MRD evaluation of AML in clinical practice: are we there yet?

Sylvie D. Freeman[1] *and Christopher S. Hourigan*[2]

[1]*Clinical Immunology Service, Institute of Immunology and Immunotherapy, College of Medical and Dental Sciences, University of Birmingham, Birmingham, United Kingdom; and* [2]*Laboratory of Myeloid Malignancies, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD*

MRD technologies increase our ability to measure response in acute myeloid leukemia (AML) beyond the limitations of morphology. When applied in clinical trials, molecular and immunophenotypic MRD assays have improved prognostic precision, providing a strong rationale for their use to guide treatment, as well as to measure its effectiveness. Initiatives such as those from the European Leukemia Network now provide a collaborative knowledge-based framework for selection and implementation of MRD assays most appropriate for defined genetic subgroups. For patients with mutated-*NPM1* AML, quantitative polymerase chain reaction (qPCR) monitoring of mutated-*NPM1* transcripts post-induction and sequentially after treatment has emerged as a highly sensitive and specific tool to predict relapse and potential benefit from allogeneic transplant. Flow cytometric MRD after induction is prognostic across genetic risk groups and can identify those patients in the wild-type *NPM1* intermediate AML subgroup with a very high risk for relapse. In parallel with these data, advances in genetic profiling have extended understanding of the etiology and the complex dynamic clonal nature of AML, as well as created the opportunity for MRD monitoring using next-generation sequencing (NGS). NGS AML MRD detection can stratify outcomes and has potential utility in the peri-allogeneic transplant setting. However, there remain challenges inherent in the NGS approach of multiplex quantification of mutations to track AML MRD. Although further development of this methodology, together with orthogonal testing, will clarify its relevance for routine clinical use, particularly for patients lacking a qPCR genetic target, established validated MRD assays can already provide information to direct clinical practice.

Learning Objectives

- Update the overview of available AML MRD assays, including newer technologies
- Understand current evidence and recommendations for the use of AML MRD assays in clinical practice
- Understand key technical aspects for selection, standardization, and interpretation of different AML MRD assays

Case presentation

A 62-year-old man with a normal white blood cell count was diagnosed with cytogenetically normal acute myeloid leukemia (AML), and subsequent molecular studies for European Leukemia Network (ELN) 2017 good and adverse risk mutations, including mutated *NPM1*, were negative. He was fit enough to be treated with 2 courses of induction chemotherapy (standard UK National Cancer Research Institute [NCRI] AML treatment). Measurable residual disease (MRD) response was assessed by flow cytometry. He achieved a morphological complete remission (CR) but was MRD positive by flow cytometry (at 0.12%) after the second cycle. Because of deterioration in performance status, his comorbidities, and preference, further treatment options did not include transplant. He was known to have *IDH1 R132* and *DNMT3A* mutations and was considered for novel regimens or azacitidine. Should he be monitored if he has further treatment and how?

Introduction

AML is a disease that consists of multiple genotypes with inter-leukemic, as well as intraleukemic, heterogeneity that is influenced by treatment. Therefore, it is unsurprising that a range of MRD biomarkers and assay platforms have been generated and are relevant to enable personalized AML MRD monitoring.[1] However, this contributes to the challenge that MRD testing in 2019 continues to present (ie, deciphering how, when, and even whether it should apply to individual patients). From a survey in 2016-2017, 69% of U.S. leukemia physicians reported routine use of AML MRD, most commonly flow cytometry, followed by polymerase chain reaction (PCR) for mutated *NPM1*, but only 21% had implemented serial PCR for longer-term monitoring.[2] For those not incorporating MRD testing into patient management, cited reasons were lack of resources and uncertainty regarding the use of the results. Of note was the extent to which MRD-directed decision-making varied in hypothetical clinical scenarios. When asked about MRD test positivity postinduction in a patient being considered for transplant, responses were divided equally among recommending against a transplant, additional chemotherapy, and changing the conditioning regimen. This variability reflects the paucity of high-quality evidence currently available to inform such AML MRD–based decisions. It also highlights

Conflict-of-interest disclosure: C.S.H. has received laboratory research funding from Merck and Sellas. S.D.F. declares no competing financial interests.
Off-label drug use: None disclosed.

the need for a more nuanced approach to MRD test interpretation, taking into account that the available evidence from MRD testing differs between AML subtypes. For example, in younger patients with mutated-*NPM1* AML, postinduction MRD in the blood by quantitative polymerase chain reaction (qPCR) predicts outcome independently of other mutations, including *FLT3*,[3] and benefit from transplant (hazard ratio = 0.25 for overall survival).[4] However, in wild-type *NPM1* intermediate-risk AML (as for the patient in the above clinical case) or in older patients, although MRD positivity is also prognostic of poor outcomes,[5-7] the effect of additional intensified chemotherapy or transplant requires further evaluation.

It is perhaps easy to forget, with the natural enthusiasm for the prospects of novel MRD technologies, that the well-tested MRD platforms (ie, qPCR and flow cytometry) can provide sophisticated information for MRD levels in most patients, identifying those most likely to relapse with current treatment schedules across all AML risk groups. Clinical trials have assimilated these methods in real time, testing the effect of intensification, early intervention, and novel approaches to improve prognosis or evaluating MRD as an early surrogate for therapeutic efficacy. Outcome data from these trials, together with amalgamated experience in routine clinical practice, stepwise informs how to deal with the results of MRD tests to best help AML patients. In parallel, technical evolution and insights from leukemia biology, together with collaborative efforts for standardization, continuously progress MRD detection. This review focuses on the more recent information from MRD testing, including assay limitations and prospects, that, together with consensus recommendations, can guide current implementation.

Recent guidance from ELN and the U.S. Food and Drug Administration

Since 2018, guidance for the expanding application of MRD in AML has been proposed by 2 organizations. The ELN, through international collaboration, published a consensus document from expert AML MRD laboratories with recommendations for flow cytometric and molecular MRD assays in clinical practice.[7] This addressed some of the key factors for MRD measurement, including sampling, recommended approaches, time points, and thresholds of positivity. Perhaps as importantly, the document highlighted areas requiring further work programs to harmonize and progress assays for which efforts by participating laboratories are ongoing. The U.S. Food and Drug Administration (FDA), after a series of workshops over several years, has invited comments on its draft guidance for the use of MRD as a biomarker in regulatory submissions for hematological cancers.[8] In it, the FDA lists criteria for MRD assays (eg, that reporting MRD-negative results requires information on detection limits and, perhaps more debatable for certain assays, sensitivity should be ≥1 log below the cutoff for MRD positivity). Referring to AML, the guidance states that each selected MRD marker(s) should reflect the leukemia and not underlying clonal hematopoiesis (false positives). Additionally, data should be provided for the false-negative rate that might result from relapse from a marker-negative clone. The FDA guidance states that MRD could be used to stratify or enrich trial populations, a strategy applied for the RELAZA2 trial[9]; however, when MRD is a trial end point, any patient with missed MRD samples should be categorized as unresponsive in the analysis.

MRD assay considerations

As highlighted by the ELN and FDA articles, there are several important factors to consider when selecting the most appropriate of the current MRD assays and then interpreting results to guide clinical decisions for individual patients. These are the "S" factors:[10] specificity, sensitivity of the MRD marker(s), and its stability during AML progression. Further considerations are the context of AML subtype and, if known, associated relapse kinetics, sample type, stage taken, and extent of assay standardization. Another "S," price ($), is a practical inclusion on the list. Finally, the evidence for correlation with outcome in clinical studies is critical. Preferably, this clinical validation would include MRD testing in real time, as in clinical practice, and would be reproducible by other centers.

Specificity: true vs false MRD positives

Perfect specificity for an assay has been defined as the ability of a method to assess unequivocally the analyte in the presence of other components that are expected to be present. For AML MRD, the analyte is a biomarker of acute leukemia; however, whether the presence of this analyte results in relapse will depend on the test time point, subsequent interventions, and the competing risk of death from other causes. The "other components" consist of any cells/genetic material that are not acute leukemia, as well as artifactual assay background (such as nonspecific antibody binding and autofluorescence in flow cytometry or nonspecific primer binding in PCR). For example, MRD assays by flow cytometry and *WT1* RNA overexpression by qPCR (both applicable in the majority of AML patients) have intermediate specificity,, primarily as the result of analyte-type signals (aberrant immunophenotypes or *WT1* transcription activity) from normal cells, especially in regenerating marrows.

Acute promyelocytic leukemia (APL) is the paradigm for MRD monitoring in AML. The APL genetic driver from *PML-RARA* fusion (>95% of APLs) also provides the molecular MRD target of *PML-RARA* transcripts[11]; this is highly specific, because only treatment-resistant APL cells, including those that may, in time, be relapse initiating, will have this MRD marker. Similarly, other AML subtypes, including core-binding factor (CBF) (*RUNX1-RUNX1T1*, *CBFB-MYH11* fusions) and mutated-*NPM1* AMLs, have main genetic drivers that generate specific molecular MRD targets. These analytes are typically representative of the residual AML independently of coexisting mutations but, additionally, are stable markers during AML progression and, thus, strong predictors of relapse by longer-term qPCR MRD monitoring.[3,7,12,13] However, even for these, genetic evolution is a consideration for interpretation.[14] Mutated-*NPM1* AMLs may infrequently relapse as wild-type *NPM1*, albeit usually later (median of 43 months in a recent study), with associated coexisting and persisting clonal hematopoietic mutations, including *DNMT3A*.[15]

It is now apparent that clonal hematopoiesis of indeterminate potential (CHIP)-associated mutations, including the *DNMT3A* point mutation in the above clinical case, although frequent and stable in AML, can persist posttreatment at high levels, despite longer-term disease-free survival.[16-19] Therefore, these are insufficiently specific MRD markers for AML relapse, substantiating the FDA statement that a marker selected to assess MRD should "not reflect underlying clonal hematopoiesis."[8] However, this does not preclude the future utility of MRD assays tracking certain of these mutations, as has already been tested for *IDH* mutations,[20,21] particularly for the efficacy evaluation of appropriate targeted therapies, such as IDH1/2 inhibitors,[22,23] a treatment option in our clinical case. Moreover, clonal hematopoietic mutations that cooperate in the progression to acute leukemia could conceivably be monitored in the future as specific biomarkers for preleukemic activity of novel agents. Of

particular interest in this regard are CHIP mutations in DNA damage-response genes (eg, *TP53* and *PPM1D*) that are enriched after cytotoxic therapy and associated with an increased risk for developing leukemia.[24] Allelic burdens of *DNMT3A*, *TET2*, and *ASXL1* mutations, including as circulating tumor DNA, may also have potential utility postallogeneic stem cell transplantation to track ablation of patient clonal hematopoiesis.[25,26]

