

Acute Myeloid Leukemia: How Do We Measure Success?

Joshua P. Sasine¹ · Gary J. Schiller²

Published online: 7 September 2016
© Springer Science+Business Media New York 2016

Abstract The development and approval of novel, effective therapies for acute myeloid leukemia (AML) has lagged behind other malignancies. Judging success of therapy with meaningful endpoints is critical to development of new treatments. Overall survival (OS) has typically been the parameter necessary for regulatory approval of experimental therapy in AML. Herein, we discuss different strategies to define outcomes for patients with AML and their relative challenges.

Keywords Acute myeloid leukemia · Therapy · Overall survival · Event-free survival

Introduction

Of all the products of bone marrow production, granulocytopoiesis is perhaps the most important in the immediate clinical sense. Unlike erythrocytes and platelets, no long-term effective replacement for these cells exists, and unlike cells of the

B-lymphocyte lineage, long-term survival in the absence of neutrophils is not realistic. Thus, when a patient presents with a malignancy of the myeloid lineage, effective therapies must impair the abnormal myeloid lineage cells while attempting to spare the normal counterparts. Even with detailed molecular knowledge currently available for the study of malignancy, achieving this degree of specificity has been elusive, despite the high numbers of therapies in development with significant promise [1••]. While breakthroughs for treatment of other malignancies of the blood and bone marrow have advanced and the US Food and Drug Administration (FDA) has granted many approvals in recent years, new approvals for acute myeloid leukemia (AML) are conspicuously absent. One explanation is the high bar for success that biology demands. However, due to the complexity of variables involved, it may be that the definition of success in this disease requires a reappraisal.

To grant approval for new drugs, the US FDA has defined two main regulatory pathways: regular and various expedited programs for serious conditions [2]. The regular pathway requires proof of clinical benefit, generally defined as improvement in length of life and/or increased quality of life (QoL), or through validated surrogates for these end points [2]. The FDA instituted the Accelerated Approval Program to decrease the time to approval for drugs that fill an unmet medical need based on a surrogate endpoint, with the understanding that after meeting this surrogate endpoint, evidence of benefit would later be demonstrated by the Sponsor. Expedited programs include fast track, breakthrough therapy, accelerated approval, and priority review. Overall survival (OS) is typically the endpoint used when evaluating benefit for new cancer drugs. A surrogate endpoint is an alternative test, such as a laboratory measurement, radiographic image, or other evaluation that is closely tied to, and can reliably predict clinical benefit, but is not itself a measure of clinical benefit. The main

This article is part of the Topical Collection on *Acute Myeloid Leukemias*

✉ Joshua P. Sasine
jsasine@mednet.ucla.edu

Gary J. Schiller
gschiller@mednet.ucla.edu

¹ UCLA Department of Medicine, Division of Hematology and Oncology, Orthopedic Hospital Research Center/BSRB, 615 Charles E. Young Drive South, Room 545, Los Angeles, CA 90095, USA

² UCLA Department of Medicine, Division of Hematology and Oncology, Aramont Foundation for Clinical/Translational Research in Human Malignancies, Room 42–121 Center for Health Sciences, David Geffen School of Medicine at UCLA, Los Angeles 90095, CA, USA

potential advantages of these endpoints are faster results and, in some cases, more reliable assessment of drug effects than survival. In this article, we will discuss the merits and pitfalls of various measures of success when treating AML. Unfortunately, there is no ideal endpoint for evaluating benefit in AML.

Overall Survival

The goal for any therapy for a malignancy is ultimately to extend life, and overall survival (OS) is the most straightforward way of evaluating that goal. There are two main drawbacks to using OS in AML. Firstly, generating data requires a sufficiently large trial with enough deaths occurring to show a difference. It can take 3 years or more [3, 4] to gather sufficient survival data, especially in younger patients with AML, and this might constitute an unacceptably long delay. Secondly, confounding variables are quite common and might dissociate the effect of pharmacologic management of AML from survival attributed to other interventions, sufficient to dilute any benefit of the novel pharmacologic agent.

If a drug were able to durably eradicate all AML and pre-leukemia subclones from a substantial proportion of patients with AML without affecting normal hematopoiesis or causing other significant toxicities, there is no doubt that this therapy would prolong survival in a measurable way. But is this the bar that should be set or that can be reasonably achieved? Without a doubt, any therapy with an enormous improvement in efficacy without imposing additional risk would improve OS, but the opposite is not necessarily true. If an intervention does not measurably prolong survival in a population-based study, it might still have meaningful benefit as a means to achieve that survival endpoint.

To address this question, we must address a different issue first, namely, the causes of death in patients with AML. Most people with an epithelial malignancy (the majority of people with malignancies in general) die from progressive disease that directly causes organ failure by mass effect. Although extramedullary AML with organ infiltration can occur, the majority of patients with leukemia have disease limited principally to the blood and bone marrow. More commonly, the malignant clone is controllable, but at the expense of bone marrow failure and loss of normal hematopoiesis. Given the advances in transfusion medicine, death from uncontrolled bleeding or high-output heart failure is now less common. So, in general, people with progressive AML die from infections secondary to bone marrow failure. This adds a variable with intrinsic inconsistency into the mix; after all, some people with bone marrow failure succumb quickly to lethal infection while others can persist for many months without functional granulopoiesis. As antibiotic therapy has improved, this adds to the complexity in the interpretation of what it means to

survive longer with AML. These interventions are often not delineated in trial design and as a direct consequence, used in non-random ways, even in randomized trials, which essentially converts the trial into an observational study of many uncontrolled variables [5]. Predicting survival in AML is a complex mixture of anti-AML therapy, supportive care, biology of the disease, and patient characteristics.

