



Selection of Patients for Individual Acute Myeloid Leukemia Therapies

4

Roland B. Walter

4.1 Introduction

In 2019, an estimated 21,450 people developed acute myeloid leukemia (AML) in the United States alone [1]. Until recently, treatment options were relatively limited, and decision-making followed an algorithm that has been invariant for several decades [2, 3]. If the person was felt to be medically fit, cure was considered possible: some form of intensive chemotherapy would be offered followed by further courses of chemotherapy and/or allogeneic hematopoietic cell transplantation (HCT) if a complete remission (CR) was obtained. On the other hand, if the person was felt to be medically unfit, cure was considered rare: in this situation, some form of nonintensive, “palliative” chemotherapy would be offered, most typically low-dose cytarabine or, more recently, single-agent treatment with an azanucleoside (e.g., azacitidine or decitabine) or AML-directed therapy is forgone altogether. Over the last few years, the U.S. Food and Drug

Administration (FDA) has approved eight new drugs for AML, most of them being “targeted” therapeutics: midostaurin, gemtuzumab, ozogamicin, enasidenib, CPX-351, ivosidenib, gilteritinib, glasdegib, and venetoclax [4]. With these, the treatment options have substantially increased, and the line between intensive and non-intensive therapies has become blurrier. Still, although outcomes have gradually improved, AML remains difficult to cure. Many affected individuals will die from consequences of leukemia or treatment-associated complications, and only a minority will be long-term survivors [2–4]. There is thus ongoing need for new therapeutics in AML and need to identify patients most suitable for participation in a clinical trial. With an increasing number of available standard AML therapeutics and ongoing need for new drugs, treatment decision-making has become more complex: should my patient receive standard AML therapy, and if so, with what regimen? Or is my patient better served participating in the testing of an investigational drug? The ability to accurately predict the efficacy of individual treatments in individual patients would greatly improve clinical management as it could form the foundation for evidence-based decision-making regarding the most appropriate treatment. This review will summarize and appraise efforts taken so far to develop tools to predict the risks and benefits of AML therapies for individual patients, focusing on non-transplant treatments.

R. B. Walter (✉)

Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

Division of Hematology, Department of Medicine, University of Washington, Seattle, WA, USA

Department of Laboratory Medicine and Pathology, University of Washington, Seattle, WA, USA

Department of Epidemiology, University of Washington, Seattle, WA, USA
e-mail: rwalter@fredhutch.org

4.2 Outcomes of Interest with AML Therapy

Arguably, the most desirable AML therapy is the one that effectively eliminates leukemia cells and restores normal hematopoietic cell function without treatment-associated morbidity and mortality but with maintenance of a high quality of life (QOL). Also arguably, we are far away from having such a therapy available. In fact, even though the value of successful AML therapy was quickly recognized once the first chemotherapeutics for acute leukemia became available [5], the presumption that potential risks are not commensurate with potential benefits often leads physicians and/or patients to shy away from AML-directed therapy, at least for older people. This is indicated by recent estimates that less than half of Americans aged >65 years with newly diagnosed AML, and as few as 10–20% of those aged >80 years, receive specific chemotherapy, and only a minority does so in specialized cancer centers [6, 7].

At least in the era of non-targeted AML therapy, multiagent chemotherapy was felt to be a pivotal component of a curative treatment strategy [2–4]. Although cautious interpretation of findings is warranted, data from the U.S. and European population-based registry data suggest value of intensive chemotherapy not just for younger patients but also for older individuals up to age 80 years or perhaps beyond [7, 8]. Considering that AML primarily affects older people, many of whom will have comorbidities that could limit drug tolerance, it is therefore not surprising that most efforts have focused on developing models to estimate the fitness to tolerate intensive multiagent chemotherapy. With continued improvements in supportive care, however, our abilities to support patients throughout the periods of disease/treatment-related cytopenias have progressively increased and early deaths with intensive chemotherapy have significantly decreased [9–11]. Likewise, non-relapse mortality with allogeneic HCT has substantially declined over time [12]. Thus, intensive therapies can now be given more safely to treat AML, even in older adults. With this, primary failure of AML therapy or disease recurrence after a period of

remission—the two outcomes that constitute “therapeutic resistance”—has become the principal life-limiting problem in AML.