Sensitivity: true vs false MRD negatives

With the prerequisite of an adequate representative sample, sensitivity for MRD assay targets ranges from 10^{-2} (current next-generation sequencing [NGS] mutation profiling) to 10^{-5} to 10^{-6}. The latter is achieved by the established qPCR assays for which the target has high transcript expression (such as *NPM1* exon 12 insertion mutations). Current MRD assays cannot test for AML eradication, but reduction of MRD target below the lower limit of detection/quantification at treatment time points (complete remission without minimal residual disease, CR_{MRD-} by ELN criteria[27]) indicates AML clearance of up to 4 logs greater depth than morphology. Not surprisingly, this significantly improves prognostic discrimination in patient cohorts for survival, as well as relapse. Younger adults in CR or CR with incomplete blood count recovery (CRi) after their first course of induction in the NCRI AML17 trial had a 5-year survival of 52%; if also categorized as CR_{MRD-} by flow cytometry (sensitivity of 10^{-4}), survival increased to 63% overall and to 70% when excluding poor-risk patients.[5] Observed clinical false negatives (by relapse frequency) from single MRD assessment time points are 20% to 30% in good- or intermediate-risk patients in multiple MRD studies, even for the most sensitive MRD assays.[3,28] Monitoring at several time points, such as after each chemotherapy cycle and when applicable sequentially from end of treatment (such as in the Figure 1 schema), captures more information and, consequently, reduces false negatives.[5,7,12,13]

Sample considerations for MRD interpretation

It is a sine qua non that a good quality bone marrow (BM) is most likely to be representative of residual AML for the majority of patients without extramedullary disease, hence, the present recommendation of a first-pull BM for almost all AML MRD assays to reduce false negatives from suboptimal sensitivity. Leukocyte numbers, cell viability (affected by transit time), hemodilution, and hypoplasia all contribute to limiting the sensitivity/lower level of quantification for any BM sample, independently of the theoretical assay sensitivity. Flow cytometric assays can assess and should incorporate information on all of these factors in the MRD report. MRD negativity from an antibody combination testing 250 000 leukocytes will only reach a detection limit of 0.01% (with a 20% coefficient of variation [CV]). Molecular assays include housekeeping gene (*ABL*) copies as control for nucleated cell numbers, but current assays cannot differentiate hemodilution and cell type unless performed on presorted cells. Single-cell assays in the future may be able to combine phenotype, as well as RNA expression and mutation profile,[29] but these currently are very expensive and can only evaluate small cell numbers.

Is blood informative for AML MRD?

Despite blood (peripheral blood [PB]) providing lower sensitivity than BM (1 log less for mutated *NPM1*[3]), measuring mutated-*NPM1* transcripts in PB postinduction is highly prognostic for the mutated-*NPM1* subgroup.[3,4] Interestingly the "false-negative" relapse risk is not increased for MRD negativity in blood vs BM.[4] This implies that, in the ~25% of mutated-*NPM1* patients with only BM positivity postinduction, the MRD is at a level concomitant with clearance by consolidation or is from nonleukemic-initiating more mature BM cells. However, BM is recommended for maximal sensitivity during later sequential monitoring, because BM positivity by qPCR targets usually precedes that of blood, providing an increased time window for any interventions.[7]

Reduced sensitivity may be compensated for by a differential increase in specificity when testing blood for flow cytometric MRD[30] (because of fewer normal progenitors/precursors and, therefore, less "noise"), as observed when measuring *WT1* levels.[31] Relapse-free survival[32] and overall survival[30,33] appear to be significantly better for patients with MRD-negative blood samples postinduction and postconsolidation in heterogeneous smaller cohorts. These encouraging results merit further evaluation in older and intermediate/poor-risk AML patients, most of whom lack sensitive qPCR molecular markers. There is also preliminary evidence that PB clonal profiles may be representative of BM during treatment,[34] as well as at diagnosis, at least for higher frequency mutations. In a small series of decitabine-treated patients, NGS (non–error-corrected) measurements of variant allele frequencies (VAFs) (at >5%) for mutation profiles in PB samples with <60% lymphocytes correlated well with paired BM results.[35]

How we would monitor: CBF AMLs and mutated-*NPM1* AMLs

Patients with CBF or mutated-*NPM1* AML, by the nature of their leukemic-specific fusion transcripts (*RUNX1-RUNX1T1*, *CBFB-MYH11*) or insertion/deletion mutations (*NPM1*), can be monitored by qPCR MRD assays that combine high sensitivity with specificity and are extensively validated. Monitoring schedules (such as in Figure 1) and reporting of results for comparability should be based on ELN guidelines,[7] although these are only evidence based for younger patients.[36] After treatment completion, qPCR MRD assessments are recommended for ≥2 years at 3-month intervals when the prior result was negative or revealed low copy numbers. Average kinetics from molecular relapse to clinical relapse ranges from >3 to 4 months for *CBFB-MYH11* to 2 to 3 months for *RUNX1-RUNX1T1* or mutated-*NPM1* when *FLT3*-ITD positive.[37] There are no new data to support a survival benefit from preemptive intensification in CBF AMLs, despite the clear association between MRD status (postinduction/consolidation or off-treatment) and clinical progression for younger patients.[7,38] Published studies indicate that allogeneic transplantation can be avoided in *FLT3*-ITD–positive mutated-*NPM1* younger adults when these patients remain in complete molecular first remission following standard induction,[7] but it is uncertain whether concomitant treatment by *Flt3* inhibitors alters relapse kinetics and, therefore, optimal off-treatment sampling intervals.

qPCR assays have been developed for other fusion transcripts, such as from rearrangements involving *KMT2A* (*MLL*), *NUP98*, and *NUP214*. Although these can track response (Figure 1), there are insufficient data to provide guidance for their use; hence, it is recommended to also assess response by flow cytometric MRD postinduction.[7]

What about patients with *FLT3* internal tandem duplications?

Given the frequent prevalence in AML patients, the high probability and rapid kinetics of relapse, and the continued development of effective targeted therapy, there is great interest in monitoring of AML clones containing *FLT3* internal tandem duplications (FLT3-ITDs). *FLT3*-ITD mutations represent late events in leukemic development and, hence, are not always detectable at relapse, particularly after

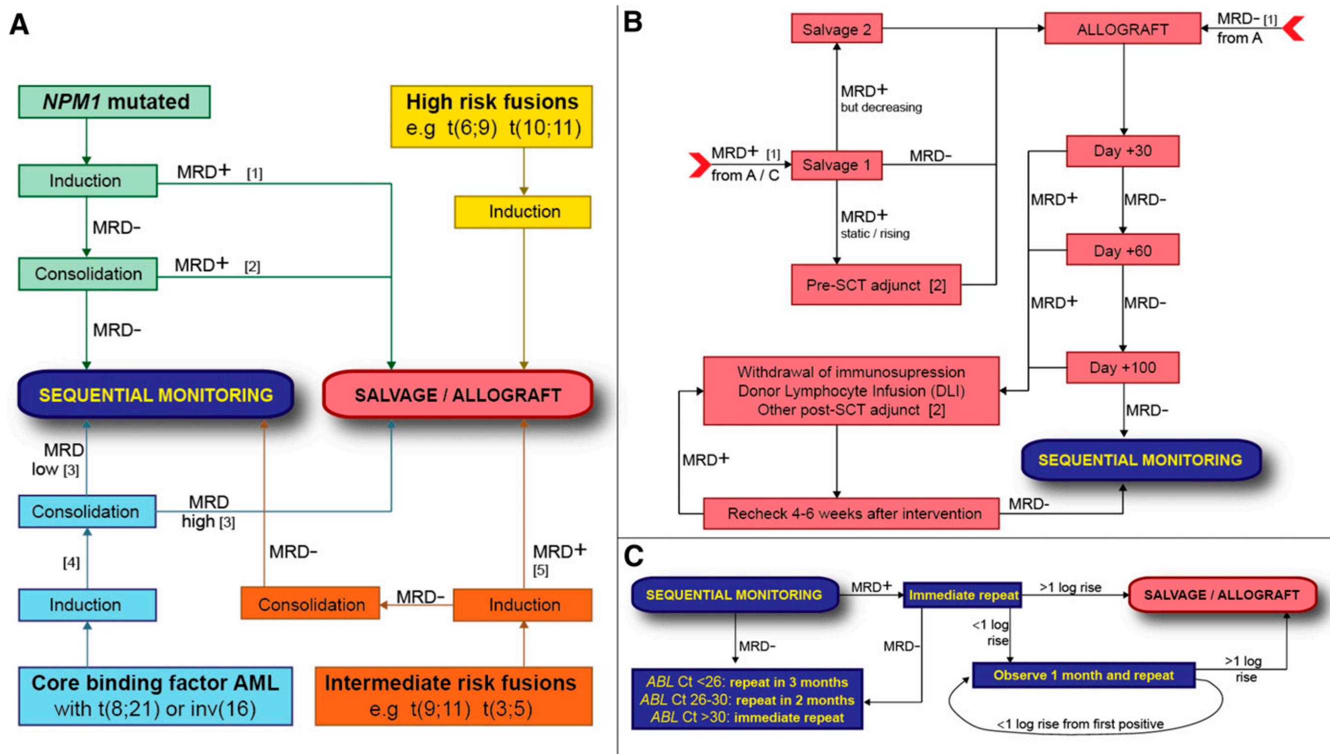


Figure 1. (A) Suggested management algorithm for patients with AML with a molecular MRD target. [1] ELN favorable risk patients with <4-log reduction in *NPM1*-mutant transcripts after first induction are shown to benefit from a CR1 allograft,[4] and any positivity in the peripheral blood (PB) after second induction is associated with a very high risk for relapse.[3] [2] MRD positivity > 200 copies per 10^5 *ABL* (ie, molecular persistence) and serially increasing transcript levels after treatment (ie, molecular progression) reliably predict relapse.[7] [3] At the end of treatment, patients with CBF AML with high or serially increasing transcript levels are destined to relapse (relevant thresholds are >500 copies [per 10^5 *ABL*] of *RUNX1/RUNX1T1* in the BM or >100 copies in the PB, and > 50 copies of *CBFB/MYH11* in the bone marrow [BM] or >10 copies in the PB).[7] Salvage according to (C) should be considered for these patients, although there is no evidence that this improves outcome. Conversely, patients with low copy numbers below these thresholds can be safely monitored according to (B). [4] Although CBF patients with an early unfavorable MRD response have a higher risk for relapse, there is insufficient evidence to warrant treatment change[7]; however, this may prompt initiation of an early donor search. Salvage may be considered in cases with extremely poor early response or if there is an increase while on treatment (ie, molecular progression). [5] Although there is no evidence that standard-risk patients who remain MRD positive benefit from transplant, this is a reasonable approach and is adopted in the current NCRI AML19 and MyeChild01 protocols. (B) Suggested algorithm for sequential monitoring after treatment. Patients with conversion to MRD positivity, confirmed on a second sample with >1-log rise, should be diagnosed with molecular relapse and treated as shown.[7] (C) Possible peri-transplant management strategy. [1] Patients with an *NPM1* mutation without *FLT3* ITD who have transcript levels below 1000 copies in the BM or 200 copies in the PB have a very good outcome after allograft, it is uncertain whether these patients benefit from salvage chemotherapy.[89] [2] Patients with high levels of MRD after salvage, without an adequate response to donor lymphocyte infusion (DLI), as well as those to whom these standard therapies cannot be given should be considered for investigational approaches. Figure by Richard Dillon, NCRI Group.

