or worse prognostic groups based on first CR duration considered as a binary variable (eg, better if first CR 6 months or 12 months, worse otherwise). Nonetheless, this practice is often followed for reasons of convenience.

Adhering to this convention, patients receiving an ara-C-containing regimen after an initial CR duration of 24 months will on average have a CR rate of 60%.[2] If the initial CR duration was 12–24 months, the CR rate will be approximately 40%, while shorter first CRs or failure to achieve a CR with initial therapy are associated with CR rates of 10–20%.[2] Cytogenetics, age, and failure of prior salvage therapy are also prognostic. In particular, the probability of CR is 1% following administration of standard therapy to patients whose initial remission lasted 1 year and whose AML failed to respond to an initial reinduction attempt.[2]

Correspondence: EH Estey
Received 20 July 1999; accepted 27 July 1999

These data should inform treatment decisions. For example, administration of standard salvage therapy to patients in whom the probability of success is 1% should be discouraged. At MD Anderson Cancer Center (MDACC), such patients receive phase I agents. Although the likelihood of success in phase I studies may be similarly low, execution of such studies is important and the patients defined above are an appropriate cohort. Later, we discuss the means to possibly improve the efficiency and conduct of phase I studies.

Role of investigational therapies

Should investigational therapy be given to patients who are about to receive an initial salvage attempt? At MDACC, such patients have received high-dose ara-C-containing (HDAC) regimens or investigational regimens largely at the discretion of the attending physician. Table 1 indicates that the HDAC regimens have produced higher CR rates in both patients with initial CR durations 12 months and in patients with initial CR durations of 12–24 months. Patients with acute promyelocytic leukemia and patients who received an allogeneic or autologous transplant either to induce CR or post such CR were not included in these analyses. The data suggest that standard regimens are generally preferable to investigational agents. This interpretation relies on the assumption that there is a relationship between CR and subsequent survival. Empirical proof of this relationship in newly diagnosed patients was provided by Freireich *et al*[3] and, more recently, by Estey *et al*.[4]

Is there a similar relationship in patients given first salvage therapy? Figure 1 compares survival in patients with initial remissions 12 months (or with no initial remissions) according to whether they received HDAC-containing therapy or investigational regimens. There is no survival advantage for the HDAC group. This reflects two phenomena: first, the brevity of remissions induced by salvage therapy, as noted previously,[5] and, second, a higher induction mortality rate following administration of HDAC (46/175 *vs* 24/168, $P = 0.006$). Assuming that the investigational and HDAC groups were similar in other respects, the data argue for initial use of investigational therapy to induce CR in AML recurrent after an initial CR of 1 year or not in CR after initial induction therapy. While a strategy of inducing a second CR with conventional therapy and then administering investigational postremission therapy might seem preferable, the higher initial death rate following use of such therapy argues against this approach.

In contrast to the findings in patients with very short or no initial CRs, there is a clear survival advantage for patients given HDAC rather than investigational agents after an initial CR of 1–2 years ($P = 0.016$, Figure 2). As a corollary, HDAC is not associated with a higher induction mortality rate. Nonetheless, the results even with HDAC are suboptimal. While a case could be thus made for use of investigational therapy in such patients, the data argue for use of HDAC sometime during treatment, for example to produce an initial CR.

At MDACC we will soon begin a study randomly assigning

Table 1 Response to first salvage therapy by type of regimen

Duration first CR (months)	*HDAC-containing regimens*		*Investigational regimens*		*P value*
	Patients	*CR (%)*	*Patients*[a]	*CR (%)*	
12 or no CR after initial therapy	175	20 (11)	168	6 (4)	0.006
12–24	43	22 (51)	27	2 (7)	0.001
24	25	14 (56)	5	1 (20)	0.15

[a]Principal regimens are as follows: Regimen (No. CR/No. patients). Topotecan (0/18); topotecan + etoposide (1/17); topotecan + CTX (0/15); CTX (q12 h) + vincristine + adriamycin + dexamethasone (2/14); 2 CDA (0/12); 9 aminocamptothecin (0/10); tallimustine (1/9); CMA-676 (3/8); deoxyazacytidine (DAC) (0/8); CI973 (0/7); BCH (0/7); thalidomide (1/6); trimetrexate (0/6); gemcitabine (0/4); PZDH (0/3); cyclophosphamide (CTX) + etoposide + carboplatin (0/3); carboplatin + gemicitabine (0/3); taxol (0/2); merbarone (0/2).