4.3 Brief Statistical Considerations

When considering approaches to estimating outcomes with AML therapy, it is important to distinguish between association and prediction (or classification) models. Association models, as the name implies, aim to identify associations between covariates and patient outcomes. Common measures of association in clinical studies are odds ratios and hazard ratios, which are interpreted as an average effect in the study population [13]. In contrast, prediction models aim to evaluate the ability of one or more covariates to predict outcomes for individual patients. Common measures of prediction models are sensitivity, specificity, positive predictive value, negative predictive value, and the area under the receiver operating characteristic curve (AUC) [14]. A strong association is usually necessary but not sufficient for a model to be able to predict well [15]. For a binary outcome (e.g., early death/no early death), the AUC measure can take values between 0.5 and 1.0, with an AUC of 0.5 being analogous to a coin flip and an AUC of 1.0 denoting perfect prediction. It is commonly accepted that AUCs of 0.6–0.7, 0.7–0.8, and 0.8–0.9 indicate poor, fair, and good predictive ability, respectively [16–18].

4.4 Predicting AML Therapy-Related Mortality

Most approaches to predict the toxicity of AML therapy have focused on deaths within 28–30 days (sometimes within 60 days) of beginning chemotherapy. A rationale behind this is the observation that weekly death rates sharply decline after 4 weeks, suggesting patients who die in this time frame are qualitatively different from those who do not [19]. However, considering early death to be equivalent to treatment-related mortality

(TRM) is flawed because deaths may be related to disease-associated myelosuppression or organ dysfunction or—as recently shown in an institutional trial using reduced-intensity CPX-351 [20]—early progression of AML rather than be a direct consequence of the therapy given. Thus, in many patients, early deaths may occur *despite* rather than *because of* the AML therapy. A cleaner way to assess treatment-related toxicities might be to model early deaths with post-remission therapy for patients who have attained full hematologic recovery with prior courses of treatment although the nature of the therapy given and the patients receiving it will limit the conclusions that could be drawn from such models.

Over the years, many factors have been associated with early death, including age and covariates such as albumin and creatinine that may serve as surrogates for biological (rather than chronological) age. Such factors allow building of multicomponent scores reflective of the probability of early death with AML therapy. Several scoring systems aimed at identifying patients at high early death risk after intensive chemotherapy have been developed [19, 21–27]. Some of these systems reach good predictive ability with AUC values above 0.8. While they differ in the details, they all indicate the accuracy of predicting early death is optimized when a combination of factors rather than just one factor (such as age or performance status) is considered. This observation underlies the recommendation by the European LeukemiaNet and the National Comprehensive Cancer Network to consider age in the context of other covariates when considering the appropriateness of intensive AML therapy [3, 28].

Although not perfect, existing models to predict early death offer an empiric approach of selecting patients who will not die early after receiving intensive chemotherapy. It is plausible that models could be improved by integrating additional covariates such as comorbidities not captured in current models, additional information on patient demographics (e.g., educational level), site of treatment, among others [4]. To what degree comprehensive geriatric assessments

could improve early death predictions after AML therapy is currently unknown but important to determine. It is becoming more and more clear that geriatric assessments provide a framework for an individual patient's fitness for therapy and can help in personalized decision-making [29]. Several studies have demonstrated geriatric assessments provide information that, independently, is associated with survival in older patients with AML [30–32], and it is possible such information could improve multicomponent early death prediction models. As a consequence of the progressively declining rates of early death with intensive AML therapy and increasing number of available treatments re-calibrating existing systems (to account for changes in the supportive care pattern) and developing new systems is becoming increasingly more challenging as larger and larger datasets of similarly-treated patients will be required to model early death mathematically.