targeted therapy; they are often "replaced" by another signaling variant (eg, *RAS*, *Kit*, or a different *FLT3* variant).[39,40] Despite the limitation of potential false-negative tests, *FLT3*-ITD MRD testing has utility, because a positive result in an AML patient otherwise thought to be in remission is highly suggestive of MRD and is associated with a high likelihood of relapse, often with a short lead time.[41-43] *FLT3*-ITD mutations consists of nucleotide sequence inserts of variable length and location between patients, making 1 universal approach to low-level quantitative assessment by conventional PCR and bioinformatic mapping of data from NGS challenging. However, this technical constraint has been mitigated by novel PCR methods[44] and sequencing approaches.[39,41,43,45] In ~50% of adult AML cases, *FLT3*-ITD mutations will co-occur with a more stable AML MRD marker, such as mutated *NPM1* or t15:17[46]; therefore, it is recommended to also track such markers,[7] particularly for patients receiving *FLT3*-ITD–directed therapy. For those without a mutation other than *FLT3*-ITD to track, the expression of *WT1* is known to be elevated in these patients,[31,47] although the use of this AML MRD target for routine testing remains controversial. Flow cytometric MRD detection is recommended for response evaluation and monitoring for those patients with AML that cannot be tracked by a validated qPCR MRD assay.

Newer molecular technologies in AML MRD detection

The past decade of focus on cancer genomics has elucidated the genetic basis of AML and provided a wide range of molecular targets suitable for new drug development, as well as potentially for leukemic disease burden tracking. Although such approaches are highly publishable, several limitations prevent direct translation to the clinic at present. Unlike conventional quantitative PCR (Figure 2A), digital PCR, developed in the 1990s, allows absolute quantification of abnormal DNA sequences by partitioning each template molecule for the PCR reaction into an individual compartment (Figure 2B). Advantages of this technology include high sensitivity as the result of a low background error rate, highly accurate quantification due to elimination of template competition, limited bioinformatic requirements, and a rapid "sample-to-result" turnaround time.[28]

Disadvantages include the need to develop and validate assays for each individual target sequence, limited multiplexing capacity, and the inability to perform discovery on serial samples (eg, to detect selection of an independent *TP53*-mutated clone during therapy), making it unsuitable to detect relapses associated with clonal evolution. This technology is likely to be most useful for orthogonal validation of other technology and for tracking of common highly conserved "hotspot" variants, such as those seen in *NPM1* and *IDH1/IDH2* genes.[20,48]

DNA sequencing, typically consisting of "panels" covering the small regions of the genome known to be often somatically mutated in AML patients, has increasingly become part of the standard initial diagnostic workup, given the importance of such mutations in risk stratification and therapy selection[27] (Figure 2C). Because detected mutations are thought to originate from the leukemic clone and may persist during remission in patients at an increased risk for relapse,[17] it may be tempting to use the same NGS test for AML MRD detection. Unfortunately, as suggested by the absence of recommendations on NGS in the current ELN AML MRD guidelines, most NGS panels currently used at the time of AML diagnosis are not fit for the purpose of MRD detection because of limited sample input, insufficient read-depth, and lack of appropriate error correction (Table 1). NGS "AML panels" at diagnosis are used to determine likely somatic variants typically with a VAF (also known as mutant allele frequency) $\geq$ 5% (ie, 5 variant reads per every 100 at that position). Accordingly, those tests use genomic DNA inputs as low as 10 ng, which represents <8000 total cells (typical range of such tests, 10-100 ng). Given losses during NGS library preparation, a requirement to detect 3 to 5 unique copies of each variant (each representing a heterozygous mutation from a unique cell) to call MRD, and Poisson distribution sampling considerations at the limit of detection (such that the FDA guidance recommended the MRD threshold be 10-fold higher than the theoretical limit of detection), it is understandable that most high-quality AML MRD NGS research studies have used DNA inputs in the 200- to 500-ng range (ie, $\geq$1 mL of marrow aspirate or blood). Similarly, most AML MRD approaches using NGS, in recognition of the massive number of false-positive results, particularly for single nucleotide variants when considering potential variants < 2% VAF, have used some form of error correction in the form of laboratory techniques (eg, use of unique molecular identifiers [UMIs]), followed by consensus clustering and/or bioinformatic techniques, including models incorporating nucleotide position-specific background error rates resulting from PCR-based library preparation and sequencing itself[17,49] (Figure 2D). The sequencing depth requirement is often substantially higher for AML MRD than the one routinely used for diagnostic testing, because, in addition to the increased sample input, each unique DNA molecule typically will be sequenced 5 or 10 times in UMI-based error-corrected sequencing. Finally, the prognostic significance of AML "MRD," as detected by NGS, may not be equivalent to that detected by qPCR or flow cytometry; indeed, there is already considerable evidence that detection of some somatic mutations may be more prognostic than others.[17,50,51]

In contrast, AML MRD targets for qPCR have already been extensively validated and have been approved by the ELN consensus guidelines.[7] This recently led to the development of an RNA sequencing assay capable of simultaneously detecting all of these molecular targets in a single standardized NGS workflow while maintaining a limit of detection comparable to qPCR.[52] Future NGS AML MRD assays may combine RNA- and DNA-based approaches.

Update for immunophenotypic MRD

The ELN article provides a framework for the harmonization/standardization of flow cytometric MRD detection.[7] Consensus guidance in this covers the most reliable of extensively tested markers for tracking leukemic aberrant immunophenotypes, use of 8-color panels (acceptable in clinical laboratories to increase single-cell information), the importance of harmonizing cytometer settings for comparable antibody-stained cell profiles, and some technical recommendations for sample processing. Knowledge of the diagnostic Leukemic Associated ImmunoPhenotypes (LAIP) prevents reporting MRD negativity in the 5% to 10% of marker-negative patients (no identifiable LAIP at diagnosis or at follow-up), whereas a different-from-normal (DfN) analysis can detect phenotypic changes from leukemia evolution.[14] Consequently, it is advised to combine information from both diagnostic LAIP and DfN analysis to minimize false negatives. Defining MRD positivity by $\geq$0.1% of leukocytes (1 log greater than the limit of detection) is proposed, because this threshold is relevant in most studies. However, the proviso for this is that some patients with quantifiable MRD at <0.1% may have residual AML that is prognostic, for example, resulting from MRD underrepresentation by hemodilution, or potentially due to smaller, but more chemo-resistant, amounts after consolidation,[53,54] or upfront higher-intensity treatment,[55] or in older patients[6]. Risk discrimination from low-level MRD may also depend on genetic subgroup; detectable, but <0.1% flow cytometric, MRD after a first induction course was associated with a significantly increased relapse risk in mutated-*NPM1* and CBF AMLs but not in wild-type *NPM1* intermediate-risk patients.[5]

Feasibility of harmonization/standardization

Experience from acute lymphoblastic leukemia flow cytometric MRD work has shown that multicenter standardization[56] can be achieved with interpretive discordance reduced among experienced laboratories by feedback schemes.[57] Recent efforts by some multicenter flow cytometry laboratory networks, such as in Germany and France, are showing that this is also feasible for AML MRD, despite differences in cytometers and so forth. Figure 3 shows an example of an implemented harmonization strategy (Adriana Plesa and Christophe Roumier, on behalf of the Acute Leukemia French Association [ALFA], written communication, 10 May 2019). The ELN is testing a consensus standardized tube as a template to simplify interlaboratory set-up and comparability for multicenter interpretation.[58]

Immunophenotypic leukemic stem cell assays

Weak/negative CD38 expression on $CD34^+$ cells ($CD34^+CD38^-$) identifies progenitors that are enriched for functional hematopoietic stem cell activity in normal BM. High frequencies of $CD34^+CD38^-$ cells in AML at diagnosis (variability shown in Figure 4) have an independent adverse prognostic impact,[59-61] consistent with this immunophenotypic subpopulation as a biomarker for leukemic stem cell (LSC)-like chemoresistance in some AMLs. Leukemic $CD34^+CD38^-$ cells can express aberrant markers[62] that provide flow cytometric "LSC" targets with higher assay specificity due to less background from normal counterparts than with bulk progenitors. The approach has been refined and standardized for a "different-from-normal" strategy by construction of a single "LSC" tube that combines multiple $CD34^+CD38^-$ aberrant "LSC" markers.[63] Immunophenotypic quantitation of an immunophenotypic LSC

