

EDITORIAL

Acute myeloid leukemia therapy and the chosen people

Leukemia (2017) 31, 269–271; doi:10.1038/leu.2016.306; published online 11 November 2016

For many are called, but few are chosen.

Mathew 22:14

In this issue of *Leukemia* Montalban-Bravo *et al.*¹ report results of an interesting clinical trial of azacitidine alone or with vorinostat in 109 subjects with acute myeloid leukemia (AML) or high-risk myelodysplastic syndrome (\geq intermediate-2 by IPSS). What is exceptional is the study cohort is composed of subjects usually excluded from clinical trials, whereas subjects usually included in clinical trials were excluded. Eligibility criteria required ≥ 1 of the following: (1) Eastern Cooperative Oncology Group (ECOG) performance score ≥ 3 ; (2) serum creatinine or bilirubin > 2 mg/dl; (3) ≥ 1 diverse co-morbidities; (4) another cancer active or in remission < 2 years. Response and survival were assessed after entry of cohorts of 3–6 subjects with accrual to terminate, should there be $< 5\%$ probability of 60-day survival (primary endpoint), was not $\geq 20\%$ better than the 50% seen previously in 181 similar subjects. There was a similar design to evaluate complete remission rate against the historical rate of 28%. Greater than 20% \geq grade-3 non-hematologic toxicity was also a stopping criterion. The 109 subjects (median age 71 years) included 55% eligible because of a prior cancer, 23% because of an abnormal bilirubin or creatinine, 13% because of an ECOG performance score ≥ 3 and 17% because of other co-morbidities. Six subjects met > 1 eligibility criterion. None of the stopping rules were met. Sixty-day survival rate was 79% and was not influenced by which specific criterion led to eligibility.

There are several potential criticisms of this study. However, this is not the focus of our editorial. Rather, it is the study eligibility criteria which we find most interesting because they provide a precedent for relaxing eligibility criteria in future clinical trials in AML, especially should they be accompanied by stringent monitoring of adverse events as occurred here. Relaxing eligibility criteria could expand numbers of persons entering clinical trials and help bridge the gap between appearance and reality with regard to AML clinical trials.

Reality in AML therapy is perhaps best represented by data from the Surveillance Epidemiology and End Results (SEER) program. Although 5-year survival rates for AML have improved to about 27% in 2006–2012 (the most recent reported interval) the pace of improvement has been slow (<http://seer.cancer.gov/statfacts/html/amyl.html>). Because 60% of persons in the United States ≥ 66 years (the approximate median age at AML diagnosis) receive no therapy within 3 months of diagnosis, the SEER data may underestimate progress among those treated.² However, there is little doubt much of the gap between appearance and reality reflects reports of advances in AML therapy typically outpacing real improvements. Ten years ago one of us (EE) reported about 70% of 91 abstracts describing 39 new therapies for AML presented at the American Society of Hematology (ASH) annual meetings 1993–2001 described promising or encouraging data.^{3,4} Only 15% of studies were declared negative; the other 15% were inconclusive. About 45 of the 63 positive abstracts eventuated in peer-reviewed publications, 38 of which reported favorable outcomes. However with a minimum 5-year follow-up, only one

drug, gemtuzumab ozogamicin, was approved by US Food and Drug Administration (FDA) or European Medicines Agency (EMA) for therapy of AML with approval subsequently withdrawn.⁵ Although factors other than efficacy influence approval and subsequent use, approval *per se* does not necessarily translate to clinically meaningful (an FDA term) benefit for most persons with AML.⁶

Our ASH survey is 15 years old. Although our conclusions may not apply to current ASH abstracts or to new drugs such as midostaurin⁷ and CPX351⁸ (each of which currently has limited indications), we suspect they still operate. This is because the constituency for success of new therapies far exceeds the constituency for objectivity. Investigators, pharmaceutical companies, medical centers and persons with AML cannot be expected to be disinterested given the benefits of seemingly promising results. The numbers of trials untranslatable to clinical practice have led to a recent emphasis on pragmatic trials, which aim to evaluate new drugs in a real-world setting.⁹ As such these trials explore recruitment of subjects and investigators, delivery of the intervention within the trial, nature of follow-up and nature, determination and analysis of outcomes. Here we focus on recruitment of subjects into AML clinical trials, emphasizing the role of selection biases such that persons entering these trials seem the chosen people.