4.5 Predicting Non-fatal Toxicities of AML Therapy

In contrast to the many efforts spent on trying to predict early death after intensive AML therapy, understanding the degree to which non-fatal toxicities can be predicted has not been of major interest, and it is not understood which patient characteristics are most strongly associated with occurrence of such toxicities. One recent study has examined these questions using data from 260 adults age 18–60 years with AML treated with 7+3 on a contemporary cooperative study group (SWOG) Phase 3 trial [33]. The following baseline covariates were assessed: age at study registration, gender, performance status, pre-study white blood cell (WBC) counts, pre-study platelets counts, pre-study hemoglobin (HGB), pre-study bone marrow blast percentage, secondary vs. de novo AML, cytogenetic risk, and *NPM1* as well as *FLT3*-ITD mutation status. In univariate models, no individual covariate was a strong predictor of toxicity. Only three pairs of toxicity/covariate had an AUC >0.65: older age predicting increased risk of endocrine

abnormalities (AUC = 0.67), higher baseline WBC predicting increased risk for bleeding (AUC = 0.67), and higher baseline HGB predicting increased risk of neurologic toxicity (AUC = 0.69). As incidence allowed, multivariable models were evaluated which showed increased AUCs compared to univariate models, but no multivariable model had an AUC larger than 0.70. Within the limitation that not all covariates important to predict toxicities may be captured in cooperative group datasets and that patients with significant organ dysfunction were excluded from trial participation, these findings indicate that there is a poor ability to predict commonly occurring grade 3 and higher toxicities that occur with multiagent AML chemotherapy.

4.6 Predicting the Efficacy of AML Therapy

At the cohort level, many disease characteristics, in particular cytogenetic and molecular abnormalities, have been associated with measures of therapeutic efficacy, e.g., achievement of complete remission (CR), relapse rates, event/disease-free survival, or overall survival [2–4]. Forecasting efficacy of therapy for individual people with AML, on the other hand, has proved relatively difficult. Using data from over 4500 adults treated with conventional intensive AML chemotherapy, it was found that there was only a fair ability to predict failure to achieve CR with the initial 1–2 courses of chemotherapy or to have a short relapse-free survival if CR was obtained. Various models that included basic patient characteristics (age, performance status) and commonly available disease characteristics (white blood cell count, secondary disease, cytogenetic risk, and *NPM1* as well as *FLT3*-ITD mutation status) had AUCs typically ranging from 0.71 to 0.78 [34]. This finding of only a fair ability to predict CR is consistent with a study by Krug et al. who observed AUCs of 0.72 and 0.68 with multivariable models in their study cohort [24]. These relatively low AUCs suggest caution to avoid overestimating our ability to predict

resistance following standard therapy of AML, which is closer to a coin-flip than certainty in many instances when commonly utilized factors are considered.

To some degree, inclusion of additional disease characteristics can refine prediction of therapeutics. For example, data from a larger number of additional commonly occurring mutations improved the predictive accuracy of simpler models minimally (from AUCs of 0.70–0.76 to 0.72–0.80 in a cohort of 298 patients treated uniformly on a cooperative study trial) [35]. Moreover, a score derived from expression data from 17 genes associated with stemness of leukemia cells (17-gene LSC score, “LSC17”)—yielding an AUC of 0.78 for the prediction of failure to achieve CR after initial induction therapy—improved a multicomponent prediction model for this endpoint from an AUC of 0.73 to an AUC of 0.82 [36]. However, it is likely that even highly sophisticated genetic models will come short of high accuracy. Data from a comprehensive genetic analysis of over 1500 patients suggested that genomic features—while being the most powerful predictors—accounted for only about two thirds of the observed variation in survival. One third of this variation was attributed to demographics, clinical, and treatment variables [37].