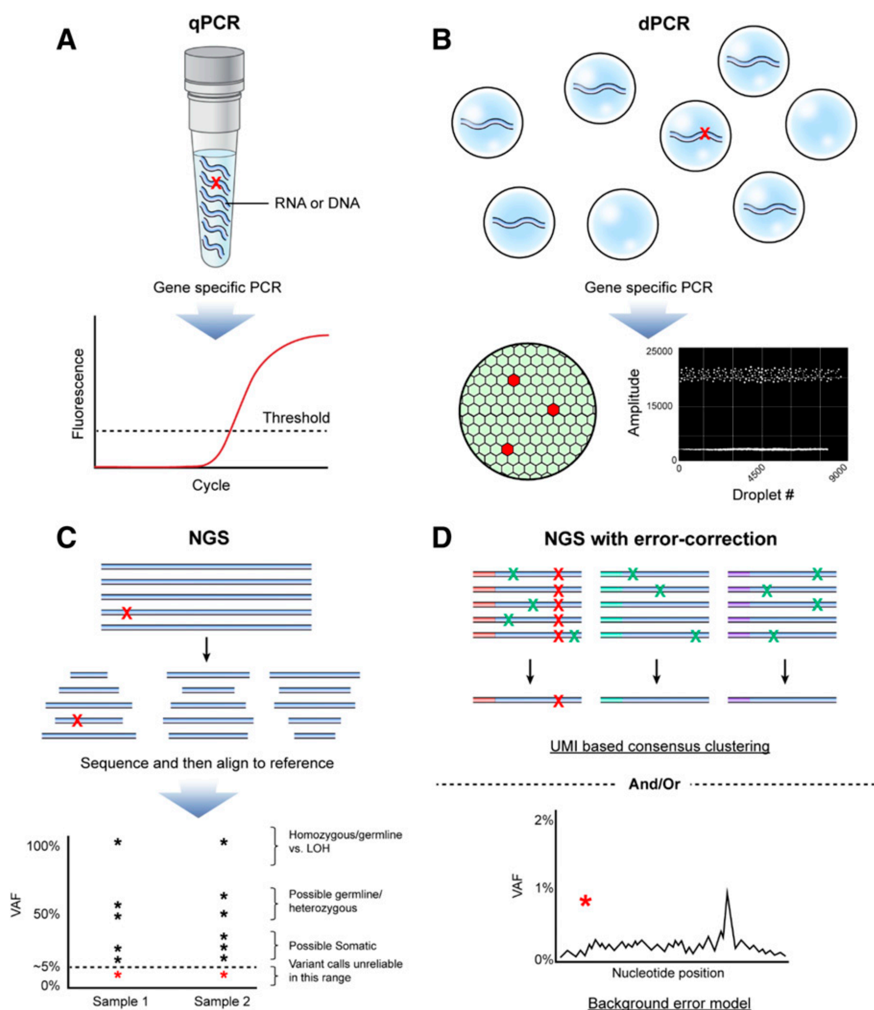


Figure 2. (A) qPCR is a common method for quantification of nucleic acid with real-time monitoring of the amplification of target of interest (eg, variant sequence shown with red X). Advantages include ubiquitous presence in most clinical laboratories, fast turnaround time, high sample throughput, and broad dynamic range. Disadvantages include limited number of suitable targets/assays available, relative lack of multiplexing ability, need to validate each target/assay individually, potential for false-negative results because of sample impurity, and limited ability to accurately discriminate between very low levels of target as seen in MRD. (B) Rather than performing the PCR reaction in "bulk," digital PCR partitions the template of interest into individual compartments (top), improving the performance compared with qPCR because of the lower background error rate (lower right), elimination of template competition, and digital result output, allowing absolute quantification (lower left). Lack of deep multiplexing ability and the need to validate each target/assay individually remain limitations. (C) NGS has revolutionized the initial clinical diagnostic evaluation of AML by allowing for simultaneous evaluation of multiple target regions typically selected from those known to be often mutated in AML. NGS is useful for discovery of mutations present in the range from 5% to 100% of a sample (VAF). However, not all variants detected will be pathogenic somatic mutations, and care should be taken to consider the possibility of identification of homozygous or heterozygous germline variants, as well as loss-of-heterozygosity (LOH) events. Variant discovery below a VAF of 5% using panels designed for profiling variants at diagnosis is challenging because of the lack of sensitivity and high false-positive rates. Red asterisks represent low-level variant calls that should be regarded with particular caution as within the range of background error for conventional NGS. (D) NGS for AML MRD performed in recent high-quality research studies has typically included error correction (upper), by incorporation of UMIs, followed by consensus determination of true (red X) variants vs false positives introduced by the technique (green X) and/or bioinformatic approaches to model background error rates at each nucleotide position in those not having a variant and determine the probability that the observed variant is a true positive (red asterisk) (lower). Figure by Erina He, National Institutes of Health Medical Arts.

population prior to allogeneic transplant was prognostic for relapse-free survival.[64] Moreover, immunophenotypic LSC frequency in CR had significant additive prognostic value to standard MRD by LAIP or mutated-*NPM1* qPCR in a large cohort, primarily as a result of increased specificity for very poor outcomes in the ~10% of patients with positivity for LSC and standard MRD.[65] Although increasing assay sensitivity to 1 in a million by testing more cells could reduce the false-negative relapse frequency observed in this study (3-year cumulative incidence of relapse [CIR] of 35% for double LSCneg/MRDneg), weak/negative CD38 may not be an appropriate or stable marker for relapse-initiating LSCs in a subset of patients.[66,67]

Future perspectives in immunophenotypic MRD

Ongoing discovery of leukemic aberrant phenotypes expands and rationalizes the repertoire of informative markers that can inform future routine AML MRD antibody panels,[68] as well as immunotherapy targets. For example, IL1RAP has been identified by

Table 1. Limitations for use of NGS to detect residual disease in AML

Problem	Significance	Solution
NGS AML panels used at diagnosis are unfit for AML MRD	Insufficient sample input/sequencing depth limits assay sensitivity. Insufficient error correction increases false-positive rate at low VAF.	NGS currently suited for research but not clinical use in AML MRD ELN guidelines on use of NGS in AML MRD coming
Association between variant tracked and residual AML clone(s)	Variants closely linked to AML (eg, FLT3-ITD) often subclonal/unstable. "Stable" variants, such as those in DMNT3A, also seen in ARCH/CHIP.	Tracking a panel of gene regions in remission will likely improve predictive power, at increased cost. Larger datasets will provide more information on which variants detected at MRD stage are the most associated with subsequent relapse.
Genetic clonal heterogeneity of AML = risk of false-negative tests	Variants detected at diagnosis are not necessarily present in the AML clone remaining after unsuccessful treatment responsible for relapse.	Tracking a panel of gene regions in remission will likely improve predictive power, at increased cost. Deep profiling of diagnostic sample to screen for minor subclones (eg:, *TP53*) may have utility.
Error rates intrinsic to NGS = risk of false-positive tests	Traditional NGS approaches cannot reliably identify novel variants present at low VAF (<2-5%) with sufficient specificity (ie, many false positive variant calls mask rare true positive variant).	ECS using UMI consensus clustering and/or bioinformatic approaches, such as background error models (Figure 2D), are helpful. Low VAF variants seen in diagnostic sample or multiple surveillance samples more likely true variants.
Correlation of NGS results with other measures of AML MRD	Tests designed to detect residual disease in AML may classify patient sample differently based on modality used (qPCR vs flow cytometry vs NGS). No single test represents a "gold standard" for detecting, in patients in remission, those cells that will subsequently lead to AML relapse.	Studies designed to integrate information from different AML MRD tests performed on the same sample cohorts are underway.
Lack of uniform reporting standards	How many UMI reads needed to call a variant (3, 5)? How many distinct UMI read families per variant needed to call MRD? How many genomic equivalents as input? Standardized filtering, consensus clustering, and variant calling needed? What about controls, duplicates, platforms?	NGS currently suited for research but not clinical use in AML MRD FDA guidance for MRD in hematological malignancies published in draft form. ELN guidelines on use of NGS in AML MRD coming

ARCH, age-related clonal hematopoiesis; CHIP, clonal hematopoiesis of indeterminate potential (an alternative term for ARCH); ECS, error-corrected sequencing.

proteomic[66] and single-cell RNA sequencing[29] strategies as a discriminatory marker, particularly in *FLT3-ITD* AML.

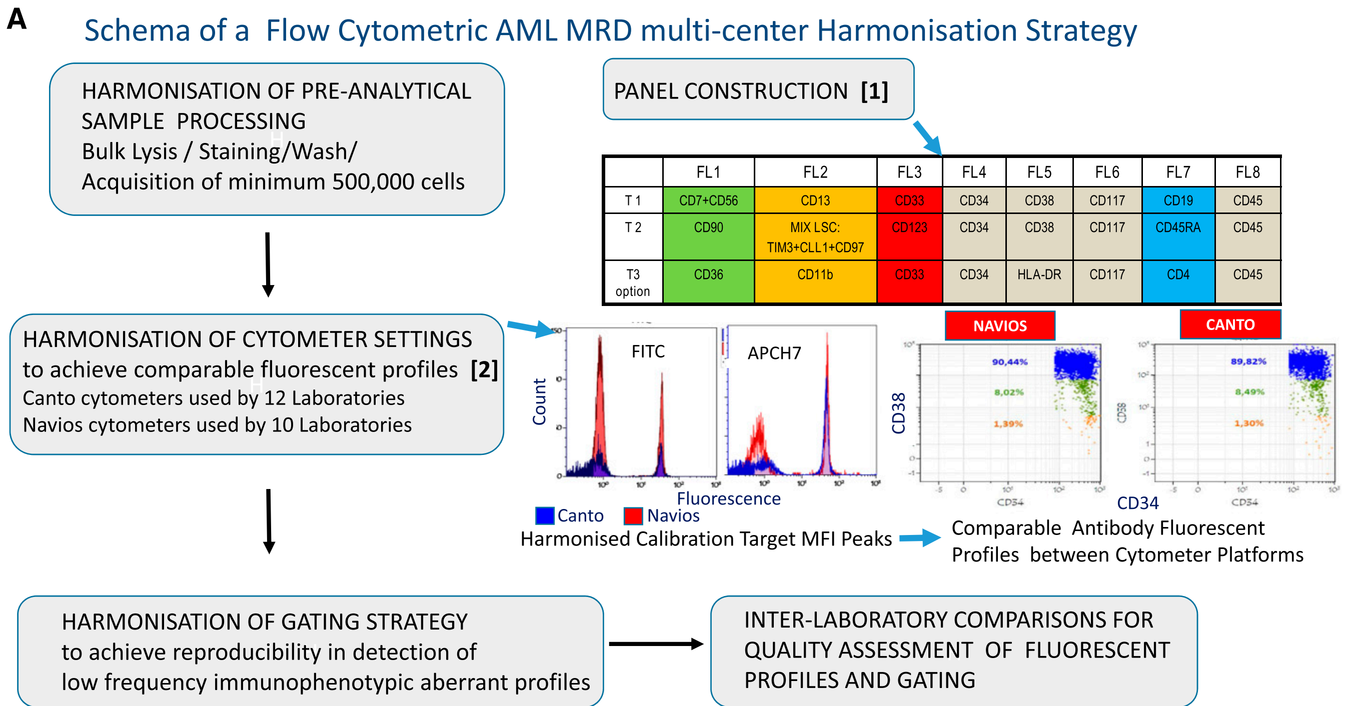
High-dimensional immunophenotyping using cytometry by time of flight (CyTOF), when available, allows deeper and/or broader population coverage because of increased parameters (>40) compared with that possible with flow cytometry. When combined with novel data-analysis algorithms,[69] CyTOF data can reveal biologically important leukemic and immune cell subpopulations that are missed by traditional approaches. Although CyTOF has proven to be a powerful discovery tool, its slow acquisition rates (500 events per second) preclude applying the technology to clinical laboratory assays that require high cell numbers, including MRD assays. However, because individual samples can be metal barcoded, there may be potential for multiplexing of diagnostic samples, achieving higher throughput and reducing data collection variation. This approach could screen for the most informative markers to guide construction of improved flow cytometric MRD panels or to enable personalized profiling for immunotherapeutic targets. The analysis algorithms developed for high-dimensional immunophenotyping have also been applied to flow cytometric data[70] and may, with further development, provide tools for unsupervised or partially unsupervised AML MRD analysis in the future.[71]

Immunophenotyping vs molecular MRD detection: friend or foe?