The most obvious cause of selection biases is restrictive study eligibility criteria. Montalban-Bravo *et al.*¹ trial is notably distinctive from the typical protocol requiring an ECOG performance score < 3 , near normal kidney, liver and heart functions, no serious co-morbidities and no uncontrolled infections. Absence of any of these is known to confer an unfavorable prognosis as is the vague but highly important group of persons, typically older, judged unfit. Excluding such subjects improves outcomes of clinical trials but limits translating results to the universe of persons with AML. In other instances, restrictive eligibility criteria may have less impact on outcomes but pose ethical issues. Consider a 60-year-old woman with breast cancer who after surgery has a 90% probability of cure without need for further therapy. One year later she develops AML with complex cytogenetics. According to European LeukemiaNet (ELN)¹⁰ and National Comprehensive Cancer Network (NCCN)¹¹ guidelines she should enter a clinical trial. However, her antecedent breast cancer (clearly not therapy-related) leads to exclusion from trials, although there are no data suggesting it affects her AML prognosis and data from persons with metastatic lung cancer and a history of a prior cancer suggest it would not.¹² A similar conclusion is suggested by data from the Montalban-Bravo *et al.*¹ study. Less restrictive eligibility criteria would potentially address this ethical issue, increase generalizability of results, accelerate accrual rates and decrease complexity and costs of conducting and monitoring clinical trials. Van Spall *et al.*¹³ reported 37% of 2709 exclusion criteria in 283 randomized controlled trials were poorly justified. About 84% of these trials included ≥ 1 such criterion.

Enrollment of newly eligible subjects could decrease effect size. This might occur if, for example, AML subtypes less likely to be benefited from a new drug and/or more likely to be harmed are enrolled. In this circumstance more subjects would be needed to maintain a given power potentially more than counterbalancing the impact of faster accrual and increasing study duration. George¹⁴ examined this issue positing various reductions in effect sizes in the newly eligible subjects and various increases in accrual

rates. If the reduction in effect size was relatively small, a 10% increase in accrual rate would provide enough subjects to maintain power of 0.9 and study length would be unaffected. If the new drug had no effect in the newly eligible subjects, study duration would need to be longer to maintain a power of 0.9 especially if accrual was dominated by newly eligible subjects. Increases in generalizability of conclusions must be balanced against increases in study duration. Interestingly, relatively simple calculations such as those described by George¹⁴ have gained no traction in clinical trials design. Obviously estimating the hazard ratio with the new treatment in currently ineligible subjects is difficult if they continue to be ineligible. Making such subjects eligible might identify additional subgroups likely to respond, especially if enrollment is accompanied by attempts to discover and verify bio-markers associated with response. There are several examples where a wider spectrum of subjects respond to a drug than originally hypothesized.¹⁵ It follows current studies may exclude subjects with no scientific bases. Kim *et al.*¹⁶ noted the rationale for relaxing eligibility criteria remains applicable in the era of precision medicine and molecular-based therapy.¹⁶

