Bone marrow evaluation for diagnosis and monitoring of acute myeloid leukemia

Mary-Elizabeth Percival\textsuperscript{a,b}, Catherine Lai\textsuperscript{c}, Elihu Estey\textsuperscript{a,b,*}, and Christopher S. Hourigan\textsuperscript{c,*}

\textsuperscript{a}Department of Medicine, University of Washington, Seattle, WA, USA
\textsuperscript{b}Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA
\textsuperscript{c}Myeloid Malignancies Section, Hematology Branch, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD, USA

Abstract

The diagnosis of acute myeloid leukemia (AML) can be made based on peripheral blood or bone marrow blasts. In this review, we will discuss the role of bone marrow evaluation and peripheral blood monitoring in the diagnosis, management, and follow up of AML patients. For patients with circulating blasts, it is reasonable to perform the necessary studies needed for diagnosis and risk stratification, including multiparametric flow cytometry, cytogenetics, and molecular analysis, on a peripheral blood specimen. The day 14 marrow is used to document hypocellularity in response to induction chemotherapy, but it is unclear if that assessment is necessary as it often does not affect immediate management. Currently, response assessments performed at count recovery for evaluation of remission and measurable residual disease rely on bone marrow sampling. For monitoring of relapse, peripheral blood evaluation may be adequate, but the sensitivity of bone marrow testing is in some cases superior. While bone marrow evaluation can certainly be avoided in particular situations, this cumbersome and uncomfortable procedure currently remains the de facto standard for response assessment.

Keywords

Acute myeloid leukemia; Bone marrow evaluation; Flow cytometry; Morphology; Measurable residual disease
1. Use of marrow in initial diagnosis

1.1 Bone marrow evaluation

When the diagnosis of acute myeloid leukemia (AML) is suspected, the treating physician typically recommends a bone marrow evaluation for further morphologic assessment. Indeed, the practice of morphologic assessment of the bone marrow is recommended in the initial diagnostic work-up for suspected AML by the European LeukemiaNet (ELN), the National Comprehensive Cancer Network (NCCN) Guidelines, and the World Health Organization (WHO) 2016 guidelines. Typically, both bone marrow aspiration and bone marrow trephine biopsy are performed. Some centers perform aspiration alone when possible, with biopsy only in cases of a “dry tap” or diagnostic uncertainty (e.g., distinguishing whether peripheral pancytopenia is related to AML or myelodysplastic syndrome); a biopsy is always indicated if there are no circulating blasts in the peripheral blood and AML is suspected. Analyses typically performed on the bone marrow sample include flow cytometry, metaphase cytogenetics, fluorescence in situ hybridization (FISH), and molecular analyses, which often utilize polymerase chain reaction-based or next-generation sequencing technology to examine specific markers (NPM1, CEBPA, FLT3) or panels of markers with demonstrated importance in myeloid malignancy. This combined information collected at the time of AML diagnosis is used in prognostication for patients overall, as well as in recommendations for individual patients regarding whether to proceed to investigational induction chemotherapy and/or allogeneic hematopoietic cell transplant (HCT).

1.2 Can AML be diagnosed and characterized without bone marrow evaluation?

The diagnosis of AML requires the presence of ≥20% blasts in the peripheral blood or bone marrow; in certain cases, the presence of recurrent cytogenetic abnormalities, such as the characteristic translocation (8;21) define AML even at a lower blast count. Frequently the diagnosis of AML can be made without resorting to invasive bone marrow sampling. For patients with high peripheral leukemic blast counts, many of the requisite tests can be performed on peripheral blood. The immunophenotype obtained by flow cytometry has been found to be the same in peripheral blood and bone marrow blasts, though the antibody panel used therein was relatively limited in scope and complexity in this relatively older study. A small case series confirmed this finding by comparing peripheral blood and bone marrow in patients with acute leukemias, which demonstrated no differences in morphology or immunophenotyping if the peripheral blood blasts were 30% or more. Differences have also been found in the cell cycle phase of blasts in these two compartments, though the clinical significance of this finding is yet to be determined. In contrast, the particular leukemia-associated immunophenotype (LAIP) for a particular patient may be variable and meaningful, particularly for later monitoring of residual disease.