Initial evidence points to molecular and flow cytometric MRD providing complementary independent information at the same time points during active treatment, particularly when the molecular assays have lower sensitivity and specificity.[17,18,21,72,73] After induction, ~30% of patients in 2 studies[17,18] had discordant results between flow cytometric and standard NGS MRD (excluding the main CHIP mutations), and they had an intermediate risk for relapse (~50% at 4 years) in the larger cohort.[17] To understand how best to improve and deploy these assays, it will be helpful in the research setting to define whether their respective false-positive/negative results cluster in certain AML subtypes and whether there is any overlap. Of particular interest in this regard are AMLs without a sensitive qPCR target, including those with greater clonal complexity, such as the chromatin-spliceosome group.

Flow cytometry vs morphology to assess remission

Flow cytometric assays continue to provide the fastest turnaround time for evaluating blast percentages and MRD so that routine results may be available concurrently with, if not before, morphology reporting. Interobserver variability for morphological BM blast counts has been reported recently as greatest for blast percentages from 2% to 10% (and only moderate agreement for <2% and >10%), whereas reference percentages calculated from digital trephine images correlated well with flow cytometric blast percentages of first-pull BM.[74] Evaluation of leukemic aberrant immunophenotypes by MRD antibody panels adds specificity and sensitivity for discriminating leukemic blasts from normal blasts. Adults[75,76] and children[77,78] who are refractory by morphology, but MRD negative by flow cytometry after first induction, have a good prognosis (60% 3-year survival in 2 adult trial cohorts[75]), equivalent to those in CR_{MRD-}. Also of note is that, in a series of 87 patients morphologically categorized as relapsed, none were MRD negative



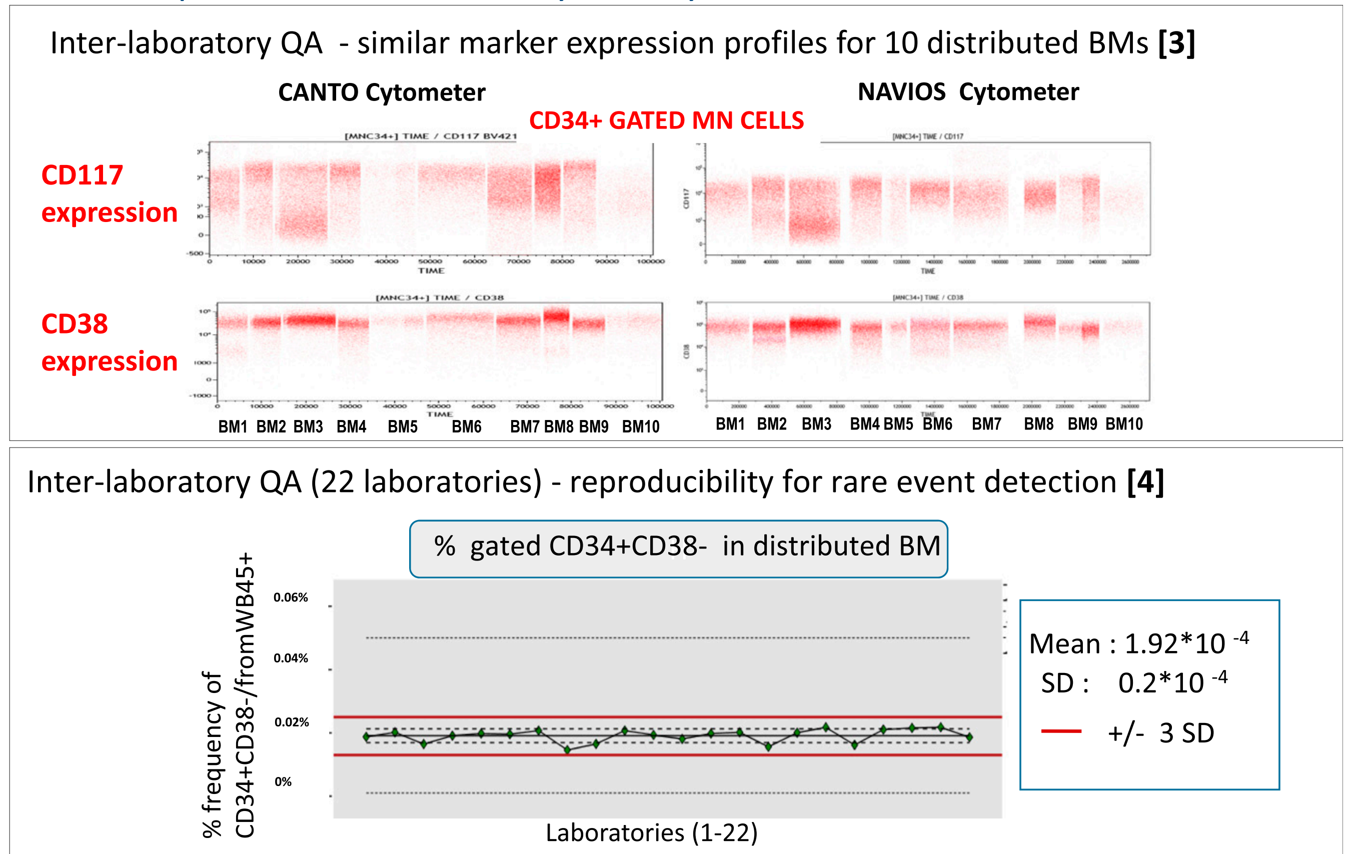


Figure 3. (A) Strategy applied for flow cytometric AML MRD multicenter harmonization by the ALFA Intergroup. [1] Rationale of AML MRD flow panel design was based on simplicity, reproducibility, and cost. Tube 1 was a core combination for LAIP detected at diagnosis and/or by different-from-normal analysis, tube 2 was targeted to aberrancies of $CD34^+CD38^-$ cells (immunophenotypic LSCs), and tube 3 was an optional development tube for monocytic aberrancies. [2] Flow cytometer fluorescent settings were harmonized ("mirrored") between Canto vs Navios cytometer platforms. Voltages

by flow cytometry.[79] Together, these results support refining the present criteria for refractory and relapsed AML to incorporate flow cytometric MRD-negative results when a validated assay is available.

Because MRD positivity in CR and morphologic refractory disease appear equivalent for outcomes preallogeneic transplant[53,80] (with the caveat of unavoidable selection bias), does blast enumeration for the criteria of resistant disease and partial remission add prognostic information? In younger adults treated in the NCRI AML17 trial, patients in partial remission (International Working Group criteria) or MRD-positive CR after a first course of induction had an equivalent intermediate prognosis (~40% 5-year survivals), with the exception of MRD-positive patients with incomplete count recovery, who had a much poorer prognosis (19% 5-year survival). Patients with resistant disease had a similar outcome to the latter group.[5] Thus, when flow cytometric MRD is incorporated into response assessment after first induction, the prognostic effect of ≥5% blasts by morphology appears to be restricted to the subgroup with resistant disease.

Treatment changes based on MRD detection

MRD status after induction

In the case of mutated-*NPM1* AML,[4] and potentially for wild-type *NPM1* intermediate-risk AML,[5] exploratory analyses support directing allogeneic transplant in first remission only to those patients testing positive for MRD postinduction chemotherapy (by qPCR for mutated-*NPM1* or flow cytometry for wild-type *NPM1*). However, when ELN 2017 adverse-risk patients are included, the benefit from allogeneic transplant as consolidation appears to be at least equivalent between MRD-negative and MRD-positive patients.[81,82] This supports allogeneic transplant in first remission as the best approach for adverse-risk patients, even when achieving MRD negativity early in treatment.

The extent to which outcome is altered by intensifying consolidation in those who are MRD positive at the postinduction time point will be further informed by forthcoming data in trials implementing this strategy in younger adults (eg, NCRI AML17/19 and HOVON 132 for ELN intermediate risk) and in older adults (NCRI AML18).

MRD status after consolidation

The CETLAM AML12 and the GIMEMA AML1310 phase 2 trials have investigated adjusting allogeneic transplant allocation by postconsolidation MRD levels in addition to genetic risk. Initial reports from these trials[83,84] show the feasibility of real-time MRD treatment stratification and, in the GIMEMA study, disease-free survival was similar between MRD-positive and MRD-negative intermediate genetic risk patients. The low patient numbers in the analyses of the effect of transplant directed by postconsolidation MRD reinforces the difficulty of testing nonbiased transplant interventions.

Monitoring off treatment

A suggested MRD-directed treatment algorithm for AML patients with stable molecular qPCR MRD targets is shown in Figure 1. When patients remain low risk by MRD levels after induction and consolidation, sequential 3-month qPCR MRD monitoring is recommended to allow a greater time window for treatment decisions, because an increase in target transcript levels (as defined by the ELN[7]) precedes clinical relapse.

For AML patients without a qPCR MRD target, the role of off-treatment sequential monitoring requires further evaluation. However, similar to MRD status after induction and consolidation, pretransplant and potentially posttransplant MRD testing can inform the risk of poorer outcomes.