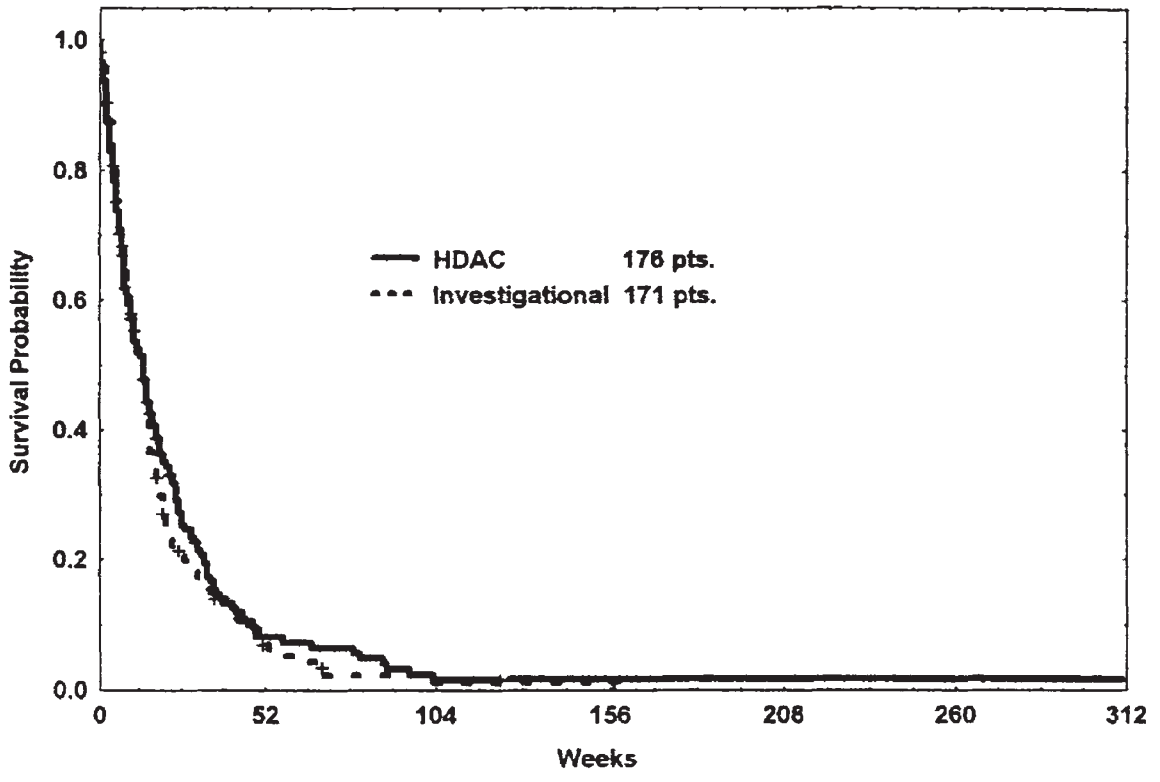

Figure 1 Survival probability by type of salvage regimen (HDAC-based *vs* investigational) in patients whose initial CR was 1 year.

patients with AML in initial relapse after a remission of 12–24 months to receive the investigational agent BCH or (the more standard) twice daily fludarabine plus ara-C (BID FA). Patients whose disease fails to respond to BCH will receive BID FA on the second course and vice versa. Our objective is to determine which strategy, standard followed by investigational or investigational followed by standard, is superior. Obviously, the results of such a study, and the results presented in this and the preceding paragraph are a function of the investigational agent(s) selected. The dose of BCH is established, and the drug has produced CR in multiply relapsed disease. In contrast, there would be obvious reluctance to use a phase I agent in this setting.

The discussion to date has ignored the role of allogeneic transplantation. In particular, should this modality, as usually practiced, be considered the best option in recurrent AML? Comparison of results following allogeneic transplant with those following chemotherapy would have to account for the possible influence of the prognostic factors described in the first paragraph as well as the possible influence of selection bias. Such an analysis remains to be done. It does seem fair to state that the results with current chemotherapy and with current transplant procedures leave much to be desired. Hence, it may be more relevant to ask whether investigational approaches can improve either modality, rather than to ask whether transplantation is superior to chemotherapy or vice versa.