Although Montalban-Bravo *et al.*¹ concluded their subjects had been safely treated relative to historical controls, loss of a safety margin with less restrictive eligibility criteria for trials of new drugs is plausible. However, medicine is fundamentally concerned with risks and benefits. It is likely many subjects would conclude the benefit-to-risk ratio with conventional therapy is so low they would prefer to be in a clinical trial. This assumption is implicit in ELN¹⁰ and NCCN¹¹ recommendations for trials in persons with poor prognoses with conventional therapies. Furthermore, reliance on restrictive eligibility criteria as the sole arbiter of study safety may be misleading, reducing medical practice to an exercise in box-checking independent of clinical judgment. We also note decreasing therapy-related mortality (TRM) rates in clinical trials in AML.¹⁷ Although this decrease was in eligible subjects it might also apply to currently ineligible subjects. The eligibility criterion of no prior cancer could be eliminated as could those of a creatinine, bilirubin or other liver function tests levels below a threshold when the drug(s) being studied is not metabolized or excreted by the relevant organ.¹⁸ Requiring normal left ventricular ejection fraction contradicts data indicating anthracycline-induced myocardial toxicity is cumulative and highly unlikely to result from doses of these drugs typically used in initial therapy of AML.¹⁹ There is no reason to require a bone marrow sample if there are many myeloblasts in the blood (save for some research questions), as results of cytogenetics and mutation analyses have been shown concordant between blood and bone marrow samples.^{20,21} Many people are referred after an initial bone marrow has been performed but the study protocol requires another sample for biologic studies of unproved value. Subjects, physicians and third party payers frequently object and the subject is not enrolled in the study. Worse, this process is very plausibly non-random serving to exclude unfit subjects. And many persons > 65 years (excluding the authors) can be healthier than younger persons. Substantial data indicate chronologic age is not the most important determinant of TRM; consideration should be given to replacing age as the criterion of eligibility with a composite index incorporating age with other covariates such as performance score, bilirubin and creatinine rather than regarding each of these in isolation.^{22,23} Such indices are widely available.^{22,23} In addition, eligibility criteria should be unambiguous in contrast to vague statements such as no active infections or reasonable life expectancy.

A more subtle form of selection bias occurs when subjects are excluded from study entry despite meeting eligibility criteria. The magnitude of this selection bias is difficult to quantify. Reasons for excluding these persons are typically non-random and reflect physicians' subjective (but often accurate) impression these persons will not do well. Few would find the results of a study

in which, for example, 30 subjects are treated out of an unknown number of eligible subjects, including those seen but never referred, more credible than the results of a study where number treated approximates numbers eligible. Our experience reviewing typescripts suggest examples of the former far outweigh the latter. Consideration should be given to requiring trialists report not only numbers of subjects treated on-study but also numbers of eligible persons not treated and reason(s) therefore. Outcomes of the latter could be reported as quasi controls.

Selection bias can also be relevant after study initiation. For example, drop-outs from a study often occurs at time of complete remission. A German AML Cooperative Study Group trial randomizing subjects at the time of study-entry to several subsequent therapy allocations, such as maintenance chemotherapy versus an autotransplant found < 30–50% of randomized subjects once in remission received their assigned therapy.²⁴ Selection bias would arise if the drop-outs were less representative of the universe of persons with AML than those receiving their assigned therapy. This might occur if, as typically, those who receive induction and post-remission therapy are younger and fitter than those receiving only induction therapy. Without knowing characteristics (and numbers) of these two groups, one might seriously overestimate the benefit of a post-remission therapy on the starting population. Another example is about 10% of subjects on ECOG trials drop out after each therapy cycle again inflating the estimated benefit on the starting population and on the totality of persons with AML (J Rowe, personal communication).

Although selection bias usually results in over-estimation of benefit, another type of selection bias potentially predisposes to false-negative results. Here new drugs are tested exclusively in subjects likely to fail, such as those with advanced AML or adverse cytogenetics or mutation profiles, unless the drug is extraordinarily effective. This bias might be decreased by successively testing new drugs in cohorts with diverse prognoses. Numbers of subjects in each cohort would depend on the results in the previous cohort with worse outcomes in the previous cohort resulting in fewer subjects in the next cohort. For example, a Bayesian strategy like that used by Montalban-Bravo *et al.*¹ to adjust randomization probability in the expansion phase of their trial might be appropriate.²⁵ This process begins with a prior probability of response for the worst prognosis cohort. Bayes theorem is then used to combine the observed results and the prior probability to yield a posterior probability, which serves as the prior probability for the next worst prognosis cohort. This strategy assumes a continuum of prognostic cohorts with increasing probabilities of a favorable outcome, a readily testable assumption.