In the small case series mentioned previously, the karyotype was insufficient for analysis in 17% of the AML peripheral blood samples (5 out of 29 patients), but in none of the bone marrow samples. However, no account was made by the authors of the total blood blast count (i.e., white blood cell count multiplied by percentage of blasts), only the blast percentage in peripheral blood. Since the comparison of karyotype between the blood and
bone marrow was performed only in patients with a high percentage of peripheral blasts, bone marrow aspiration for karyotype should be performed if few or no circulating blasts are present. Similar findings were demonstrated by the study of Hussein et al, in which patients with high numbers of circulating blasts (at least $0.1 \times 10^9$ cells/L) were likely to have successful peripheral blood karyotype (90% success rate or higher). In AML patients specifically, peripheral blood karyotyping produced successful metaphases in 32 out of 42 patients (76%). Conventional karyotyping with chromosome banding is the current standard in the diagnosis and work-up of a patient with AML, and several AML entities are defined by their recurrent karyotypic abnormalities. It is possible that next generation sequencing approaches will make conventional chromosomal banding redundant in the future.

FISH testing, typically performed as part of an AML or myelodysplastic syndrome-specific panel, is less clearly correlated between blood and bone marrow. In a review of 48 cases of AML with paired peripheral blood and bone marrow samples, abnormal peripheral blood FISH results were found in 69% of patients with abnormal bone marrow FISH results (18 of 26), but also in 23% of cases with normal bone marrow FISH results (5 of 22). There is uncertainty whether patients with abnormal cytogenetics as assessed by FISH but not standard 20-metaphase karyotype have a prognosis more befitting the FISH results or the standard results; pathologist consultation may help the treating physician in cases of discordance. Molecular analysis has similarly been compared between peripheral blood and bone marrow samples, and peripheral blood has high sensitivity and specificity for detection of FLT3-ITD and NPM1 mutations when the blast count is >2000 cells per microliter. More recently, the protein expression pattern (so-called proteome) has been shown to be closely correlated between peripheral blood and bone marrow blasts, though this finding is not clinically relevant at the current time.

Overall, if peripheral blood blasts are high at the time of AML diagnosis (>2000 cells per microliter), we posit that bone marrow examination is an unnecessary adjunct to peripheral blood sampling, which is able to provide morphologic, immunophenotypic, and molecular data; however, discrepancies still remain between the consistency of FISH results in peripheral blood and bone marrow.

1.3 Is morphologic assessment at diagnosis necessary?

Morphologic assessment of the bone marrow for the evaluation of acute leukemias was initially standardized through pathologic review by the French-American-British (FAB) cooperative group in 1976, and revised a decade later. While the small study mentioned previously showed good correlation between morphology in the peripheral blood and bone marrow, the question remains whether morphology is necessary at all for the diagnosis of AML. In fact, at times, the bone marrow may be inaspirable, making morphology moot. In limited subtypes, including acute megakaryoblastic leukemia, acute panmyelosis with myelofibrosis, and acute myeloid leukemia with myelodysplasia-related changes, some have argued that immunohistochemistry performed on a bone marrow biopsy is a crucial adjunct to peripheral blood analysis in order to make the final diagnosis. It should be noted that one way for the diagnosis of AML with myelodysplasia-related changes to be made in the
2016 WHO classification is with multilineage dysplasia (defined as >50% of cells with dysplasia in at least two cell lines), which can only be assessed on bone marrow sampling. It has recently been suggested that the presence of a mutation in any of SRSF2, SF3B1, U2AF1, ZRSR2, ASXL1, EZH2, BCOR, or STAG2 in AML is greater than 95% specific for the diagnosis of secondary AML suggesting the possibility that future revisions of the WHO classification will continue the trend to move away from morphological descriptors as surrogates of underlying etiology.

However, shortly after the FAB classification was released, flow cytometry was developed and rapidly incorporated into clinical diagnosis, allowing for precise surface marker characterization of acute leukemias at diagnosis and beyond. Furthermore, it is increasingly becoming evident that morphological and even immunophenotypic assessments cannot accurately reflect the diverse genetic etiology of this class of diseases, including even clonal heterogeneity within a single patient. Increasingly sophisticated molecular tools are able to better refine an AML diagnosis based on genetic abnormality, with the important caveats that 1) somatic mutations are not typically in themselves disease defining and can be seen in healthy older adults or patients in prolonged remissions after chemotherapy and that 2) such genomic profiling pre-treatment is currently suboptimal in predicting resistance to induction therapy.