4.7 Predicting the QOL Impact of AML Therapy

QOL is severely reduced in people diagnosed with AML and is affected over time as a patient goes through AML-directed therapy successfully or unsuccessfully [38]. Current evidence indicates that different treatments will affect the QOL in different ways. For example, QOL may further decrease early after receiving intensive chemotherapy but then improve, whereas it may be stable initially with non-intensive treatment but worsen over time. QOL considerations therefore need to play an important role in the daily care of AML patients. Undoubtedly, QOL is linked to treatment toxicities and efficacy, but there will be some elements that are independent. For example,

for two therapies that are equally toxic and effective, there might be strong preference to receive this treatment at home (e.g., as oral medication) or in the clinic rather than requiring administration in the hospital. To date, no efforts have been made to predict QOL (or, rather, changes in QOL) with different types of AML therapy. One barrier to modeling QOL endpoints in AML is the lack of a disease-specific QOL instrument that can efficiently capture the major QOL deficits in this population. Efforts to correct this deficiency are ongoing [39].

4.8 Outcome Prediction in the Era of Targeted AML Therapy and Rapidly Evolving Treatment Algorithms

With recent regulatory approval of several small molecule inhibitors and one antibody–drug conjugate, we have now entered the era of targeted therapy in AML. Unlike the treatments that target PML-RARA in acute promyelocytic leukemia (APL), however, so far none of the existing targeted AML drugs has near-perfect efficacy even in patients selected by the presence of documented abnormalities in the drug target. In a disease as heterogeneous as AML where genetic abnormalities typically appear to work in concert rather than single handedly to drive the leukemic process, they never may. Thus, having good tools available that can help select individual therapies will remain as important as it is today. To what degree relevant outcomes with targeted AML therapies can be predicted is not known. It will take a considerable amount of time to gather large-enough datasets that allow development and validation of such models.

Since outcome prediction models are reflective of the time when they were developed, capturing not only the anti-AML therapy given but also the supportive care provided, they are—by default—outdated at the time they are introduced into clinical use. That is true even if the general therapeutic strategy does not substantially change. Fortunately, after many years of no

change, we have now seen rapid introduction of several new drugs for AML. With every additional drug approved for clinical use, there will be more treatments and treatment combinations available to choose from. As seen today by the shift away from low-dose cytarabine or azanucleoside monotherapy to lower-intensity doublet (or triplet) therapies, some treatments may become quickly obsolete and replaced by others. Constant shifts in treatment paradigms pose a real challenge for physicians and patients interested in empiric approaches to help choose the most appropriate treatment given the time it takes to establish a validated treatment selection tool. Rather than estimating outcomes with the treatment of interest, they may be left with estimating outcomes with “older” treatments and, indirectly, have to use that information to decide whether it is worth pursuing an alternative treatment. Similar to how we might think about deciding between “standard” and investigational therapy.

4.9 Conclusion

There is no shortage in the scoring systems aimed to identify patients at high risk of either early death or treatment resistance after conventional intensive AML therapy. They offer an empiric approach of selecting patients who will do well with standard AML chemotherapy. However, there are important caveats physicians and patients need to be aware of when utilizing these tools. First and foremost, the accuracy of these prediction models is imperfect even at the time of their development, highlighting our limitations in comprehensively capturing and mathematically describing the factors relevant for outcomes of AML therapies. As the rather small improvements in accuracy between relatively simple and complex resistance models indicate, it is very unlikely that we will reach perfect (or near-perfect) prediction accuracy, at least not when trying to forecast the results with conventional AML therapy. Second, scoring systems are likely not agnostic to the type of AML therapy given. Especially at times when newly approved drugs become available for routine use and the standard

of care approach changes, existing models are no longer capturing the clinical reality. It will take great, concerted effort and large patient datasets to refine prediction models based on data derived from patients receiving new standard therapies. And finally, scoring systems are a reflection of factors that mattered at the time the patient received treatment that contributed to the models. Already imperfect at the time of development, the models' accuracy will likely decrease over time with changes in AML care. For example, the rate of early death following intensive induction chemotherapy has declined considerably over the last 20 years because of improvement in supportive care [9–11]. Thus, early death prediction tools—but also resistance prediction tools that are affected by non-leukemia-related deaths—need to be re-assessed and re-calibrated periodically to account for our increasing ability to keep AML patients alive. The task of updating mortality prediction models will become more and more difficult as death rates decline.