Another aspect of treatment that can be a source of confounding is post-remission therapy, namely allogeneic hematopoietic transplant (allo-HCT). Due to advances in donor selection, advancement in haploidentical transplants, and increasingly effective supportive care, this therapy is expanding as fewer patients are ineligible due to advanced patient age or lack of a donor. While this is an advance for transplant, it also increasingly complicates the interpretation of survival data since treatment-related mortality (TRM) or non-relapse mortality (NRM) of the procedure will inevitably dilute the effects of investigational therapy. If lack of OS benefit is confirmed in that setting, it might be challenging to parse out the etiology, since it could be due to transplant effects or due to true lack of clinical benefit. One potential example of the impact of transplant on interpretation of study results comes from the randomized trial of vosaroxin, the VALOR trial [6]. Overall, there was a significant benefit in the rate of remission induction, though no OS difference; but on subgroup analysis, patients older than age 60 had a survival benefit (median OS 7.1 vs 5.0 months in the placebo arm) with a relatively lower allo-HCT rate (20 vs 46 %). Importantly, transplant rates between placebo and study groups were not significantly different. Perhaps the survival benefit induced by achieving CR in the younger patients was diluted to insignificance by the higher rates of allo-HCT in the younger population, or, alternatively, the drug might just be more efficacious for older patients owing to different disease biology. Use of other endpoints can be occasionally helpful in answering a question like this, but are no less problematic, as discussed later. For example, in the VALOR trial, the remission difference from placebo group was slightly higher in the older cohort, but whether this can fully explain the OS difference seen with older patients is unclear given the multiplicity of variables and the lack of a straightforward correlation between remission rates and OS in general.

Another drawback to using survival in any trial, AML or otherwise, is the difficulty in incorporating crossover design in late-phase studies, something important in recruitment to clinical trials. Allowing patients initially in the control group to cross over will dilute the survival benefit, if any. Add to this list the fact that AML is a biologically heterogeneous disease and that survival data is usually presented as a median. If fewer than half of the patients benefit, the median is unchanged. Given the biological heterogeneity of AML [7–10], the possibility of only a subpopulation gaining benefit is realistic. With these variables in mind, it is clear why

clinicians may believe that survival data are insufficient to define success of therapy for AML. But are other surrogates more useful?

Event-Free Survival

Event-free survival (EFS) is a composite endpoint measuring the time between treatment and a major event which includes death or relapse, but could also include other serious anticipated complications. In this respect, it captures progression-free survival and is often identical to that endpoint, but could include other events. This highlights a weakness of using EFS: the inherent non-uniformity. The definition of events is important and might differ in relevant ways across trials. EFS in the setting of AML often includes failure to obtain a CR. Given the fact that therapy for other cancers has been approved based on improvements in progression-free survival [11, 12], it might be reasonable to extend this approach to AML. However, while EFS is correlated with OS in AML, the correlation is not tight, and its usefulness as a surrogate has been called into question [13]. One recent study [14] showed only a moderate correlation of EFS with OS. Other studies have suggested that the correlation is not always reliable enough to justify replacement as a surrogate endpoint [13]. The correlation of EFS with OS may also vary with intensity of therapy. For intense therapies, the events might actually worsen in the immediate period due to the therapy; but if a substantial portion of patients have long-lasting control of AML, this could be ultimately beneficial. However, the EFS might not reflect this. EFS also highlights the importance of patient selection. Patients who do not respond to the therapy but also have toxicity will counterbalance the results of patients who respond favorably. Also, cumulative toxicity of a drug might not be evident until later on and this could be a way for EFS to improve but if those who do not respond to the therapy do worse than the control arm overall, the EFS might not reflect it. One potential example of this effect is the addition of gemtuzumab ozogamicin (GO) to low-dose cytarabine (LDAC) in older patients with AML [15]. In this trial that compared LDAC to the same plus GO, among patients who did not achieve remission, survival was significantly better with LDAC alone (15 vs 9 %; HR 1.27 (1.03–1.56), $P=0.03$). There was no difference in 30-day mortality to account for the difference so this could suggest that added cumulative toxicity of the drug with lack of disease control was a contributor, but this is unclear. Additionally, among patients who initially responded to treatment but relapsed, survival time following relapse was better in the LDAC arm (37 %) than the LDAC + GO arm (11 %; HR 1.49 (0.93–2.48), $P=0.09$). GO improved the remission rate from 11 to 21 % (OR 0.46 (0.29–0.75) $P=0.002$) but taken together, possibly

owing to the effects above, there was no overall survival difference noted.

The definition of the events is important in EFS, and patients may have an improvement in QoL despite having an event such as lack of achieving CR. For example, CR may be achieved after the time-point specified in the study due to another line of therapy, and this next line of therapy might only have been possible because the disease was controlled or modified by the investigational therapy. Also, CR will only be seen in the minority of patients treated with a hypomethylating agent, but they may have sufficient hematologic recovery to have an improved QoL while being spared the QoL decrement imposed by standard induction chemotherapy [16, 17]. Thus, increasing quality of life could actually come at the expense of other endpoints like CR and EFS. In this tug-of-war for endpoints, predetermination of endpoints based on experience with the drug and expected outcome is pivotal in trial design.