A large number of investigational agents are available (Table 2). Conventionally, these drugs are studied first in phase I and then phase II. The majority of phase I trials still employ the '3 + 3' design. However, numerous computer simulations have demonstrated that this algorithm is inferior to the 'continuous reassessment method' (CRM),[6] or modifications of the CRM,[7] in identifying the dose that produces that level of toxicity thought to be a precondition of efficacy. In particular, the 3 + 3 design frequently selects a dose that is too low. Although a detailed comparison of the 3 + 3 and the CRM is beyond the scope of this paper, the superiority of the

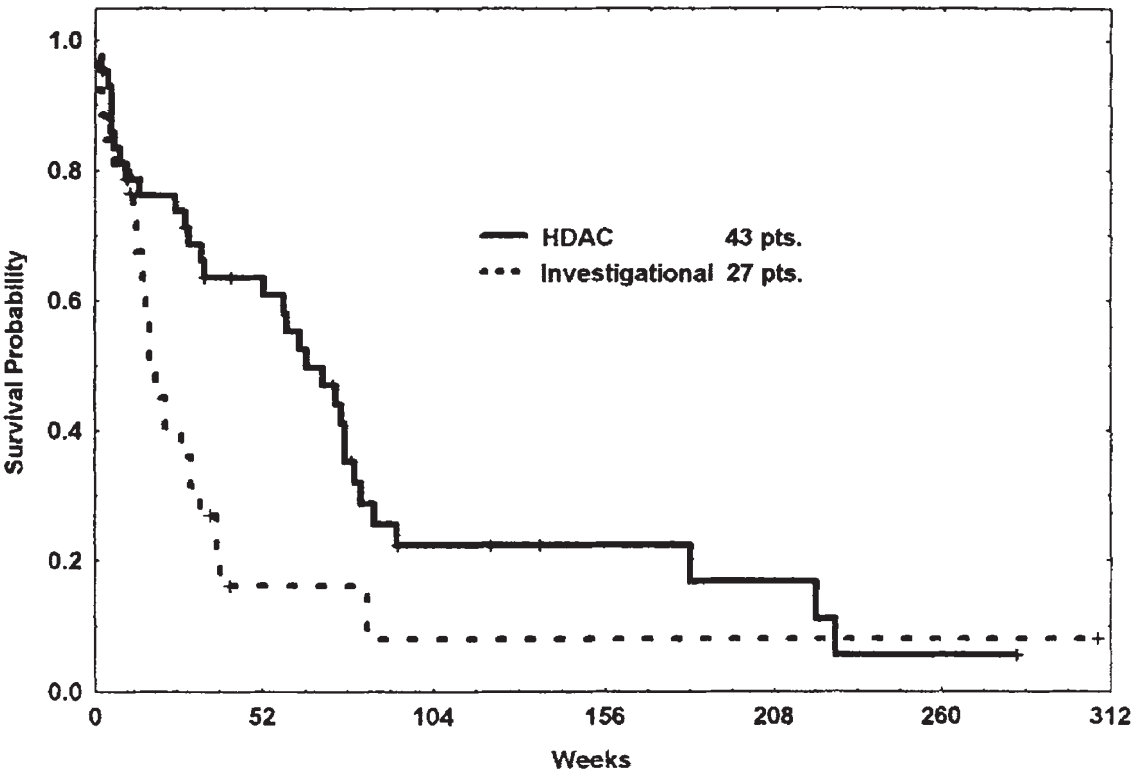

Figure 2 Survival probability by type of salvage regimen (HDAC-based *vs* investigational) in patients whose initial CR was 1–2 years.

CRM to some extent reflects its use of information from every patient. In contrast, if, say, one-third of patients have toxicity at dose X, the 3 + 3 calls for entry of another three patients at that dose regardless of results at lower doses.

The distinction between phase I and phase II is itself artificial. The ethical underpinning of a phase I study is the expectation of response. Hence, both toxicity and response should be monitored during such a study. Thall and Russell[8] and Thall *et al*[9] have described a hybrid phase I–II design that selects a dose for the next cohort based on both response and toxicity data from previous cohorts. Similarly, toxicity should be monitored in phase II since the data on which the phase II dose is selected are often minimal. Typically, such monitoring is done only informally. Thall *et al*[10] and Thall and Sung[11] have described Bayesian designs that permit simultaneous monitoring of efficacy and toxicity in a phase II trial.

Selecting phase II agents for study

An issue that has received limited attention is selection of agents for phase II testing. Typically, a large number of agents are available (Table 2), but the number of patients is limited. Furthermore, preclinical data do not allow prediction as to which agent is best. Indeed, experience with agents such as 2-CdA in hairy cell leukemia or with alpha interferon in chronic myelogenous leukemia suggests that the explanation for clinical success can remain elusive many years after the initial clinical observation. Given this scenario, Thall and Estey[12] have proposed a Bayesian selection design which calls for randomizing a small number of patients among several investigational treatments, early stopping if there is high probability that a specified success rate is unlikely, and selecting therapies that are not stopped early for subsequent large-scale phase II study.