2. Use of marrow in response assessment

2.1 Utility of the day 14 marrow

Soon after initial diagnosis with AML, fit patients are treated with intensive induction chemotherapy. The NCCN guidelines for AML recommend performing bone marrow evaluation 7-10 days after completion of induction chemotherapy, which is around day 14 if traditional 7+3 chemotherapy (combining continuous low-dose cytarabine with three days of an anthracycline) is used. If persistent or residual disease is identified (generally >5-10% blasts, though the background marrow cellularity may also be important), the recommendation is to administer more chemotherapy. A study of clinical flow cytometry examined a novel method to calculate the degree of cytoreduction with induction chemotherapy; when combined with knowledge that time to complete remission (CR) is an important prognostic factor, it seems intuitive that a day 14 marrow would provide important clinical information. Indeed, Liso et al. examined the prognostic value of a day 14 marrow in 198 de novo AML patients in an attempt to derive a predictive tool based on blast percentage. Similarly, a German study evaluated outcomes in 449 patients enrolled in the German AML Cooperative Group 1992 trial, and found that day 16 blasts as a continuous variable were significantly related to rates of CR and persistent disease, as well as to overall and relapse-free survival. It should be noted that patients in this study and similar studies from European groups received a double induction regardless of early marrow status and achievement of CR after first induction course. Additionally, there is wide variation when assessing blast clearance in the aforementioned and other studies, having a nadir bone marrow blast cut-off ranging from ‘too few to count’ to <22%. Patients who had blast cutoffs below these ranges had higher rates of CR and improved outcomes. However, intra-observer variability remains an important factor when evaluating hypoplastic marrows.
More recent studies have questioned the utility of the day 14 marrow, as the marrow results are not always correlated with level of disease and may not reliably predict achievement of CR.

In a retrospective analysis of 194 untreated AML patients, Hussein et al. found that day 14 marrow was highly sensitive in predicting CR (90% sensitivity), but did not predict overall survival. In fact, some patients with a high BM blast percentage at day 14 were still able to achieve CR at day 21 or day 28 without re-induction chemotherapy, a finding also seen in other retrospective analyses using morphology and even flow cytometry. The method of assessment of residual disease at day 14 or later by flow cytometry may also be important, evaluating a leukemia-associated immunophenotype (so-called LAIP) or a “different-from-normal” phenotype. Further, the treatment algorithm for patients with evidence of disease on a day 14 marrow is not standardized even by practitioners at a single institution, as summarized in a recent review.

2.2 Marrow sampling after induction

It is difficult to debate the necessity of an end of treatment bone marrow performed at count recovery after the first cycle of induction chemotherapy. Remission status is formally assessed around day 28-35, as the peripheral blood counts recover from induction chemotherapy. The definition of CR requires peripheral blood count recovery, generally defined as neutrophils > 1000/microliter, platelets > 100,000/microliter, and independence from red blood cell transfusion, along with a concomitant decrease in marrow blasts to <5%. Such criteria were first proposed in 1956 and were updated in 2003 and 2010, and are expected to be updated again within the next year in the 2017 ELN guidelines. However, despite patients being in CR by morphology, patients often have measurable residual disease (MRD) detectable by more sensitive flow or polymerase chain reaction (PCR)-based assays. Patients with MRD at the end of treatment or prior to transplant have similar outcomes as patients who have bone marrow blasts > 5%. Prospective randomized studies need to be performed evaluating patients in CR but with MRD to determine if further treatment improves outcomes (see section 3.2 Measurable Residual Disease below).