Complete Remission

Bone marrow blasts <5 % with predetermined definitions of normal bone marrow function has been the standard definition for defining remission in AML and achievement of complete remission (CR) is associated with a better outcome [18–20]. CR has formed the basis for regulatory approval in other malignancies of the bone marrow [2, 21]. Achieving a CR was originally associated with OS and the prolongation of survival was proportional to the interval of CR [22]. This is consistent with a theme that runs throughout AML literature, higher burden of disease is associated with risk of relapse, and although intensification of therapy can modify this risk, all other things being equal, higher leukemia burden is a risk of progression, relapse, and death. This also holds true if allo-HCT is performed [23]. Allo-HCT during remission is associated with a lower likelihood of relapse compared to allo-HCT with active disease (especially if the bone marrow blasts are >25 % or circulating) [23]. Despite that, encouraging results have been observed using myeloablative allo-HCT to treat highly selected patients who have active refractory disease with a good performance status, no circulating blasts, a prior CR duration of >6 months, and no poor-risk cytogenetics. The 3-year OS was 42 % with matched-sibling transplants [24] in this population. More sensitive ways of detecting disease would not address the problem of transplant selection in patients with active disease, but given the high numbers of patients who relapse despite being in a morphologic CR, this is clearly a problematic way of assessing response. More sensitive measures are required to subdivide patients in a CR into separate risk groups. Moreover, several trials have now shown a decoupling of CR from OS (Table 1), casting doubt on its value as a surrogate endpoint.

Table 1 Strategies that increase rates of remission without affecting overall survival in AML

Therapy	Patients	Study design	CR	OS
Gemtuzumab ozogamicin [53]	495 patients with AML, unfit for standard induction	LDAC + GO vs. LDAC	30 % with LDAC + GO vs. 17 % with LDAC ($P < 0.006$)	27 % LDAC + GO vs 25 % with LDAC at 12 months ($P = 0.9$)
Clofarabine [54]	406 older patients, AML or high-risk MDS, unfit for standard induction	Clofarabine + LDAC vs LDAC	22 % with clofarabine vs. 11 % LDAC ($P < 0.005$)	13 % with clofarabine vs 12 % with LDAC at 2 years surviving ($P = 0.7$)
Clofarabine [55]	320 patients with relapse/refractory AML, median age 67	Clofarabine + Ara-C vs. Ara-C + placebo	35.2 % with clofarabine + Ara-C versus 17.8 % with Ara-C + placebo ($P < 0.01$)	Median OS 6.6 months with clofarabine + Ara-C vs 6.3 with Ara-C + placebo ($P = 1.0$)
Vosaroxin [6]	711 patients, relapsed/refractory AML, age >18	Phase III, vosaroxin + Ara-C vs. Ara-C + placebo	30 % with vosaroxin + cytarabine vs. 16 % with cytarabine + placebo ($P < 0.0001$)	Median OS 7.5 months with vosaroxin vs. 6.1 with Ara-C + placebo ($P = 0.06$)
	Pre-specified subgroup analyses of the original 711 patients with the 451 age >60 only with relapsed/refractory AML		32 % with vosaroxin + cytarabine vs. 14 % with cytarabine + placebo ($P < 0.0001$)	Median OS 7.1 months with vosaroxin vs. 5.0 months placebo ($P = 0.003$)

Ara-C cytarabine, LDAC low-dose cytarabine, GO gemtuzumab ozogamicin

Minimal Residual Disease by Flow Cytometry

Molecular and cytogenetic abnormalities are one way to risk-stratify patients with AML, but increasingly, minimal residual disease (MRD) assessment by flow cytometric methods is recognized as a complementary tool. The ideal time to assess for MRD is not clear and the degree of minimal residual disease by flow cytometry (flow-MRD) to be considered “positive” is not yet defined in a standard way, but the level of detectable disease correlates with relapse risk [25]. Pre-transplant flow-MRD has been associated with risk for relapse after myeloablative allo-HCT [26] and non-myeloablative allo-HCT [27], and the risk of relapse is similar for each [27]. Flow-MRD might be useful for informing decisions of whether to proceed with allo-HCT in some settings, particularly with patients that have intermediate-risk disease by other criteria, although this is not done routinely [25]. Some reasons for the lack of widespread adoption of MRD assessment include the technically challenging nature of the test and the non-uniform thresholds for risk-stratification, along with the heterogeneity in assessment-timing reported in the literature [28]. Despite these limitations, flow-MRD has widespread applicability and can be used in >90 % of persons with AML [29]. Although it has now been established that flow-MRD is independent from cytogenetics, it has not yet been clear if the modality of intensification therapy (chemotherapy or transplant) modifies the influence of the MRD level assessed after

induction therapy. These questions have yet to be addressed in a prospective fashion.

However, despite the challenges associated with flow-MRD, a few striking conclusions can be made. First, patients undergoing allo-HCT while in morphologic CR with any level of detectable flow-MRD have a substantially higher relapse rate, on the order of 65 to 70 % at 3 years, and 3-year OS estimates of only 25 %, with MRD being the dominant risk factor for adverse outcome [30]. Interestingly, the outcomes for adults with pre-transplant morphologically detectable disease mirror those for patients in morphologic CR with any level of disease detected by flow-MRD, suggesting that flow-MRD is a better tool for prediction of transplant outcome than morphologic CR [30]. If a patient has disease that is flow-MRD positive prior to transplant, checking for clearance of flow-MRD again post-transplant is not useful since only pre-transplant flow-MRD is a predictor of relapse [31]. Flow-MRD at various time points during and after induction and consolidation therapy as a marker for relapse has been confirmed in many studies [32–36].

Many questions surround flow-MRD in the evaluation of AML. One counterintuitive aspect of flow-MRD testing is that a substantial proportion of patients with detectable disease (20–30 %) do not relapse sometimes despite lack of further therapy, pointing to an unexpectedly and unacceptably high rate of false positives (the rate is variable depending on the conditions and study, but almost never zero). This limitation is one of the key challenges of trial design, since assigning interventions to a population that includes many already cured of their disease is problematic. Prospective studies are

required to determine whether driving flow-MRD positive patients to negativity is a strategy that will improve outcome, but if so, this endpoint would allow for much faster response assessment with a direct, quantifiable means to evaluate treatment response. Unfortunately, doing an allo-HCT and driving a flow-MRD-positive patient to negative post-transplant does not seem to be a promising strategy [31]. In one study [31], patient's flow-MRD positive pre-transplant had similarly poor relapse rates regardless of whether they became flow-MRD negative post-transplant. Because of this, only pre-transplant but not post-transplant flow-MRD was independently associated with OS and relapse risk.