MRD status pretransplant

Although it is clear that persistent MRD positivity in AML patients in morphological remission prior to allogeneic transplant[49,53,80,85] is associated with increased relapse and decreased survival after transplant, because of the lack of randomized clinical trials it is unproven whether additional intervention can improve clinical outcomes rather than simply increase treatment toxicity. Results from ultradeep NGS on pretransplant blood of AML patients in the BMT CTN phase 3 randomized trial 0901 were recently presented; in patients with an AML-associated variant detected, increased posttransplant relapse risk and inferior overall survival were noted in those randomized to reduced intensity conditioning compared with those randomized to myeloablative conditioning.[86] This study provides the best current evidence that intervention on the MRD state in AML may improve clinical outcomes.

MRD posttransplant

For those patients who develop MRD positivity off treatment, preemptive treatment may be feasible even after allogeneic transplant. In the RELAZA2 phase 2 trial, azacitidine converted 36% of 53 patients back to CR_{MRD-}, with 20% maintaining this status during the 2-year follow-up. Although such nonintensive interventions directed to younger patients progressing to MRD positivity posttransplant are of interest for further evaluation, optimizing delivery will be challenging because of the variables of relapse kinetics and tolerability.

Comment on case presentation

A 62-year-old man with a normal white blood cell count was diagnosed with cytogenetically normal AML, and subsequent molecular studies for ELN 2017 adverse and good risk mutations, including mutated-*NPM1*, were negative. He was fit enough to be treated with 2 courses of induction chemotherapy (standard UK NCRI AML treatment). MRD response was assessed by flow cytometry. He achieved a CR but was MRD positive by flow cytometry (at 0.12%) after the second cycle. Because of deterioration in performance status, his comorbidities, and preference, further treatment options did not

Figure 3 (continued) were set to reach target mean fluorescence intensity (MFI) values by acquisition of rainbow calibration beads without compensation for fluorescent channels FL1 to FL8 on the Canto cytometers; these rainbow bead settings were transposed to Navios cytometers by applying MFI target = Canto target/256. Mirrored (superimposable) target peaks for both cytometers are shown for FL1 and FL8 fluorescent channels with an example of resulting comparable antibody profiles between cytometer platforms. (B) [3] Quality controls for reproducibility of staining profiles from harmonized cytometer settings/sample processing between cytometers/laboratories. Examples shown are for CD117 and CD38 expression intensity on $CD34^+$ gated mononuclear cells of tube 1 from 10 shared BM samples stained and then acquired on Canto or Navios flow cytometers. Intensity profiles are similar between cytometer platforms for each sample. [4] External quality assessment for all harmonized steps from preanalytical to final gating analyses by distribution of a normal BM to 22 participating laboratories (cytometer platforms: 12 Cantos, 10 Navios). Example shows that strong reproducibility can be achieved in the detection of rare events (shown for $CD34^+CD38^-$) among 22 participating laboratories. Figure by Christophe Roumier and Adriana Plesa, ALFA.

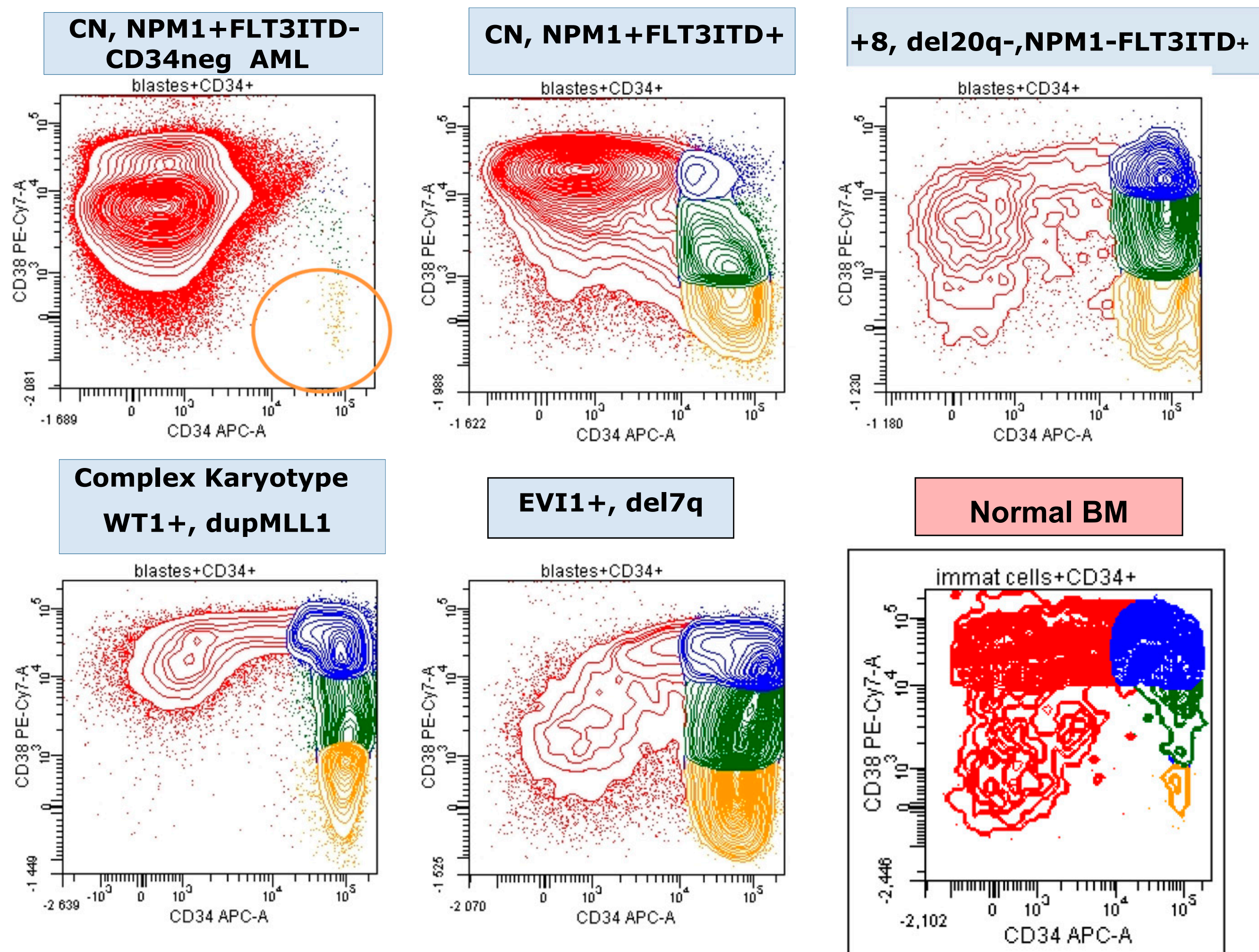


Figure 4. CD34/CD38 expression pattern of blasts from 5 diagnostic AML samples showing the variability in the frequency of the most immature leukemia cells ($CD34^{+}CD38^{-}$; orange) compared with normal BM. CN, normal cytogenetics. Figure by Adriana Plesa and Christophe Roumier, ALFA.

include transplant. He was known to have *IDH1 R132* and *DNMT3A* mutations and was considered for novel regimens or azacitidine. Should he be monitored if he has further treatment and how?

Even if this patient were younger, he has a very high risk for relapse by his postinduction MRD status, even when treated with 3 or 4 courses of standard induction/consolidation chemotherapy (3-year CIR 89%, data from NCRI AML17 trial for younger adults[5]). Because he could not proceed to allogeneic transplant, other treatment options include azacitidine maintenance (although it is uncertain whether MRD-positive patients derive benefit) or, when available, novel regimens incorporating Bcl2[87] or IDH1 inhibition (venetoclax and ivosidenib, respectively).[22] There are no recommendations available for MRD monitoring off-trial in the setting of less intensive treatment. Because he is MRD positive by flow cytometry, conversion to MRD negativity by this assay (observed in 32% of older AML patients attaining CR/CRi with venetoclax and low-dose cytarabine[87] and in 40% treated by hypomethylating agents[88]) may be encouraging for a more durable response, although not as yet a surrogate for survival. Parallel monitoring of *IDH1 R132* mutation[20-22] may be particularly informative for tracking on-target effect of an IDH1 inhibitor. Clearance of *IDH1* mutations by digital PCR (limit of detection 2 to 4×10^{-4}) in this context appears promising as a surrogate for outcome.[22]

Concluding remarks

AML MRD evaluation in clinical practice is happening and will continue to increase. Upfront intensification that includes allogeneic transplantation plus incorporation of available novel agents and maintenance schedules expand current possibilities to improve outcomes but require evidence from randomized clinical trials to establish benefit. MRD-guided therapy added to diagnostic genetic profiling has the potential to target these therapeutic options appropriately and, thus, improve the ratio of benefit to toxicity and costs. Although MRD at single time points has strong prognostic (and, for some treatments, predictive) value at a cohort level, consecutive measurements of currently recommended MRD targets during and after treatment are more likely to provide accurate information for individual patients by tracking any increase in their MRD levels. This also reduces the potential impact from false-negative and false-positive MRD results inherent in any single MRD test.

As in the case of acute lymphoblastic leukemia, chronic myeloid leukemia, chronic lymphocytic leukemia, and myeloma, trial participation is likely to continue as a key factor for enabling familiarity with the clinical use of MRD, as well as advancing methodology. Clinical trial networks provide resources and direction in addition to investigating unresolved questions. This has been observed, for

example, in the United Kingdom; MRD-informed treatment of non–acute promyelocytic leukemia has been integrated into UK NCRI trials since 2012 with availability of coordinated advice for interpretation of results and management decisions. An important task for the next few years will be the evaluation of error-corrected NGS MRD assays, particularly in those AML subgroups without available sensitive PCR assays. Because there is no uniform marker of leukemia or clonality in AML, unlike chronic myeloid leukemia and lymphoid malignancies, the most appropriate prognostic MRD platform and whether orthogonal testing best measures leukemic reservoirs of relapse will depend on AML subtype. In this regard, AMLs with greater intratumoral genetic heterogeneity (and, consequently, also with increased mechanisms for treatment escape) are particularly challenging for tracking relevant MRD by mutations.

Acknowledgments

The authors thank Richard Dillon (UK NCRI Group) and Adriana Plesa and Christophe Roumier (coordinators of Flow AML MRD on behalf of the Acute Leukemia French Association (ALFA) laboratories and lead investigators) for the analyses and preparation of their contributed figures, as well as helpful discussions. They also thank Naeem Khan (University of Birmingham) for assistance with mass cytometry and gratefully acknowledge their ELN colleagues for major contributions to information in this review, including a continuing tribute to David Grimwade. S.D.F. and Richard Dillon thank Nigel Russell, Alan Burnett, Robert Hills, Charlie Craddock, Paresh Vyas, Brian Huntly, and other members of the NCRI AML Working Group for advice and enabling evaluation of MRD in the NCRI AML trials. S.D.F. also gratefully acknowledges present and past laboratory colleagues of the NCRI MRD network, including Kathleen Gallagher, Georgia Andrew, Nick McCarthy, Marlen Meztner, Edward Theakston, Adam Ivey, Jelena Jovanovic, Nicola Potter, and Amanda Gilkes.