2.3 Use of marrow in patients receiving less intensive therapy

Increasingly, less intensive therapies are being used in the management of patients with AML who are considered not to be candidates for induction chemotherapy; these are typically “less fit” newly-diagnosed patients or relapsed/refractory patients who have not responded to conventional chemotherapy. While the ELN response guidelines do not specify required intensity of treatment, remission status is generally determined after one or two cycles of induction chemotherapy. Less intensive therapies, including hypomethylating agents such as azacitidine and decitabine, may take months in order to achieve CR. The frequency at which bone marrow evaluation should be performed is not clear; the AZA-AML-001 study, in which patients were randomized to azacitidine or a conventional care regimen, specified that peripheral blood and bone marrow aspirate/biopsy would be collected every second cycle beginning at cycle 3, but the authors did not comment on time to best response. A retrospective analysis of patients treated with three cycles of decitabine at MD Anderson suggested patients were more likely to achieve CR if they had a significant
(p < 0.05) reduction in the ratio of number of blasts to number of non-blasts in the bone marrow (Estey, unpublished data).

Targeted therapies, whether used alone or combined with intensive chemotherapy, are growing in importance, as investigators seek to exploit the molecular heterogeneity of the disease.64 FLT3 inhibitors have been under investigation for over a decade, but the multikinase inhibitor midostaurin is the first to show an overall survival benefit when studied in a randomized controlled fashion in combination with 7+3 chemotherapy in newly-diagnosed FLT3-mutated patients.65 While midostaurin has primarily been studied in combination with other drugs, the IDH inhibitors used in patients with IDH1 and IDH2 mutations have been used as single agents to date.66 Anecdotally, these drugs have led to a differentiation syndrome with high numbers of blasts seen both in the circulation and in the marrow after multiple weeks of therapy, a finding which seems to correspond to clinical response.67 Such phenomena are provocative, and require a reassessment of the standard timing of blood and marrow response assessments for patients receiving these novel targeted drugs. Simultaneously, some might argue that bone marrow assessment may be advantageous in these patients to avoid prolonged, expensive treatment without demonstrable clinical benefit.

3. Peripheral Blood Monitoring

3.1 Prognostic value of clearing peripheral blood blasts

One non-invasive marker to consider using in place of an invasive day 14 marrow would be kinetics of peripheral blood blast clearance in applicable patients. Clearance of leukemic blasts in the periphery has been correlated with day 14 marrow when analyzed by prospective daily flow cytometry in a small group of 30 patients; in 17 of 19 patients who had a decrease in peripheral blasts of > 2 logs by day 6 of therapy with induction chemotherapy, CR was achieved.68,69 In another study, time to blast clearance monitored by manual differentials in 162 AML patients receiving induction chemotherapy showed that early blast clearance (prior to day 6 of treatment) was able to predict for early marrow blast clearance, CR, relapse free survival, and overall survival.70 Similarly, a retrospective analysis by Elliott et al stratified relapse-free survival in 73 patients with de novo AML who ultimately achieved CR; time of peripheral blast clearance (at or before day 3, on days 4 or 5, and on day 6 or later) was highly significant, with early clearance associated with a relapse rate of 12.5% and late clearance with a much higher relapse rate of 78%.71 These findings were confirmed by an analysis examining peripheral blood blast clearance by more sensitive multiparametric flow cytometry in 130 AML patients.72 Mathematical modeling of the peripheral blood blast clearance has been performed by at least two groups to evaluate kinetics, and rapid blast clearance is strong and independent predictor of CR.72,73 Lacombe et al. evaluated the slope of blast cell decrease in each individual patient over the first four days of treatment, and the slope was strongly correlated with the achievement of CR and risk of relapse.72 Vainstein V et al. examined peripheral blood blast dynamics by modeling an exponential decay curve for 106 patients, and using this methodology calculated an area under the ROC curve of 0.79.73
A major limitation of examining peripheral blood blasts is that not all patients have circulating blasts at diagnosis. Changes in treatment course in response to rate of peripheral blood blast clearance have not been studied in a prospective fashion. An association between mutation clearance (as measured by next-generation sequencing of bone marrow) and clinical outcome has been reported in a retrospective cohort. It is possible that changes measured with high sensitivity tools, such as used for MRD, on peripheral blood samples during induction may provide early information regarding clinical response; a clinical trial is currently underway at the NIH to test this hypothesis (NCT02527447).