Real-Time Quantitative Polymerase Chain Reaction

The appeal of PCR-based MRD testing is the higher sensitivity and the greater ease of standardized testing. But it shares a disadvantage with flow-MRD: both immunophenotype and mutational composition can be altered from diagnosis to evaluation of treatment response, so a search for the initial abnormalities could be wasted effort. As next-generation sequencing becomes more widespread, the dynamic nature of the mutational landscape in AML is more apparent [9, 37–39]. It is entirely possible to see one subclone disappear, only for another to re-emerge in a game of clonal whack-a-mole. And gene selection makes a “positive vs. negative” interpretation fraught with difficulty, owing to the hierarchical structure of the mutational landscape in AML. For example, the discovery of molecular MRD by real-time quantitative polymerase chain reaction (RT-PCR) of mutations that is typically a late-occurring genomic event in the progression of AML (e.g., FLT3-ITD, NPM1) would likely be a more ominous finding than a mutation that could be indicative of a pre-leukemic stem cell (e.g., DNMT3A) [40, 41]. Patients can have disease that persists with readily detectable pre-leukemia stem cells for prolonged periods, but certain mutations are only found at the time of relapse or shortly preceding relapse [39, 40]. It has also been established that pre-leukemia clones can expand greatly after induction therapy and herald relapse [39, 40, 42]. Therefore, when a patient is MRD-positive by mutation analysis, the situation is not simply a yes or a no; it must be interpreted within the context of the AML biology. Unfortunately, only a fraction of all potential individual mutations are known that can signify a pre-leukemia stem cell or overt disease.

It is also worth noting that the mutation-detection technology is heterogeneous in many respects, but importantly, with respect to detection limit. Next-generation sequencing typically has a defined limit of 1–10 % allele frequency [43] and may not capture mutations present while a patient is in remission. In a recent study [44], persistently detectable NPM1 mutations in the blood by RT-PCR using a mutation-specific primer with

a common primer and probe were found to be associated with risk for relapse. Patients with persistently detectable circulating NPM1 mutant cells after the second chemotherapy cycle represented 15 % of the group and their 3-year risk of relapse was significantly higher (82 vs. 30 %; hazard ratio, 4.80; 95 % confidence interval (CI), 2.95 to 7.80; $P < 0.001$), with a lower rate of survival (24 vs. 75 %; hazard ratio for death, 4.38; 95 % CI, 2.57 to 7.47; $P < 0.001$). Interestingly, there were higher rates of detection of MRD in the bone marrow than in peripheral blood, but after taking into account blood MRD status following the second chemotherapy cycle, no other measurement of MRD provided additional prognostic value, and this was the only independent prognostic factor for death in a multivariate analysis (hazard ratio, 4.84; 95 % CI, 2.57 to 9.15; $P < 0.001$). These results were validated in an independent cohort. NPM1 mutations were detected in 69 of 70 patients at the time of relapse. Of note, even in patients with concomitant higher-risk mutations, such as FLT3-ITD and DNMT3A mutations, a negative result on the NPM1 RT-qPCR assay of peripheral blood after the second chemotherapy cycle was associated with a 3-year survival rate of 70 %. Conversely, slow clearance of MRD in patients with cytogenetically normal AML and NPM1 mutation without higher-risk features, normally a favorable-risk group, had a very poor outcome as evidenced by the vast majority relapsing within 2 years. It is tempting to use this data to predict who might benefit from allo-HCT; but in this study, patients with MRD-positive disease did not enjoy improved survival with transplant, with the caveat that the number of patients in the analysis was small. This study corroborated an earlier one showing NPM1 qRT-PCR mutation MRD after two induction cycles prognosticated remission duration and OS [45] and another that demonstrated a worsened risk of relapse and OS associated with detection of persistent leukemia-associated mutations in at least 5 % of bone marrow cells in day 30 remission samples [46].

Despite the advances in increasing the detection limit, even in a disease such as chronic myeloid leukemia with a highly reliable marker of the leukemia clone (BCR/ABL), there is a still a >50 % rate of false-negative PCR tests when used to predict eradication [47]. And in AML, founder mutations in DNMT3A and IDH genes can often persist for prolonged periods even in patients with long-term remissions, indicating that elimination of pre-leukemia stem cells may not be necessary for disease control [39–41], even though it is associated with increased relapse risk. These data suggest that caution is warranted in the interpretation of results of MRD by RT-PCR, despite the promising results.

Quality of Life

Precedent exists for drug approval for management of other myeloid neoplasms on the basis of substantial improvements

in quality of life (QoL) and ruxolitinib is a good example. The COMFORT-I and COMFORT-II trials using ruxolitinib [48, 49] showed a significant decrease in disease related symptoms in patients with myelofibrosis. The RESPONSE trial showed similar results with symptom improvement when used treating patients with polycythemia vera [50]. The increase in QoL was largely due to reduction in symptomatic splenomegaly and cytokine-mediated symptoms that are frequently a direct consequence of the disease process. While patients with AML certainly have decreased QoL as a result of their disease, these are distinct in that the disease itself typically does not cause symptoms, rather, the result of immune dysfunction and iatrogenic complications (typically in the form of allo-HCT) are the culprits.