Conflict of Interest The author declares no competing financial interests.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin*. 2019;69(1):7–34.
2. Döhner H, Weisdorf DJ, Bloomfield CD. Acute myeloid leukemia. *N Engl J Med*. 2015;373(12):1136–52.
3. Döhner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood*. 2017;129(4):424–47.
4. Estey EH. Acute myeloid leukemia: 2019 update on risk-stratification and management. *Am J Hematol*. 2018;93(10):1267–91.
5. Freireich EJ, Gehan EA, Sulman D, Boggs DR, Frei E 3rd. The effect of chemotherapy on acute leukemia in the human. *J Chronic Dis*. 1961;14(6):593–608.
6. Meyers J, Yu Y, Kaye JA, Davis KL. Medicare fee-for-service enrollees with primary acute myeloid leukemia: an analysis of treatment patterns, survival, and healthcare resource utilization and costs. *Appl Health Econ Health Policy*. 2013;11(3):275–86.
7. Oran B, Weisdorf DJ. Survival for older patients with acute myeloid leukemia: a population-based study. *Haematologica*. 2012;97(12):1916–24.
8. Juliusson G, Antunovic P, Derolf Å, et al. Age and acute myeloid leukemia: real world data on decision to treat and outcomes from the Swedish Acute Leukemia Registry. *Blood*. 2009;113(18):4179–87.
9. Halpern AB, Culakova E, Walter RB, Lyman GH. Association of risk factors, mortality, and care costs of adults with acute myeloid leukemia with admission to the intensive care unit. *JAMA Oncol*. 2017;3(3):374–81.
10. Othus M, Kantarjian H, Petersdorf S, 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(2):289–92.
11. Percival ME, Tao L, Medeiros BC, Clarke CA. Improvements in the early death rate among 9380 patients with acute myeloid leukemia after initial therapy: a SEER database analysis. *Cancer*. 2015;121(12):2004–12.
12. Gooley TA, Chien JW, Pergam SA, et al. Reduced mortality after allogeneic hematopoietic-cell transplantation. *N Engl J Med*. 2010;363(22):2091–101.
13. Harrell FE Jr. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. New York, NY: Springer; 2001.
14. Pepe MS. The statistical evaluation of medical tests for classification and prediction. New York, NY: Oxford University Press; 2004.
15. Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidemiol*. 2004;159(9):882–90.
16. Ishihara R, Yamada T, Iishi H, et al. Quantitative analysis of the color change after iodine staining for diagnosing esophageal high-grade intraepithelial neoplasia and invasive cancer. *Gastrointest Endosc*. 2009;69(2):213–8.
17. Muller MP, Tomlinson G, Marrie TJ, et al. Can routine laboratory tests discriminate between severe acute respiratory syndrome and other causes of community-acquired pneumonia? *Clin Infect Dis*. 2005;40(8):1079–86.
18. Roelen CA, van Rhenen W, Groothoff JW, van der Klink JJ, Bultmann U, Heymans MW. The development and validation of two prediction models to identify employees at risk of high sickness absence. *Eur J Pub Health*. 2013;23(1):128–33.
19. Walter RB, Othus M, Borthakur G, 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(33):4417–23.
20. Walter RB, Othus M, Orłowski KF, et al. Unsatisfactory efficacy in randomized study of reduced-dose CPX-351 for medically less fit adults with newly diagnosed acute myeloid leukemia or other high-grade myeloid neoplasm. *Haematologica*. 2018;103(3):e106–9.
21. Giles FJ, Borthakur G, Ravandi F, et al. The haematopoietic cell transplantation comorbidity index