3.2 Measurable Residual Disease (MRD)

As discussed above, CR requires peripheral blood count recovery in addition to morphologic remission in the marrow with blasts <5%. Subtypes of CR include CR with incomplete platelet recovery (CRp) and CR with incomplete neutrophil recovery (CRI), and these subtypes have a significantly worse overall prognosis in terms of both response to chemotherapy and survival. It is possible to further risk stratify patients in a CR by using high sensitivity techniques to detect biomarkers associated with increased relapse risk. These can include flow cytometry, PCR for gene expression, PCR for abnormal gene sequence, and increasingly next generation sequencing. The 2017 ELN guidelines for the diagnosis and treatment of AML that are currently under development will move toward such MRD-based response criteria. That is, the most stringent definition of CR will require no evidence of MRD, as detected by multiparametric flow cytometry (MFC) or molecular techniques where appropriate for individual patients in the bone marrow. This change is made in response to the fact that post-induction factors, particularly MRD status, have a very strong correlation with outcomes after either further chemotherapy or after allogeneic HCT. MRD provides prognostic information independent of type of response to induction chemotherapy, which can be important in future treatment planning for younger and older patients. Additionally, presence of MRD is a critical factor in determining outcomes for AML patients following allogeneic HCT, to such a degree that patients with MRD, but morphologic remission, behave similarly poorly to those with active disease at the time of allogeneic HCT These observations have led investigators to suggest that “minimal” in the traditional definition of MRD should be replaced with “measurable,” since any detectable evidence of disease leads to a worse prognosis. The sensitivity of any particular MRD technique used is likely of lesser importance than issues of amount, type, and frequency of sampling; clonal heterogeneity and antigen drift; technical reproducibility; and interpretation and integration of such measurements into clinical care. The sensitivities of multiple targets assessed by MFC or PCR, which range from 1:100 to 1:200,000, have been comprehensively summarized by Hokland et al. Persistence of cytogenetic abnormalities for those patients in remission after therapy is known to be associated with worse outcomes, though this technique is not sensitive for the presence of residual disease. A number of more sensitive tools exist to detect MRD, many of which may be used on peripheral blood, lessening the need for invasive bone marrow sampling. No clear superiority of one MRD technology over another in AML has...
been proven, with typically at least one hundred fold improvement in sensitivity compared with morphology alone, however flow cytometry methods may suffer from greater variability between centers than molecular approaches. 81 PCR-based monitoring of disease in the peripheral blood has been used successfully to monitor for patients with favorable risk AML for translocation(15;17), inversion(16) and translocation(8;21) AML, 94-97 and more recently somatic mutations such as in NPM1. 98 The ELN performed extensive testing on expression based MRD using WT1. 99 Despite being expressed in approximately 90% cases of AML, it was overexpressed to a level useful for MRD monitoring in only around 50% of cases. This limitation may be mitigated, in part, by using a multiple gene approach as studied in patients receiving allogeneic HCT 59,100 and autologous HCT. 101 Though sampling of peripheral blood every three months is typically used for monitoring, different molecular aberrations may require more frequent testing or even bone marrow sampling, due to distinct differences in the doubling time of abnormal clones. 102,103

MRD assessment using flow cytometric techniques has traditionally been done on marrow samples, but there is increasing evidence that peripheral blood can be used to monitor for MRD. A recent study examined the cumulative incidence of relapse and 3-year overall survival for patients with MRD detected by immunophenotyping of the peripheral blood, and found that both differences were significant. Specifically, the cumulative incidence of relapse at 1 year for patients with peripheral blood MRD positivity was 89% vs. 29% (p<0.001). 104 Caveats include that the study included only 114 AML patients with paired bone marrow and peripheral blood samples, primarily at a single center.