Replacing or restoring bone marrow function is an active area of research and if successful, could lead to significant decreases in infectious complications and transfusion requirements. The short lifespan of granulocytes and lack of ability to culture hematopoietic stem and progenitor cells *ex vivo* are major barriers to success. Other strategies to replace the innate immune system have had conflicting or disappointing results, such as the use of G-CSF [51–53] or prophylactic granulocyte transfusions [54]. Antibiotic, antiviral, and antifungal prophylaxis has been evaluated with a wide spectrum of results [55–59], but the emphasis in these trials has typically leaned on infectious endpoints without primarily focusing on QoL directly. This is a rational approach since an increase in QoL without a decrease in “harder” endpoints like infections would be difficult to explain. Considering infectious complications represent a major reason for hospital admissions, pain, and reduction in quality and quantity of life, this is an area that might have untapped potential.

Alleviating transfusion dependence (platelets and red blood cells) is another area of active investigation, but is limiting in the setting of uncontrolled AML. The vast majority of drugs studied to treat AML target leukemia cells with the goal of inducing death specifically in that population, yet, for reasons previously discussed, they almost always cause collateral damage to normal hematopoietic stem and progenitor cells. More and more, evidence suggests that the leukemia cells can alter the bone marrow stem cell niche to their advantage [60–67], and this could be targetable. If the environment were tipped to favor normal hematopoietic stem and progenitor populations, one could imagine an improvement in functional bone marrow output without achieving remission status. None of our current endpoints would capture this improvement except survival and potentially QoL, although there is a chance neither would be significantly prolonged due to reasons previously discussed.

The need for transfusion of red blood cells could potentially be decreased by the use of longer lasting oxygen-carrying blood substitutes, long-lasting recombinant hemoglobin, and many other techniques. Many academic centers and for-profit

medical research companies are working to produce safe and effective means to synthetically replace red blood cells, which would be particularly useful in combat or emergency situations, among others. However, after a meta-analysis of 13 trials demonstrated an increased risk for death and myocardial infarction with some of these agents [68], the safety of these blood substitutes was more critically evaluated, and it is unclear what the future holds for this field. Bypassing the need for platelets seems a much taller order, but nanoparticles might be able to provide a partial substitute. Animal data exists showing nanoparticles that can mimic key attributes of platelets and potentially decrease bleeding [69, 70]. The safety and efficacy have not yet been tested on humans, but given the problem of platelet alloimmunization, this could be a very useful adjunct.

Bone pain associated with active AML is evident in a minority of patients, but can be quite debilitating and difficult to treat when present. There are groups that have made headway tackling this difficult issue, but as of yet, no therapy has been evaluated in a clinical setting for this purpose. Since bone pain is a feature of many malignancies that might share a common mechanism, a QoL improvement might be feasible without resorting to opioids and their side effects.

It is worth noting that QoL could also be significantly changed by other factors outside of a regulatory approval process. A “neutropenic diet” is commonplace in many centers on the hypothetical basis that uncooked foods might increase rates of infection, but clinical trial data have not entirely endorsed this idea [71]. The extent to which we should restrict the lifestyle of our patients is an area that can have significant effects on QoL, but suffers from lack of robust data (and funding for research) to back it.

Extending the quantity of life of our patients with AML has been challenging; improving the quality of life should be another important goal. It is probably not commonly used as an endpoint due to difficulties in assessment caused by heterogeneity in disease course (one infection can change everything) and the fact that the quality of life decrease imposed by AML is mainly due to non-uniform, secondary effects that are largely unpredictable on an individual patient level. The unpredictable nature requires a large population to demonstrate a benefit in a study. Most of the investigational drugs we use inhibit the normal bone marrow progenitors as well, so using transfusion need as a surrogate for QoL is challenging until we have different targets and mechanisms in trials that rely on bypassing bone marrow function or restoring it.

Is There an Ideal Endpoint?

Interventions that improve quality of life would be quite valuable, but measurement difficulties and large changes in each patient course due to different, largely unpredictable events

(on an individual level) make this challenging. Most of the issues with QoL in AML, besides transfusions, do not lend themselves well to investigational therapeutic intervention, apart from dealing with the AML itself. Nevertheless, many patients with AML spent countless days hospitalized receiving intensive remission induction strategies. Even if a new therapy did not improve survival, if an oral therapy could control AML equally well while allowing a patient to spend their time outside the hospital or decreased transfusion needs, it might dramatically improve QoL. Although EFS is not a perfect correlate with survival, it may be a better assessment of a new drug's efficacy since the endpoint is less affected by subsequent uncontrolled, potentially biased variables and interventions. EFS is an endpoint hampered by non-uniform definitions of events and often inadequate correlation with OS. CR, EFS, or both may correlate with better with quality of life given their tighter correlation with bone marrow function. MRD by flow cytometry and/or by high-resolution mutational evaluation is likely the best way we currently have to measure the depth of response to a therapy and hopefully future studies will continue to make this a priority. MRD improvements must also be taken together with safety data for a more complete evaluation of efficacy. One could also imagine using EFS and defining "failure to achieve MRD" as an event for a combined approach, and this might represent an endpoint that is more reflective of a therapies immediate potential.

Overall survival data is the ultimate goal of any therapy and a sufficiently safe and effective therapy for AML will increase survival. A large OS benefit is typically sufficient to conclude benefit in a well-controlled trial, but the lack of OS increase does not exclude benefit. Using this as the bar for approval will have the fallout of excluding potentially helpful therapies where the benefit was later diluted by confounders. In addition, if, optimistically, a very clear benefit was seen in an early stage trial, designing a later phase trial that does not allow crossover might have serious ethical implications. But allowing crossover would likely make interpretation of the survival data difficult or impossible.