Unlike other adaptive designs, the selection design permits monitoring on a patient-by-patient basis. Although, given the small number of patients (eg, a maximum of 20 per treatment

Table 2 Investigational therapies for relapsed/refractory AML under study or planned at MDACC

Agent	*Comment*
BCH	Nucleoside analogue
Liposomal annamycin	Anthracycline, possibly unaffected by MDR
Dolastatin	
DX8951	Reactive with topoisomerase I (like topotecan)
UCN01	Cell cycle checkpoint modulator (combined with ara-C)
PS341	Proteosome inhibitor
SU516	Angiogenesis inhibitor
Anti-CD52 monoclonal antibody	15% of newly diagnosed AML is CD52 positive
Anti-CD33 monoclonal antibody	In particular, combined with an anthracycline (eg, calicheamycin) or other toxin (eg, gelanin)
Anti-GM-CSF receptor monoclonal antibody	Combined with diphtheria toxin
Farnesyltransferase (RAS) inhibitors	May not require RAS mutation for efficacy
Fludarabine + ara-C + haploidentical donor lymphocyte infusions	Goal is to induce graft-versus-leukemia effect

arm), the false negative rates associated with the design can approach 30%, this rate should be compared with the effective false negative rate of, for example, 75% that results if one of four agents is arbitrarily selected for large-scale phase II study in the absence of compelling data justifying the selection. Use of such informal processes is not infrequent.

A final, relatively unexplored issue is the criterion for treating a more favorable cohort with a new regimen. At the risk of oversimplification, there are two schools of thought. The first believes that a valuable new regimen will show a modicum of success even in the least prognostically favorable cohort. Barring such success, the drug should not be investigated in a more prognostically favorable cohort. The second school of thought believes that such a demand will result in premature loss of interest in potentially effective drugs. One possible approach to this issue is to test a new drug in a spectrum of prognostic cohorts with the leniency of the stopping rule in a prognostically better cohort, however, depending on previous results in a prognostically worse cohort.

References

1 Hiddemann W, Martin WR, Sauerland CM, Heinecke A, Buchner T. Definition of refractoriness against conventional chemotherapy in acute myeloid leukemia: a proposal based on the results of retreatment by thioguanine, cytosine arabinoside, and daunorubicin (TAD 9) in 150 patients with relapse after standardized first line therapy. *Leukemia* 1990; **4**: 184–188.
2 Estey E, Kornblau S, Pierce S, Kantarjian H, Beran M, Keating M. A stratification system for evaluating and selecting therapies in patients with relapsed or primary refractory acute myelogenous leukemia. *Blood* 1996; **88**: 756.
3 Freireich EJ, Gehan EA, Sulman D, Boggs DR, Frei E III. The effect of chemotherapy on acute leukemia in the human. *J Chron Dis* 1961; **14**: 593.
4 Estey E, Shen Y, Thall PF. Effect of time to achieve CR on subsequent survival time and disease-free survival time in patients with AML, RAEB-t, or RAEB in whom response is established after the first course of chemotherapy. *Blood* (in press).
5 Vey N, Keating M, Giles F, Cortes J, Beran M, Estey E. Effect of complete remission on survival in patients with acute myelogenous leukemia receiving first salvage therapy. *Blood* 1999; **93**: 3149–3150.
6 O'Quigley J, Pepe M, Fisher L. Continual reassessment method: a practical design for phase I clinical trials in cancer. *Biometrics* 1990; **46**: 33–48.
7 Goodman SN, Zahurak ML, Piantadosi S. Some practical improvements in the continual reassessment method for phase I studies. *Stat Med* 1995; **14**: 1149–1161.
8 Thall PF, Russell KE. A strategy for dose-finding and safety monitoring based on efficacy and adverse outcomes in phase I/II clinical trials. *Biometrics* 1998; **54**: 251–264.
9 Thall PF, Estey EH, Sung His-Guang. A new statistical method for dose-finding based on efficacy and toxicity in early phase clinical trials. *Invest New Drugs* (in press).
10 Thall PF, Simon RM, Estey EH. Bayesian sequential monitoring designs for single-arm clinical trials with multiple outcomes. *Stat Med* 1995; **14**: 357–379.
11 Thall PF, Sung HG. Some extensions and applications of a Bayesian strategy for monitoring multiple outcomes in clinical trials. *Stat Med* 1998; **17**: 1563–1580.
12 Thall PF, Estey EH. A Bayesian strategy for screening cancer treatments prior to phase II clinical evaluation. *Stat Med* 1993; **12**: 1197–1211.