Importantly, however, though each of these methods has shown that detection of disease is associated with a worse prognosis, early intervention for MRD-positive patients has not been studied in a systematic manner to demonstrate improvement in outcome when MRD is detected. There is provocative evidence in childhood AML that a risk-stratified approach based on genetic classification and MRD may improve outcomes. 105

In the future, there may also be a role for whole-genome or whole-exome sequencing to follow patients for MRD. In an analysis of comprehensive sequencing data for 50 patients with paired samples from diagnosis and remission, the 24 patients who had persistent leukemia-associated mutations had significantly worse survival. 74 Though the cost of large-scale sequencing has decreased considerably in recent years, concerns still remain about the interpretation and utility of the large amount of data generated for each individual AML patient. Additionally, the sensitivity and specificity of MRD on outcomes, however the MRD is detected, are such that it can be difficult to counsel individual patients about treatment planning; indeed, though a patient with MRD after induction chemotherapy may be more likely to relapse after allogeneic HCT than one without MRD, that same patient may be more likely to benefit from a graft-versus-leukemia effect than from more cytotoxic chemotherapy.

The optimal frequency of monitoring for the development of MRD is unknown at this time, and likely depends on the specific type of mutations that are identified, as discussed above, because of differences in both test sensitivity and leukemic clone doubling time. 98,102,103,106 Whether early intervention will be beneficial for relapsed disease is also unknown; for
example, an older study suggested that routine bone marrow examination was not beneficial during first CR, and it remains to be proven that early detection of MRD leads to improved survival outcomes. Given the current limitations of the technology, our practice is not to make clinical decisions on the basis of a single MRD result; our viewpoints regarding necessary times for bone marrow evaluation are summarized in Table 1.

4. Novel approaches to track disease burden

Bone marrow biopsies at diagnosis and count recovery time-points are currently still the “gold standard.” With the improvement of peripheral blood monitoring techniques, in combination with better imaging modalities, it may be possible to create a new standard for evaluating response to treatment. There are limited studies exploring imaging as a prognostic and predictive indicator of response and survival, likely related to cost and time needed to complete these studies. While provocative, these radiologic studies are not yet ready for incorporation into routine clinical practice.

4.1 FDG PET

Positron emission tomography (PET) is a functional imaging technique used to evaluate metabolic processes. In fludeoxyglucose (FDG) PET, a biologically active analogue of glucose is used as a tracer and is very sensitive at measuring glucose uptake as a function of metabolic activity. However, FDG PET is not specific for distinguishing inflammation secondary to tumor versus infection in the majority of cases. Interestingly, FDG PET has shown efficacy in visualizing extramedullary disease (EMD). In a small study of 10 patients, FDG PET was able to detect known EMD in 90% of patients and additional EMD in 60% with an SUV max range 2.1-8.1. Cribe et al. evaluated 26 patients with newly diagnosed AML in which FDG PET found 65% of patients to have EMD compared to 31% by clinical exam. There was a high degree of concordance with bone marrow response and FDG PET response at the end of treatment, with 4 of 6 patients achieving a PR on FDG PET but CR on bone marrow biopsy experiencing an early relapse. The utility of FDG PET is unknown, but given the sensitivity of the imaging, FDG PET may be useful as an adjunct at diagnosis for patients with EMD AML to determine extent of disease, and at the end of treatment to document response.

4.2 18F-FLT PET

18F-FLT PET may be more suitable for the evaluation of AML patients given that 3′-deoxy-3′-18F-fluorothymidine (FLT) is a thymidine analog that is resistant to in vivo degradation and accumulates in proliferating tissues, including rapidly dividing hematopoietic stem cells in the bone marrow. The first demonstration of FLT PET in AML patients showed higher rates of biodistribution in the bone marrow, spleen and EMD compared to normal healthy controls. In a pilot study of eight patients, 18F-FLT PET was used as an early assessment of treatment response. Eight newly diagnosed AML patients were treated with induction chemotherapy and completed 18F-FLT PET during therapy (range from 2-6 days from start of treatment). Patients with a CR showed SUV uptake < 2 while patients with resistant disease (RD) displayed SUV > 2. SUV mean and SUV max were also significantly lower in patients with CR compared to RD and normal controls had SUVs
similar to patients in CR.\textsuperscript{113} While the numbers in this study are too small to generalize to larger populations, it addresses an interesting question of using imaging as an early assessment tool of response. For patients who are not responding, it may be worth changing therapy early to avoid unnecessary toxicity from an unsuccessful regimen. Taken in combination with peripheral blood monitoring, there is potential to predict response minimizing the need for an invasive testing. Currently, ECOG-ACRIN Cancer Research Group is conducting a phase 2 study of FLT PET/CT at the time of the nadir bone marrow (days 10-17) in newly diagnosed AML patients being treated with standard induction chemotherapy (NCT02392429).