Taken together, there is no ideal endpoint and probably not even a "best" endpoint, considering all the factors at work. For now, we would recommend induction of remission with normalization of hematologic parameters over a pre-specified time period as the best marker for regulatory approval of new drugs. Ideally, new drugs will advance to the point of obviously and immediately improving every endpoint, making our quibbles of endpoint selection obsolete in the pages of history.

Compliance with Ethical Standards

Conflict of Interest Joshua P. Sasine declares no potential conflicts of interest. Gary J. Schiller declares research support and scientific advisory board activity from Sunesis Pharmaceuticals.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest published recently, have been highlighted as:

- Of importance
- Of major importance

- 1.•• Sasine JP, Schiller GJ. Emerging strategies for high-risk and relapsed/refractory acute myeloid leukemia: novel agents and approaches currently in clinical trials. *Blood Rev.* 2015;29(1):1–9.
2. Appelbaum FR, Rosenblum D, Arceci RJ, et al. End points to establish the efficacy of new agents in the treatment of acute leukemia. *Blood.* 2007;109(5):1810–6.
3. Willemze R, Suci S, Meloni G, et al. High-dose cytarabine in induction treatment improves the outcome of adult patients younger than age 46 years with acute myeloid leukemia: results of the EORTC-GIMEMA AML-12 trial. *J Clin Oncol.* 2014;32(3):219–28.
4. Lowenberg B, Ossenkoppele GJ, van Putten W, et al. High-dose daunorubicin in older patients with acute myeloid leukemia. *N Engl J Med.* 2009;361(13):1235–48.
5. Hernan MA, Hernandez-Diaz S, Robins JM. Randomized trials analyzed as observational studies. *Ann Intern Med.* 2013;159(8):560–2.
6. Ravandi F, Ritchie EK, Sayar H, et al. Vosaroxin plus cytarabine versus placebo plus cytarabine in patients with first relapsed or refractory acute myeloid leukaemia (VALOR): a randomised, controlled, double-blind, multinational, phase 3 study. *Lancet Oncol.* 2015;16(9):1025–36.
7. Kandoth C, McLellan MD, Vandin F, et al. Mutational landscape and significance across 12 major cancer types. *Nature.* 2013;502(7471):333–9.
8. The Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med* 2013;368:2059–74.
9. Grimwade D, Ivey A, Huntly BJ. Molecular landscape of acute myeloid leukemia in younger adults and its clinical relevance. *Blood.* 2016;127(1):29–41.
10. Schiller GJ. High-risk acute myelogenous leukemia: treatment today ... tomorrow. *Hematol Am Soc Hematol Educ Program.* 2013;2013:201–8.
11. Andre T, Boni C, Mounedji-Boudiaf L, et al. Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. *N Engl J Med.* 2004;350(23):2343–51.
12. Herndon TM, Demko SG, Jiang X, et al. U.S. Food and Drug Administration approval: peginterferon-alfa-2b for the adjuvant treatment of patients with melanoma. *Oncologist.* 2012;17(10):1323–8.
13. Estey E, Othus M, Lee SJ, Appelbaum FR, Gale RP. New drug approvals in acute myeloid leukemia: what's the best end point? *Leukemia.* 2016;30(3):521–5.
14. Lusk MR, Lee J-W, Fernandez HF, Lazarus H, Rowe J, Tallman M et al. Results of the ECOG E1900 trial in younger adults with AML using an event free survival endpoint are concordant with results based on overall survival: potential for a surrogate endpoint to facilitate rapid approval of therapies in AML. *Blood* 2014;124:2599.