We hope Montalban-Bravo *et al.*¹ typescript and this editorial stimulate discussion of the important role of selection biases in determining outcomes of AML therapy. Further improvements in AML therapy may depend on improved understanding of the difference in normal and neoplastic myelopoiesis and the consequent development of truly effective drugs. The pace at which this will occur is, to some extent, beyond our control, difficult as this may be to admit. Consequently, we should focus on controllable factors such as selection biases in clinical trials. Let's increase numbers of chosen people so we might have a better idea of the real impact of new therapies in AML.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

Charles Schiffer (Wayne State University), Hagop Kantarjian (MD Anderson Hospital, Mikkael Sekeres (Cleveland Clinic) and Jacob Rowe (Shari Tzedek Medical Centre, Jerusalem) kindly reviewed the typescript and made it much better than we could

have done alone. RPG acknowledges support from the National Institute of Health Research (NIHR) Biomedical Research Centre funding scheme.

E Estey¹ and RP Gale²

¹*Division of Hematology, University of Washington and Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA and*

²*Division of Experimental Medicine, Department of Medicine, Haematology Research Centre, Imperial College London, London, UK*
E-mail: eestey@uw.edu

REFERENCES

- Montalban-Bravo G, Huang X, Jabbour E, Borthakur G, DiNardo C, Pemmaraju N *et al.* A clinical trial for patients with acute myeloid leukemia or myelodysplastic syndromes not eligible for standard clinic trials. *Leukemia* 2017; **31**: 318–324.
- Medeiros BC, Satram-Hoang S, Hurst D, Hoang KQ, Momin F, Reyes C *et al.* Big data analysis of treatment patterns and outcomes among elderly acute myeloid leukemia patients in the United States. *Ann Hematol* 2015; **94**: 1127–1138.
- Estey EH, Bedikian SH, Witter DC, Pierce SA, Giles F. The predictive value of a 'Positive' ASH abstract in AML therapeutics. *Blood* 2006; **108**: 555a–556a.
- Walter RB, Appelbaum FR, Tallman MS, Weiss NS, Larson RA, Estey E *et al.* Shortcomings in the clinical evaluation of new drugs: acute myeloid leukemia as paradigm. *Blood* 2010; **116**: 2420–2428.
- Ravandi F, Estey EH, Appelbaum FR, Lo-Coco F, Schiffer CA, Larson RA. Gemtuzumab ozogamicin: time to resurrect? *J Clin Oncol* 2012; **30**: 3921–3923.
- Chul K, Vinay P. Cancer drugs approved on the basis of a surrogate end point and subsequent overall survival; An analysis of 5 years of US Food and Drug Administration approvals. *JAMA* 2015; **175**: 1992–1994.
- Stone RM, Mandrekar S, Sanford BL, Geyer S, Bloomfield CD, Döhner K *et al.* The Multi-Kinase Inhibitor Midostaurin (M) Prolongs Survival Compared with Placebo (P) in Combination with Daunorubicin (D)/Cytarabine (C) Induction (ind), High-Dose C Consolidation (consol), and As Maintenance (maint) Therapy in Newly Diagnosed Acute Myeloid Leukemia (AML) Patients (pts) Age 18–60 with FLT3 Mutations (mut): an International Prospective Randomized (rand) P-Controlled Double-Blind Trial (CALGB 10603/RATIFY [Alliance]). *Blood* 2015; **126**: 6.