### 4.3 DCE-MRI

Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) provides global and functional imaging of bone marrow angiogenesis as compared to traditional MRI which uses radio waves in a magnetic field to identify anatomy. In the studies conducted, DCE-MRI has been examined as a predictor for overall survival using a calculated peak enhancement ratio (Peak) and quantification of vascularity (Amp). In 78 de novo AML patients, those with a low Peak and Amp at diagnosis had an improved disease free survival and overall survival compared to patients with a high Peak and Amp.\textsuperscript{114} Another study by the same group looked at DCE-MRI at day 0 and day 7 of chemotherapy and found that patients with a decrease in Peak values (decrease in angiogenesis compared to baseline) had a higher chance of achieving CR and longer disease free survival compared to patients that had an increase in Peak values (increase in angiogenesis compared to baseline).\textsuperscript{115} Similar to the results seen with FLT PET, imaging modalities have the potential to strengthen our current testing methods for response assessment.

Future imaging studies have the potential to answer some important outstanding questions before imaging technology can be incorporated into standard assessments for AML monitoring: 1) Which imaging technique most accurately reflects the total burden of disease? 2) When is the ideal time, during or after treatment, for imaging to take place? 3) Can imaging accurately predict which patients will go into a complete remission and measure depth of response? and 4) What is the role of imaging in surveillance?

### 5. Future directions

Bone marrow evaluation, therefore, remains an important adjunct to peripheral blood analysis in patients with AML, and perhaps always will since some patients do not have circulating peripheral blood blasts. However, we feel that some “standard” tests such as the day 14 marrow are of questionable importance in the management of AML patients and should not be incorporated into routine clinical practice. The recently published WHO guidelines for the diagnosis of AML still espouse morphology as the most important characteristic for the diagnosis and management of myeloid neoplasms,\textsuperscript{3} though more modern techniques may threaten the hegemony of morphology, as the biology underlying this diverse set of malignancies is better elucidated. While it is likely that highly sensitive tools will be increasingly used on peripheral blood for response assessment and to monitor for clinical relapse, at present, sampling of the bone marrow compartment remains an
important component of initial AML diagnosis and at the end of induction treatment in the majority of patients. Figure 1 summarizes the current recommendations for bone marrow evaluation, as well as areas for possible future modifications to the algorithm as methods for diagnosis and monitoring of AML are refined.

In an aging population experiencing an increased incidence in AML (de novo, progression from MDS, and treatment related), it is imperative that we consider the patient's ability, willingness, and pain threshold in continuing to do bone marrows. Table 2 summarizes the pros and cons of peripheral blood versus bone marrow sampling for diagnosis and monitoring of AML. Notably, clinical trials often include patients with the best performance status, and the findings generated by such patients may not hold true for the general population. Work on alternatives to bone marrow examination in AML will continue, primarily using sensitive assays on the peripheral blood and newer imaging technologies, but for now bone marrow evaluation remains an important diagnostic tool in the care of AML patients.

6. Practice points

• Morphologic diagnosis of AML, immunophenotyping, and karyotype can all be performed on a peripheral blood sample if the absolute blast count is > 2,000/μl.

• The day 14 marrow is of questionable clinical utility, since it is predictive of CR rates, but not of OS.

• Marrow sampling at count recovery after induction chemotherapy is critical for response assessment; patients with a morphologic CR but evidence of MRD have worse outcomes.

7. Research agenda

• The 2017 ELN guidelines for AML will include a response category of so-called “stringent” CR (i.e., CR without MRD); meanwhile, new techniques are being developed for monitoring MRD, including next-generation sequencing and array-based approaches.

• The recommended frequency of marrow sampling for less intensive therapies, such as with targeted inhibitors, is unclear.

• Novel imaging technology, using both PET and MRI, may have a role in future monitoring of AML patients, though many questions still remain about the predictive ability, utility, and cost.