15. Burnett AK, Hills RK, Hunter AE, et al. The addition of gemtuzumab ozogamicin to low-dose Ara-C improves remission rate but does not significantly prolong survival in older patients with acute myeloid leukaemia: results from the LRF AML14 and NCRI AML16 pick-a-winner comparison. *Leukemia*. 2013;27(1):75–81.
16. Kantarjian HM, Thomas XG, Dmoszynska A, et al. Multicenter, randomized, open-label, phase III trial of decitabine versus patient choice, with physician advice, of either supportive care or low-dose cytarabine for the treatment of older patients with newly diagnosed acute myeloid leukemia. *J Clin Oncol*. 2012;30(21):2670–7.
17. Dombret H, Seymour JF, Butrym A, et al. International phase 3 study of azacitidine vs conventional care regimens in older patients with newly diagnosed AML with >30% blasts. *Blood*. 2015;126(3):291–9.
18. Walter RB, Kantarjian HM, Huang X, et al. Effect of complete remission and responses less than complete remission on survival in acute myeloid leukemia: a combined eastern cooperative oncology group, southwest oncology group, and M. D. Anderson cancer center study. *J Clin Oncol*. 2010;28(10):1766–71.
19. Cheson BD, Bennett JM, Kopecky KJ, et al. Revised recommendations of the International working group for diagnosis, standardization of response criteria, treatment outcomes, and reporting standards for therapeutic trials in acute myeloid leukemia. *J Clin Oncol*. 2003;21(24):4642–9.
20. Dohner H, Estey EH, Amadori S, 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(3):453–74.
21. Przepiorka D, Ko CW, Deisseroth A, et al. FDA approval: blinatumomab. *Clin Cancer Res*. 2015;21(18):4035–9.
22. Freireich EJ, Gehan EA, Sulman D, Boggs DR, Frei 3rd E. The effect of chemotherapy on acute leukemia in the human. *J Chronic Dis*. 1961;14:593–608.
23. Todisco E, Ciceri F, Oldani E, et al. The CIBMTR score predicts survival of AML patients undergoing allogeneic transplantation with active disease after a myeloablative or reduced intensity conditioning: a retrospective analysis of the gruppo Italiano trapianto Di midollo Osseo. *Leukemia*. 2013;27(10):2086–91.
24. Duval M, Klein JP, He W, et al. Hematopoietic stem-cell transplantation for acute leukemia in relapse or primary induction failure. *J Clin Oncol*. 2010;28(23):3730–8.
25. Vidriales MB, Perez-Lopez E, Pegenaute C, et al. Minimal residual disease evaluation by flow cytometry is a complementary tool to cytogenetics for treatment decisions in acute myeloid leukaemia. *Leuk Res*. 2016;40:1–9.
26. Walter RB, Buckley SA, Pagel JM, et al. Significance of minimal residual disease before myeloablative allogeneic hematopoietic cell transplantation for AML in first and second complete remission. *Blood*. 2013;122(10):1813–21.
27. Walter RB, Gyurkocza B, Storer BE, et al. Comparison of minimal residual disease as outcome predictor for AML patients in first complete remission undergoing myeloablative or nonmyeloablative allogeneic hematopoietic cell transplantation. *Leukemia*. 2015;29(1):137–44.
28. Grimwade D, Freeman SD. Defining minimal residual disease in acute myeloid leukemia: which platforms are ready for “prime time”? *Blood*. 2014;124(23):3345–55.
29. Al-Mawali A, Gillis D, Lewis I. The role of multiparameter flow cytometry for detection of minimal residual disease in acute myeloid leukemia. *Am J Clin Pathol*. 2009;131(1):16–26.
30. Araki D, Wood BL, Othus M, et al. Allogeneic hematopoietic cell transplantation for acute myeloid leukemia: time to move toward a minimal residual disease-based definition of complete remission? *J Clin Oncol*. 2016;34(4):329–36.
31. Zhou Y, Othus M, Araki D, et al. Pre- and post-transplant quantification of measurable (“minimal”) residual disease via multiparameter flow cytometry in adult acute myeloid leukemia. *Leukemia*. 2016.
32. Freeman SD, Virgo P, Couzens S, et al. Prognostic relevance of treatment response measured by flow cytometric residual disease detection in older patients with acute myeloid leukemia. *J Clin Oncol*. 2013;31(32):4123–31.
33. Terwijn M, van Putten WL, Kelder A, et al. High prognostic impact of flow cytometric minimal residual disease detection in acute myeloid leukemia: data from the HOVON/SAKK AML 42A study. *J Clin Oncol*. 2013;31(31):3889–97.
34. Kern W, Voskova D, Schoch C, Hiddemann W, Schnittger S, Haferlach T. Determination of relapse risk based on assessment of minimal residual disease during complete remission by multiparameter flow cytometry in unselected patients with acute myeloid leukemia. *Blood*. 2004;104(10):3078–85.
35. Venditti A, Buccisano F, Del Poeta G, et al. Level of minimal residual disease after consolidation therapy predicts outcome in acute myeloid leukemia. *Blood*. 2000;96(12):3948–52.
36. Kohnke T, Sauter D, Ringel K, et al. Early assessment of minimal residual disease in AML by flow cytometry during aplasia identifies patients at increased risk of relapse. *Leukemia*. 2015;29(2):377–86.
37. Farrar JE, Schuback HL, Ries RE, et al. Genomic profiling of pediatric acute myeloid leukemia reveals a changing mutational landscape from disease diagnosis to relapse. *Cancer Res*. 2016.
38. Bodini M, Ronchini C, Giaco L, et al. The hidden genomic landscape of acute myeloid leukemia: subclonal structure revealed by undetected mutations. *Blood*. 2015;125(4):600–5.
39. Kim T, Yoshida K, Kim YK, et al. Clonal dynamics in a single AML case tracked for 9 years reveals the complexity of leukemia progression. *Leukemia*. 2016;30(2):295–302.
40. Shlush LI, Zandi S, Mitchell A, et al. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature*. 2014;506(7488):328–33.
41. Ploen GG, Nederby L, Guldberg P, et al. Persistence of DNMT3A mutations at long-term remission in adult patients with AML. *Br J Haematol*. 2014;167(4):478–86.
42. Wong TN, Miller CA, Klcó JM, et al. Rapid expansion of preexisting nonleukemic hematopoietic clones frequently follows induction therapy for de novo AML. *Blood*. 2016;127(7):893–7.
43. Luthra R, Patel KP, Reddy NG, et al. Next-generation sequencing-based multigene mutational screening for acute myeloid leukemia using MiSeq: applicability for diagnostics and disease monitoring. *Haematologica*. 2014;99(3):465–73.
44. Ivey A, Hills RK, Simpson MA, et al. Assessment of minimal residual disease in standard-risk AML. *N Engl J Med*. 2016;374(5):422–33.
45. Kronke J, Schlenk RF, Jensen KO, et al. Monitoring of minimal residual disease in NPM1-mutated acute myeloid leukemia: a study from the German-Austrian acute myeloid leukemia study group. *J Clin Oncol*. 2011;29(19):2709–16.
46. Klcó JM, Miller CA, Griffith M, et al. Association between mutation clearance after induction therapy and outcomes in acute myeloid leukemia. *Jama*. 2015;314(8):811–22.
47. Mahon FX, Rea D, Guilhot J, et al. Discontinuation of imatinib in patients with chronic myeloid leukaemia who have maintained complete molecular remission for at least 2 years: the prospective, multicentre stop Imatinib (STIM) trial. *Lancet Oncol*. 2010;11(11):1029–35.
48. Verstovsek S, Mesa RA, Gotlib J, et al. A double-blind, placebo-controlled trial of ruxolitinib for myelofibrosis. *N Engl J Med*. 2012;366(9):799–807.
49. Harrison C, Kiladjian JJ, Al-Ali HK, et al. JAK inhibition with ruxolitinib versus best available therapy for myelofibrosis. *N Engl J Med*. 2012;366(9):787–98.