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

Correspondence

Sylvie D. Freeman, Institute of Immunology and Immunotherapy, University of Birmingham, Birmingham B15 2TT, United Kingdom; e-mail: s.freeman@bham.ac.uk.

References

1. Hourigan CS, Gale RP, Gormley NJ, Ossenkoppele GJ, Walter RB. Measurable residual disease testing in acute myeloid leukaemia. *Leukemia*. 2017;31(7):1482-1490.
2. Epstein-Peterson ZD, Devlin SM, Stein EM, Estey E, Tallman MS. Widespread use of measurable residual disease in acute myeloid leukemia practice. *Leuk Res*. 2018;67:92-98.
3. Ivey A, Hills RK, Simpson MA, et al; UK National Cancer Research Institute AML Working Group. Assessment of minimal residual disease in standard-risk AML. *N Engl J Med*. 2016;374(5):422-433.
4. Balsat M, Renneville A, Thomas X, et al. Postinduction minimal residual disease predicts outcome and benefit from allogeneic stem cell transplantation in acute myeloid leukemia with NPM1 mutation: a study by the Acute Leukemia French Association Group. *J Clin Oncol*. 2017;35(2):185-193.
5. Freeman SD, Hills RK, Virgo P, et al. Measurable residual disease at induction redefines partial response in acute myeloid leukemia and stratifies outcomes in patients at standard risk without NPM1 mutations. *J Clin Oncol*. 2018;36(15):1486-1497.
6. 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):4123-4131.
7. Schuurhuis GJ, Heuser M, Freeman S, et al. Minimal/measurable residual disease in AML: a consensus document from the European LeukemiaNet MRD Working Party. *Blood*. 2018;131(12):1275-1291.
8. U.S. Department of Health and Human Services. Hematologic malignancies: regulatory considerations for use of minimal residual disease in development of drug and biological products for treatment. https://www.fda.gov/media/117035/download. Accessed 2 May 2019.
9. Platzbecker U, Middeke JM, Sockel K, et al. Measurable residual disease-guided treatment with azacitidine to prevent haematological relapse in patients with myelodysplastic syndrome and acute myeloid leukaemia (RELAZA2): an open-label, multicentre, phase 2 trial. *Lancet Oncol*. 2018;19(12):1668-1679.
10. Selim AG, Moore AS. Molecular minimal residual disease monitoring in acute myeloid leukemia: challenges and future directions. *J Mol Diagn*. 2018;20(4):389-397.
11. Grimwade D, Lo Coco F. Acute promyelocytic leukemia: a model for the role of molecular diagnosis and residual disease monitoring in directing treatment approach in acute myeloid leukemia. *Leukemia*. 2002;16(10): 1959-1973.
12. Hoffmann H, Thiede C, Glauche I, et al. The prognostic potential of monitoring disease dynamics in NPM1-positive acute myeloid leukemia. *Leukemia*. 2019;33(6):1531-1534.
13. Höllein A, Jeromin S, Meggendorfer M, et al. Minimal residual disease (MRD) monitoring and mutational landscape in AML with RUNX1-RUNX1T1: a study on 134 patients. *Leukemia*. 2018;32(10):2270-2274.
14. Martínez-Losada C, Serrano-López J, Serrano-López J, et al. Clonal genetic evolution at relapse of favorable-risk acute myeloid leukemia with *NPM1* mutation is associated with phenotypic changes and worse outcomes. *Haematologica*. 2018;103(9):e400-e403.
15. Höllein A, Meggendorfer M, Dicker F, et al. *NPM1* mutated AML can relapse with wild-type *NPM1*: persistent clonal hematopoiesis can drive relapse. *Blood Adv*. 2018;2(22):3118-3125.
16. Gaidzik VI, Weber D, Paschka P, et al; German-Austrian Acute Myeloid Leukemia Study Group (AMLSG). DNMT3A mutant transcript levels persist in remission and do not predict outcome in patients with acute myeloid leukemia. *Leukemia*. 2018;32(1):30-37.
17. Jongen-Lavrencic M, Grob T, Hanekamp D, et al. Molecular minimal residual disease in acute myeloid leukemia. *N Engl J Med*. 2018;378(13): 1189-1199.
18. Morita K, Kantarjian HM, Wang F, et al. Clearance of somatic mutations at remission and the risk of relapse in acute myeloid leukemia. *J Clin Oncol*. 2018;36(18):1788-1797.
19. Ottone T, Alfonso V, Iaccarino L, et al. Longitudinal detection of DNMT3A^{R882H} transcripts in patients with acute myeloid leukemia. *Am J Hematol*. 2018;93(5):E120-E123.
20. Ferret Y, Boissel N, Helevaut N, et al. Clinical relevance of *IDH1/2* mutant allele burden during follow-up in acute myeloid leukemia. A study by the French ALFA group. *Haematologica*. 2018;103(5):822-829.
21. Ok CY, Loghavi S, Sui D, et al. Persistent *IDH1/2* mutations in remission can predict relapse in patients with acute myeloid leukemia. *Haematologica*. 2019;104(2):305-311.
22. DiNardo CD, Stein EM, de Botton S, et al. Durable remissions with ivosidenib in IDH1-mutated relapsed or refractory AML. *N Engl J Med*. 2018;378(25):2386-2398.
23. Stein EM, DiNardo CD, Fathi AT, et al. Molecular remission and response patterns in patients with mutant-*IDH2* acute myeloid leukemia treated with enasidenib. *Blood*. 2019;133(7):676-687.
24. Desai P, Hassane D, Roboz GJ. Clonal hematopoiesis and risk of acute myeloid leukemia. *Best Pract Res Clin Haematol*. 2019;32(2):177-185.
25. Kim T, Moon JH, Ahn JS, et al. Next-generation sequencing-based posttransplant monitoring of acute myeloid leukemia identifies patients at high risk of relapse. *Blood*. 2018;132(15):1604-1613.
26. Nakamura S, Yokoyama K, Shimizu E, et al. Prognostic impact of circulating tumor DNA status post-allogeneic hematopoietic stem cell transplantation in AML and MDS. *Blood*. 2019;133(25):2682-2695.
27. 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):424-447.