Acknowledgments

Funding: This work was supported in part by the Intramural Research Program of the National Heart, Lung, and Blood Institute of the National Institutes of Health.
References


74. Kc B, DiNardo CD. Evidence for Clinical Differentiation and Differentiation Syndrome in Patients With Acute Myeloid Leukemia and IDH1 Mutations Treated With the Targeted Mutant IDH1 Inhibitor, AG-120. Clin Lymphoma Myeloma Leuk. 2016

Figure 1.
The possible future landscape for diagnosis and monitoring of acute myeloid leukemia at various time points during treatment and subsequent surveillance. The current schema follows the guidelines of the National Comprehensive Cancer Network and others. In the future, we posit that advances in flow cytometry and sequencing (and possibly imaging) may circumvent our current reliance on morphology and cytogenetics. Though the current sensitivity of bone marrow testing is generally 10-fold higher than in the peripheral blood, many tests may be done on peripheral blood only in the future. Timing of surveillance monitoring for measurable residual disease on the peripheral blood will likely depend on the abnormalities being followed for a particular patient.
Table 1

The authors' viewpoint on necessity of bone marrow evaluation at standard times during the course of AML diagnosis therapy.

<table>
<thead>
<tr>
<th>Time point</th>
<th>Is bone marrow evaluation necessary?</th>
<th>Why or why not?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnosis</td>
<td>Sometimes</td>
<td>If absolute peripheral blast count &gt;2,000/μl, peripheral blood is sufficient for diagnosis and characterization of AML (morphology, karyotype, and molecular)</td>
</tr>
<tr>
<td>Day 14 after induction</td>
<td>No</td>
<td>Enough discordance exists in the AML community about how to utilize this information in terms of further chemotherapy that we do not advocate performing a marrow until at least day 21 or at the time of count recovery</td>
</tr>
<tr>
<td>Count recovery after induction</td>
<td>Yes</td>
<td>MRD is extremely important from a prognostic standpoint, and sensitivity of peripheral blood studies is as yet insufficient to detect low levels of MRD</td>
</tr>
<tr>
<td>Monitoring for MRD</td>
<td>Possibly</td>
<td>The benefit of early intervention for relapse has not yet been established in AML, meaning that we do not typically monitor disease with routine bone marrow evaluation after the completion of planned treatment. Additionally, for gene expression PCR-based MRD monitoring, peripheral blood may be superior to bone marrow due to high background rates in the bone marrow</td>
</tr>
</tbody>
</table>

Abbreviations: AML (acute myeloid leukemia); MRD (measureable residual disease).
Table 2
Pros and cons of sampling from the peripheral blood and bone marrow in AML.

<table>
<thead>
<tr>
<th>Peripheral blood</th>
<th>Bone marrow</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pros</strong></td>
<td></td>
</tr>
<tr>
<td>• Easily accessible, especially for patients with semi-permanent intravenous catheters in place</td>
<td>• Higher sensitivity by 10× than expected from peripheral blood for MRD monitoring, except for gene expression PCR-based tests</td>
</tr>
<tr>
<td>• If absolute peripheral blast count &gt;2,000/μl, sufficient for diagnosis of AML and further characterization (morphology, karyotype, and molecular studies)</td>
<td>• Generally referred to in guidelines for treatment/monitoring; often required for clinical trials</td>
</tr>
<tr>
<td>• Higher sensitivity than bone marrow for MRD monitoring for gene expression PCR-based tests (e.g., WT1)⁹⁹</td>
<td></td>
</tr>
<tr>
<td><strong>Cons</strong></td>
<td></td>
</tr>
<tr>
<td>• Decreased sensitivity</td>
<td>• Painful procedure for the patient, which patient may not want to undergo with frequency</td>
</tr>
<tr>
<td>• Unclear role of monitoring for response with less intensive therapies, including hypomethylating agents and targeted inhibitors</td>
<td>• More expensive</td>
</tr>
<tr>
<td></td>
<td>• May require extra staff and monitoring (e.g., for conscious sedation)</td>
</tr>
<tr>
<td></td>
<td>• Unclear role of monitoring for response with less intensive therapies, including hypomethylating agents and targeted inhibitors</td>
</tr>
</tbody>
</table>

Abbreviations: AML (acute myeloid leukemia); MRD (measurable residual disease)