- Lancet JE, Uy GL, Cortes JE, Newell LF, Lin TL, Ritchie E *et al.* Final results of a phase III randomized trial of CPX-351 versus 7+3 in older patients with newly diagnosed high risk (secondary) AML. *J Clin Oncol* 2016; **34**(suppl): abstr 7000.
- Ford I, Norrie J. Pragmatic trials. *N Engl J Med* 2016; **375**: 454–463.
- Döhner H, Estey EH, Amadori S, Appelbaum FR, Buchner T, Burnett A *et al.* Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood* 2010; **115**: 453–474.
- O'Donnell MR, Tallman MS, Abboud CN, Altman JK, Appelbaum FR, Arber D *et al.* Acute myeloid leukemia, version 2. 2013 *J Natl Compr Canc Netw* 2013; **11**: 1047–1055.
- Gerber DE, Pruitt SL, Halm EA. Should criteria for inclusion in cancer clinical trials be expanded? *J Comp Eff Res* 2015; **4**: 289–291.
- Van Spall HG, Toren A, Kiss A, Fowler RA. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. *JAMA* 2007; **297**: 1233–1240.
- George SL. Reducing patient eligibility criteria in cancer clinical trials. *J Clin Oncol* 1996; **14**: 1364–1370.
- Estey E, Levine RL, Lowenberg B. Current challenges in clinical development of 'targeted therapies': the case of acute myeloid leukemia. *Blood* 2015; **125**: 2461–2466.
- Kim ES, Bernstein D, Hilsenbeck SG, Chung CH, Dicker AP, Ersek JL. Modernizing eligibility criteria for molecularly driven trials. *J Clin Oncol* 2015; **33**: 2815–2820.
- Othus M, Kantarjian H, Petersdorf S, Ravandi F, Godwin J, Cortes J *et al.* Declining rates of treatment-related mortality in patients with newly diagnosed AML given 'intense' induction regimens: a report from SWOG and MD Anderson. *Leukemia* 2014; **28**: 289–292.
- Veatch JR, Sandhu V, Becker PS, Pagel JM, Appelbaum FR, Estey E. The NCI common toxicity criteria and treatment-associated mortality in acute myeloid leukemia. *Blood* 2013; **122**: 293–294.
- Anderlini P, Benjamin RS, Wong FC, Kantarjian HM, Andreeff M, Kornblau S *et al.* Idarubicin cardiotoxicity: a retrospective study in acute myeloid leukemia and myelodysplasia. *J Clin Oncol* 1995; **13**: 2827–2834.
- Weinkauff R, Estey EH, Starostik P, Hayes K, Huh YO, Hirsch-Ginsberg C *et al.* Use of peripheral blood blasts vs bone marrow blasts for diagnosis of acute leukemia. *Am J Clin Pathol* 1999; **111**: 733–740.
- Tong WG, Sandhu VK, Wood BL, Hendrie PC, Becker PS, Pagel J *et al.* Correlation between peripheral blood and bone marrow regarding FLT3-ITD and NPM1 mutational status in patients with acute myeloid leukemia. *Haematologica* 2015; **100**: e97–e98.
- Walter RB, Othus M, Borthakur G, Ravandi F, Cortes JE, Pierce S *et al.* Prediction of early death after induction therapy for newly diagnosed acute myeloid leukemia with pretreatment risk scores: a novel paradigm for treatment assignment. *J Clin Oncol* 2011; **29**: 4417–4423.
- Sorrór ML, Storer BE, Elsayy M, Fathi AT, Brunner A, Gerd S *et al.* Impact of comorbidities at diagnosis of acute myeloid leukemia on one-year mortality. *Blood* 2015; **126**: 532.
- Krug U, Berdel WE, Gale RP, Haferlach C, Schnittger S, Müller-Tidow C *et al.* Increasing intensity of therapies assigned at diagnosis does not improve survival of adults with acute myeloid leukemia. *Leukemia* 2016; **30**: 1230–1236.
- Berry DA. A case for Bayesianism in clinical trials. *Stat Med* 1993; **12**: 1377–1393.