- score is predictive of early death and survival in patients over 60 years of age receiving induction therapy for acute myeloid leukaemia. *Br J Haematol*. 2007;136(4):624–7.
22. Kantarjian H, O'Brien S, Cortes J, et al. Results of intensive chemotherapy in 998 patients age 65 years or older with acute myeloid leukemia or high-risk myelodysplastic syndrome: predictive prognostic models for outcome. *Cancer*. 2006;106(5):1090–8.
 23. Kantarjian H, Ravandi F, O'Brien S, et al. Intensive chemotherapy does not benefit most older patients (age 70 years or older) with acute myeloid leukemia. *Blood*. 2010;116(22):4422–9.
 24. Krug U, Röllig C, Koschmieder A, et al. Complete remission and early death after intensive chemotherapy in patients aged 60 years or older with acute myeloid leukaemia: a web-based application for prediction of outcomes. *Lancet*. 2010;376(9757):2000–8.
 25. Malfuson JV, Etienne A, Turlure P, et al. Risk factors and decision criteria for intensive chemotherapy in older patients with acute myeloid leukemia. *Haematologica*. 2008;93(12):1806–13.
 26. Savic A, Kvrjic V, Rajic N, et al. The hematopoietic cell transplantation comorbidity index is a predictor of early death and survival in adult acute myeloid leukemia patients. *Leuk Res*. 2012;36(4):479–82.
 27. Valcárcel D, Montesinos P, Sánchez-Ortega I, et al. A scoring system to predict the risk of death during induction with anthracycline plus cytarabine-based chemotherapy in patients with de novo acute myeloid leukemia. *Cancer*. 2012;118(2):410–7.
 28. O'Donnell MR, Tallman MS, Abboud CN, et al. Acute myeloid leukemia, version 3.2017, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Cancer Netw*. 2017;15(7):926–57.
 29. Michaelis LC, Klepin HD, Walter RB. Advancements in the management of medically less-fit and older adults with newly diagnosed acute myeloid leukemia. *Expert Opin Pharmacother*. 2018;19(8):865–82.
 30. Deschler B, Ihorst G, Platzbecker U, et al. Parameters detected by geriatric and quality of life assessment in 195 older patients with myelodysplastic syndromes and acute myeloid leukemia are highly predictive for outcome. *Haematologica*. 2013;98(2):208–16.
 31. Klepin HD, Geiger AM, Tooze JA, et al. Geriatric assessment predicts survival for older adults receiving induction chemotherapy for acute myelogenous leukemia. *Blood*. 2013;121(21):4287–94.
 32. Sherman AE, Motyckova G, Fega KR, et al. Geriatric assessment in older patients with acute myeloid leukemia: a retrospective study of associated treatment and outcomes. *Leuk Res*. 2013;37(9):998–1003.
 33. Othus M, Garcia-Manero G, Appelbaum FR, Erba HP, Walter RB. Predicting induction toxicity with 7+3: analysis of SWOG trial S1203 [abstract]. *Blood*. 2018;132(Suppl 1):1403.
 34. Walter RB, Othus M, Burnett AK, et al. Resistance prediction in AML: analysis of 4601 patients from MRC/NCRI, HOVON/SAKK, SWOG and MD Anderson Cancer Center. *Leukemia*. 2015a;29(2):312–20.
 35. Walter RB, Othus M, Paietta EM, et al. Effect of genetic profiling on prediction of therapeutic resistance and survival in adult acute myeloid leukemia. *Leukemia*. 2015b;29(10):2104–7.
 36. Ng SW, Mitchell A, Kennedy JA, et al. A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature*. 2016;540(7633):433–7.
 37. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic classification and prognosis in acute myeloid leukemia. *N Engl J Med*. 2016;374(23):2209–21.
 38. Buckley SA, Kirtane K, Walter RB, Lee SJ, Lyman GH. Patient-reported outcomes in acute myeloid leukemia: where are we now? *Blood Rev*. 2018a;32(1):81–7.
 39. Buckley SA, Jimenez-Sahagun D, Othus M, Walter RB, Lee SJ. Quality of life from the perspective of the patient with acute myeloid leukemia. *Cancer*. 2018b;124(1):145–52.