50. Vannucchi AM, Kiladjan JJ, Griesshammer M, et al. Ruxolitinib versus standard therapy for the treatment of polycythemia vera. *N Engl J Med*. 2015;372(5):426–35.
51. Lehrnbecher T, Zimmermann M, Reinhardt D, Dworzak M, Stary J, Creutzig U. Prophylactic human granulocyte colony-stimulating factor after induction therapy in pediatric acute myeloid leukemia. *Blood*. 2007;109(3):936–43.
52. Sung L, Aplenc R, Alonzo TA, Gerbing RB, Lehrnbecher T, Gamis AS. Effectiveness of supportive care measures to reduce infections in pediatric AML: a report from the children's oncology group. *Blood*. 2013;121(18):3573–7.
53. Wheatley K, Goldstone AH, Littlewood T, Hunter A, Burnett AK. Randomized placebo-controlled trial of granulocyte colony stimulating factor (G-CSF) as supportive care after induction chemotherapy in adult patients with acute myeloid leukaemia: a study of the United Kingdom medical research council adult leukaemia working party. *Br J Haematol*. 2009;146(1):54–63.
54. Strauss RG, Connett JE, Gale RP, et al. A controlled trial of prophylactic granulocyte transfusions during initial induction chemotherapy for acute myelogenous leukemia. *N Engl J Med*. 1981;305(11):597–603.
55. Pagano L, Cairam M, Candoni A, et al. Evaluation of the practice of antifungal prophylaxis use in patients with newly diagnosed acute myeloid leukemia: results from the SEIFEM 2010-B registry. *Clin Infect Dis*. 2012;55(11):1515–21.
56. Vehreschild JJ, Bohme A, Buchheidt D, et al. A double-blind trial on prophylactic voriconazole (VRC) or placebo during induction chemotherapy for acute myelogenous leukaemia (AML). *J Infect*. 2007;55(5):445–9.
57. Annino L, Chierichini A, Anaclerico B, et al. Prospective phase II single-center study of the safety of a single very high dose of liposomal amphotericin B for antifungal prophylaxis in patients with acute myeloid leukemia. *Antimicrob Agents Chemother*. 2013;57(6):2596–602.
58. Kern W, Behre G, Rudolf T, et al. Failure of fluconazole prophylaxis to reduce mortality or the requirement of systemic amphotericin B therapy during treatment for refractory acute myeloid leukemia: results of a prospective randomized phase III study, German AML cooperative group. *Cancer*. 1998;83(2):291–301.
59. Bergmann OJ, Mogensen SC, Ellermann-Eriksen S, Ellegaard J. Acyclovir prophylaxis and fever during remission-induction therapy of patients with acute myeloid leukemia: a randomized, double-blind, placebo-controlled trial. *J Clin Oncol*. 1997;15(6):2269–74.
60. Glait-Santar C, Desmond R, Feng X, et al. Functional niche competition between normal hematopoietic stem and progenitor cells and myeloid leukemia cells. *Stem Cells*. 2015;33(12):3635–42.
61. Geyh S, Rodriguez-Paredes M, Jager P, et al. Functional inhibition of mesenchymal stromal cells in acute myeloid leukemia. *Leukemia*. 2016;30(3):683–91.
62. Cogle CR, Goldman DC, Madlambayan GJ, et al. Functional integration of acute myeloid leukemia into the vascular niche. *Leukemia*. 2014;28(10):1978–87.
63. Boyd AL, Campbell CJ, Hopkins CI, et al. Niche displacement of human leukemic stem cells uniquely allows their competitive replacement with healthy HSPCs. *J Exp Med*. 2014;211(10):1925–35.
64. Krevvata M, Silva BC, Manavalan JS, et al. Inhibition of leukemia cell engraftment and disease progression in mice by osteoblasts. *Blood*. 2014;124(18):2834–46.
65. Hanoun M, Zhang D, Mizoguchi T, et al. Acute myelogenous leukemia-induced sympathetic neuropathy promotes malignancy in an altered hematopoietic stem cell niche. *Cell Stem Cell*. 2014;15(3):365–75.
66. Bosse RC, Wasserstrom B, Meacham A, et al. Chemosensitizing AML cells by targeting bone marrow endothelial cells. *Exp Hematol*. 2016;44(5):363–377.e365.
67. Le Y, Fraigneau S, Chandran P, et al. Adipogenic mesenchymal stromal cells from bone marrow and their hematopoietic supportive role: towards understanding the permissive marrow microenvironment in acute myeloid leukemia. *Stem Cell Rev*. 2016;12(2):235–44.
68. Natanson C, Kern SJ, Lurie P, Banks SM, Wolfe SM. Cell-free hemoglobin-based blood substitutes and risk of myocardial infarction and death: a meta-analysis. *Jama*. 2008;299(19):2304–12.
69. Anselmo AC, Modery-Pawłowski CL, Menegatti S, et al. Platelet-like nanoparticles: mimicking shape, flexibility, and surface biology of platelets to target vascular injuries. *ACS Nano*. 2014;8(11):11243–53.
70. Bertram JP, Williams CA, Robinson R, Segal SS, Flynn NT, Lavik EB. Intravenous hemostat: nanotechnology to halt bleeding. *Science translational medicine*. 2009;1(11):11ra22.
71. Gardner A, Mattiuzzi G, Faderl S, et al. Randomized comparison of cooked and noncooked diets in patients undergoing remission induction therapy for acute myeloid leukemia. *J Clin Oncol*. 2008;26(35):5684–8.