28. Parkin B, Londoño-Joshi A, Kang Q, Tewari M, Rhim AD, Malek SN. Ultrasensitive mutation detection identifies rare residual cells causing acute myelogenous leukemia relapse. *J Clin Invest*. 2017;127(9):3484-3495.
29. van Galen P, Hovestadt V, Wadsworth II MH, et al. Single-cell RNA-seq reveals AML hierarchies relevant to disease progression and immunity. *Cell*. 2019;176(6):1265-1281.e24.
30. Zeijlemaker W, Kelder A, Oussoren-Brockhoff YJ, et al. Peripheral blood minimal residual disease may replace bone marrow minimal residual disease as an immunophenotypic biomarker for impending relapse in acute myeloid leukemia. *Leukemia*. 2016;30(3):708-715.
31. Cilloni D, Renneville A, Hermitte F, et al. Real-time quantitative polymerase chain reaction detection of minimal residual disease by standardized WT1 assay to enhance risk stratification in acute myeloid leukemia: a European LeukemiaNet study. *J Clin Oncol*. 2009;27(31):5195-5201.
32. Maurillo L, Buccisano F, Spagnoli A, et al. Monitoring of minimal residual disease in adult acute myeloid leukemia using peripheral blood as an alternative source to bone marrow. *Haematologica*. 2007;92(5):605-611.
33. Guénot C, Lacombe F, Allou K, et al; Groupe d'Etude Immunologique des Leucémies (GEIL). Peripheral blood minimal/measurable residual disease assessed in flow cytometry in acute myeloblastic leukemia. *Leukemia*. 2019;33(7):1814-1816.
34. Wong HY, Sung AD, Lindblad KE, et al. Molecular measurable residual disease testing of blood during AML cytotoxic therapy for early prediction of clinical response. *Front Oncol*. 2019;8:669.
35. Duncavage EJ, Uy GL, Petti AA, et al. Mutational landscape and response are conserved in peripheral blood of AML and MDS patients during decitabine therapy. *Blood*. 2017;129(10):1397-1401.
36. Buccisano F, Dillon R, Freeman SD, Venditti A. Role of minimal (measurable) residual disease assessment in older patients with acute myeloid leukemia. *Cancers (Basel)*. 2018;10(7):10.
37. Ommen HB. Monitoring minimal residual disease in acute myeloid leukaemia: a review of the current evolving strategies. *Ther Adv Hematol*. 2016;7(1):3-16.
38. Boddu P, Gurguis C, Sanford D, et al. Response kinetics and factors predicting survival in core-binding factor leukemia. *Leukemia*. 2018; 32(12):2698-2701.
39. Bibault JE, Figeac M, Hélevaut N, et al. Next-generation sequencing of FLT3 internal tandem duplications for minimal residual disease monitoring in acute myeloid leukemia. *Oncotarget*. 2015;6(26):22812-22821.
40. McMahon CM, Ferng T, Canaani J, et al. Clonal selection with RAS pathway activation mediates secondary clinical resistance to selective FLT3 inhibition in acute myeloid leukemia [published online ahead of print 14 May 2019]. *Cancer Discov*. doi:10.1158/2159-8290.CD-18-1453.
41. Levis MJ, Perl AE, Altman JK, et al. A next-generation sequencing-based assay for minimal residual disease assessment in AML patients with *FLT3*-ITD mutations. *Blood Adv*. 2018;2(8):825-831.
42. Schnittger S, Schoch C, Dugas M, et al. Analysis of FLT3 length mutations in 1003 patients with acute myeloid leukemia: correlation to cytogenetics, FAB subtype, and prognosis in the AMLCG study and usefulness as a marker for the detection of minimal residual disease. *Blood*. 2002;100(1):59-66.
43. Zuffa E, Franchini E, Papayannidis C, et al. Revealing very small FLT3 ITD mutated clones by ultra-deep sequencing analysis has important clinical implications in AML patients. *Oncotarget*. 2015;6(31): 31284-31294.
44. Lin MT, Tseng LH, Dudley JC, et al. A novel tandem duplication assay to detect minimal residual disease in FLT3/ITD AML. *Mol Diagn Ther*. 2015;19(6):409-417.
45. Blätte TJ, Schmalbrock LK, Skambraks S, et al. getITD for FLT3-ITD-based MRD monitoring in AML [published online ahead of print 14 May 2019]. *Leukemia*. doi:10.1038/s41375-019-0483-z.
46. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic classification and prognosis in acute myeloid leukemia. *N Engl J Med*. 2016;374(23): 2209-2221.
47. Candoni A, De Marchi F, Zanini F, et al. Predictive value of pretransplantation molecular minimal residual disease assessment by WT1 gene expression in FLT3-positive acute myeloid leukemia. *Exp Hematol*. 2017;49:25-33.
48. Mencia-Trinchant N, Hu Y, Alas MA, et al. Minimal residual disease monitoring of acute myeloid leukemia by massively multiplex digital PCR in patients with NPM1 mutations. *J Mol Diagn*. 2017;19(4):537-548.
49. Thol F, Gabdoulline R, Liebich A, et al. Measurable residual disease monitoring by NGS before allogeneic hematopoietic cell transplantation in AML. *Blood*. 2018;132(16):1703-1713.
50. Bhatnagar B, Eisfeld AK, Nicolet D, et al. Persistence of DNMT3A R882 mutations during remission does not adversely affect outcomes of patients with acute myeloid leukaemia. *Br J Haematol*. 2016;175(2):226-236.
51. Debarri H, Lebon D, Roumier C, et al. IDH1/2 but not DNMT3A mutations are suitable targets for minimal residual disease monitoring in acute myeloid leukemia patients: a study by the Acute Leukemia French Association. *Oncotarget*. 2015;6(39):42345-42353.
52. Dillon LW, Hayati S, Roloff GW, et al. Targeted RNA-sequencing for the quantification of measurable residual disease in acute myeloid leukemia. *Haematologica*. 2019;104(2):297-304.
53. 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):329-336.
54. Buccisano F, Maurillo L, Spagnoli A, et al. Cytogenetic and molecular diagnostic characterization combined to postconsolidation minimal residual disease assessment by flow cytometry improves risk stratification in adult acute myeloid leukemia. *Blood*. 2010;116(13):2295-2303.
55. Ravandi F, Jorgensen J, Borthakur G, et al. Persistence of minimal residual disease assessed by multiparameter flow cytometry is highly prognostic in younger patients with acute myeloid leukemia. *Cancer*. 2017;123(3):426-435.
56. Brüggemann M, Schrauder A, Raff T, et al; International Berlin-Frankfurt-Münster Study Group (I-BFM-SG). Standardized MRD quantification in European ALL trials: proceedings of the Second International Symposium on MRD assessment in Kiel, Germany, 18-20 September 2008. *Leukemia*. 2010;24(3):521-535.
57. Keeney M, Wood BL, Hedley BD, et al. A QA program for MRD testing demonstrates that systematic education can reduce discordance among experienced interpreters. *Cytometry B Clin Cytom*. 2018;94(2):239-249.
58. Brooimans RA, van der Velden VHJ, Boeckx N, et al. Immunophenotypic measurable residual disease (MRD) in acute myeloid leukemia: is multicentric MRD assessment feasible? *Leuk Res*. 2019;76:39-47.
59. Jentzsch M, Bill M, Nicolet D, et al. Prognostic impact of the CD34+/ CD38− cell burden in patients with acute myeloid leukemia receiving allogeneic stem cell transplantation. *Am J Hematol*. 2017;92(4):388-396.
60. Khan N, Freeman SD, Virgo P, et al. An immunophenotypic pretreatment predictor for poor response to induction chemotherapy in older acute myeloid leukaemia patients: blood frequency of CD34+ CD38 low blasts. *Br J Haematol*. 2015;170(1):80-84.
61. Terwijn M, Zeijlemaker W, Kelder A, et al. Leukemic stem cell frequency: a strong biomarker for clinical outcome in acute myeloid leukemia. *PLoS One*. 2014;9(9):e107587.
62. Pollyea DA, Jordan CT. Therapeutic targeting of acute myeloid leukemia stem cells. *Blood*. 2017;129(12):1627-1635.
63. Zeijlemaker W, Kelder A, Oussoren-Brockhoff YJ, et al. A simple one-tube assay for immunophenotypical quantification of leukemic stem cells in acute myeloid leukemia. *Leukemia*. 2016;30(2):439-446.
64. Bradbury C, Houlton AE, Akiki S, et al. Prognostic value of monitoring a candidate immunophenotypic leukaemic stem/progenitor cell population in patients allografted for acute myeloid leukaemia. *Leukemia*. 2015; 29(4):988-991.
65. Zeijlemaker W, Grob T, Meijer R, et al. $CD34^{+}CD38^{-}$ leukemic stem cell frequency to predict outcome in acute myeloid leukemia. *Leukemia*. 2019;33(5):1102-1112.
66. Boyd AL, Aslostovar L, Reid J, et al. Identification of chemotherapy-induced leukemic-regenerating cells reveals a transient vulnerability of human AML recurrence. *Cancer Cell*. 2018;34(3):483-498.e5.
67. Ho TC, LaMere M, Stevens BM, et al. Evolution of acute myelogenous leukemia stem cell properties after treatment and progression. *Blood*. 2016;128(13):1671-1678.

68. Coustan-Smith E, Song G, Shurtleff S, et al. Universal monitoring of minimal residual disease in acute myeloid leukemia. *JCI Insight*. 2018;3(9): pii:98561.
69. Kimball AK, Oko LM, Bullock BL, Nemenoff RA, van Dyk LF, Clambey ET. A beginner's guide to analyzing and visualizing mass cytometry data. *J Immunol*. 2018;200(1):3-22.
70. Hu Z, Jujjavarapu C, Hughey JJ, et al. MetaCyto: a tool for automated meta-analysis of mass and flow cytometry data. *Cell Reports*. 2018;24(5): 1377-1388.
71. Lacombe F, Dupont B, Lechevalier N, et al. New concepts of flow cytometry analysis in oncohematology: application to diagnosis and follow up (minimal residual disease) in AML, ALL, and MDS. *Blood*. 2017;130(suppl 1):1421.
72. Marani C, Clavio M, Grasso R, et al. Integrating post induction WT1 quantification and flow-cytometry results improves minimal residual disease stratification in acute myeloid leukemia. *Leuk Res*. 2013;37(12):1606-1611.
73. Getta BM, Devlin SM, Levine RL, et al. Multicolor flow cytometry and multigene next-generation sequencing are complementary and highly predictive for relapse in acute myeloid leukemia after allogeneic transplantation. *Biol Blood Marrow Transplant*. 2017;23(7):1064-1071.
74. Hodes A, Calvo KR, Dulau A, Maric I, Sun J, Braylan R. The challenging task of enumerating blasts in the bone marrow. *Semin Hematol*. 2019; 56(1):58-64.
75. Freeman SD, Hills RK, Russell NH, et al. Induction response criteria in acute myeloid leukaemia: implications of a flow cytometric measurable residual disease negative test in refractory adults. *Br J Haematol*. 2019; 186(1):130-133.
76. Ouyang J, Goswami M, Tang G, et al. The clinical significance of negative flow cytometry immunophenotypic results in a morphologically scored positive bone marrow in patients following treatment for acute myeloid leukemia. *Am J Hematol*. 2015;90(6):504-510.
77. Inaba H, Coustan-Smith E, Cao X, et al. Comparative analysis of different approaches to measure treatment response in acute myeloid leukemia. *J Clin Oncol*. 2012;30(29):3625-3632.
78. Loken MR, Alonzo TA, Pardo L, et al. Residual disease detected by multidimensional flow cytometry signifies high relapse risk in patients with de novo acute myeloid leukemia: a report from Children's Oncology Group. *Blood*. 2012;120(8):1581-1588.
79. Zhou Y, Wood BL, Walter RB, et al. Is there a need for morphologic exam to detect relapse in AML if multi-parameter flow cytometry is employed? *Leukemia*. 2017;31(11):2536-2537.
80. Hourigan CS, Goswami M, Battiwalla M, et al. When the minimal becomes measurable. *J Clin Oncol*. 2016;34(21):2557-2558.
81. Versluis J, Kalin B, Zeijlemaker W, et al. Graft-versus-leukemia effect of allogeneic stem-cell transplantation and minimal residual disease in patients with acute myeloid leukemia in first complete remission. *JCO Precision Oncology*. 2017;1(1):1-13.
82. Lambert J, Lambert J, Thomas X, et al. Early detection of WT1 minimal residual disease predicts outcome in acute myeloid leukemia and identify patients with high risk of relapse independently of allogeneic stem cell transplantation. *Blood*. 2017;130(suppl 1):29.
83. Sierra J, Garrido A, Vives S, et al. Therapy for acute myeloid leukemia (AML) adjusted to genetic data and minimal residual disease: results of the AML12 Trial of the Spanish Cetlam Group in adults up to the age of 70 years. *Blood*. 2017;130(suppl 1):567.
84. Venditti A, Piciocchi A, Candoni A, et al. Risk-adapted, MRD-directed therapy for young adults with newly diagnosed acute myeloid leukemia: results of the AML1310 trial of the GIMEMA group. Paper presented at the 22nd Congress of EHA, 23 June 2017, Madrid, Spain.
85. Buckley SA, Wood BL, Othus M, et al. Minimal residual disease prior to allogeneic hematopoietic cell transplantation in acute myeloid leukemia: a meta-analysis. *Haematologica*. 2017;102(5):865-873.
86. Hourigan CS, Dillon L, Logan B, et al. Impact of conditioning intensity of allogeneic transplantation for acute myeloid leukemia with genomic evidence of residual disease. Paper presented at the 24th Congress of EHA, 14 June 2019, Amsterdam, Netherlands.
87. Wei A, Strickland SA, Hou J-Z, et al. Venetoclax with low-dose cytarabine induces rapid, deep, and durable responses in previously untreated older adults with AML ineligible for intensive chemotherapy. *Blood*. 2018;132:284.
88. Boddu P, Jorgensen J, Kantarjian H, et al. Achievement of a negative minimal residual disease state after hypomethylating agent therapy in older patients with AML reduces the risk of relapse. *Leukemia*. 2018; 32(1):241-244.
89. Dillon R, Hills RK, Freeman SD, et al. Pre-transplant *NPM1* mutant transcript level is highly predictive of outcome after allograft and thresholds are dependent on *FLT3* ITD status. *Blood*. 2018;132:2739.