MRD evaluation of AML in clinical practice

Freeman, Sylvie D.; Hourigan, Christopher S.

DOI:
10.1182/hematology.2019000060

Document Version
Peer reviewed version

Citation for published version (Harvard):
Freeman, SD & Hourigan, CS 2019, 'MRD evaluation of AML in clinical practice: are we there yet?', *Hematology*, vol. 2019, no. 1, pp. 557-569. https://doi.org/10.1182/hematology.2019000060

MRD EVALUATION OF AML IN CLINICAL PRACTICE: ARE WE THERE YET?

Sylvie D Freeman[1] & Christopher S Hourigan[2]

1 Clinical Immunology Service, Institute of Immunology & Immunotherapy, College of Medical & Dental Sciences, University of Birmingham Birmingham, UK.

2 Laboratory of Myeloid Malignancies, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD, USA.

Abstract

MRD technologies increase our ability to measure response in 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 measure its effectiveness. Initiatives such as from the European Leukemia Network (ELN) 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 of relapse. In parallel to this data, advances in genetic profiling have not only extended understanding of the etiology and the complex dynamic clonal nature of AML but also 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. There however remain challenges inherent in the NGS approach of multiplex quantification of mutations to track AML MRD. While 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.

Case Presentation

A 62yr old man with a normal WBC was diagnosed with cytogenetically normal AML and subsequent molecular studies for ELN 2017 good and adverse risk mutations including mutated-NPM1 were negative. He was fit enough to be treated with two 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. Due to deterioration in performance status, his co-morbidities and preference, further treatment options did not include transplant. He was known to have an IDH1 R132 *and DNMT3A mutations and was considered for novel regimens or azacytidine. Should he be monitored if he has further treatment and how?*

Introduction

AML is a disease that consists of multiple genotypes with not only inter- but also intra-leukemic heterogeneity that is influenced by treatment. It is therefore unsurprising that a range of MRD biomarkers and assay platforms have been generated and are relevant to enable personalised AML MRD monitoring.[1] This however contributes to the challenge that MRD testing in 2019 continues to present, i.e. deciphering how, when and even whether it should apply to individual patients. From a survey in 2016-17, 69% of USA leukemia physicians reported routine use of AML MRD, most commonly flow cytometry (MFC) followed by PCR for mutated-*NPM1* but only 21% however 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 post induction in a patient being considered for transplant, responses were divided equally between recommending against a transplant vs additional chemotherapy vs changing conditioning regimen. This variability reflects the paucity of high-quality evidence currently available to inform such AML MRD based decisions. It also highlights 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, post induction MRD in the blood by qPCR predicts outcome independently of other mutations including *FLT3* [3] and benefit from transplant (HR of 0.25 for overall survival).[4] However in wild type *NPM1* intermediate risk AML (as for patient in 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, namely qPCR and flow cytometry can provide sophisticated information for MRD levels in most patients, identifying across all AML risk groups those most likely to relapse with current treatment schedules. 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, together with amalgamated experience in routine clinical practice step-wise 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 standardisation continuously progresses MRD detection. This review focuses on the more recent information from MRD testing, including assay limitations and prospects, that with consensus recommendations can guide current implementation.

Recent Guidance from ELN and FDA

Since 2018, guidance for the expanding application of MRD in AML has been proposed by two organisations. The ELN through international collaboration published a consensus document from expert AML MRD laboratories with recommendations for both 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 harmonise and progress assays for which efforts by participating laboratories are ongoing. The FDA after a series of workshops over several years, have invited comments on their draft guidance for the use of MRD as a biomarker in regulatory submissions for haematological cancers.[8] In this the FDA list criteria for MRD assays, for example that reporting MRD negative results requires information on detection limits and, perhaps more debatable for certain assays, sensitivity should be at least a log below the cut-off 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 comments that MRD could be used to stratify or enrich trial populations, a strategy applied for the RELAZA2 trial [9] but when MRD is a trial endpoint any patient with missed MRD samples should be categorised as unresponsive in the analysis.

MRD Assay Considerations

As highlighted by the ELN and FDA papers, 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 comprise 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, stage taken and extent of assay standardisation. Another 'S', price ($), is a practical inclusion to 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 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 artefactual assay background (such as non-specific antibody binding and autofluorescence in flow cytometry or non-specific primer binding in PCR). For example, MRD assays by flow cytometry and *WT1* RNA overexpression by qPCR (both applicable in the majority of all AML patients) have intermediate specificity due mainly to 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 as only treatment resistant APL cells including those that may in time be relapse initiating with have this MRD marker. Similarly, other AML subtypes including 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 co-existing 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 report) with associated co-existing and persisting clonal hematopoietic mutations including *DNMT3A*. [15]

It is now apparent that CHIP associated mutations including the *DNMT3A* point mutation in the above clinical case, although frequent and stable in AML, can persist post treatment at high levels despite longer-term disease free survival.[16-19] These are therefore insufficiently specific MRD markers for AML relapse, substantiating the FDA statement that marker selected to assess MRD should "not reflect underlying clonal hematopoiesis". This does not preclude however 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 bio-markers for pre-leukemic activity of novel agents. Of particular interest in this regard are CHIP mutations in DNA damage response genes such as *TP53* and *PPM1D*

that are enriched after cytotoxic therapy and associated with an increased risk of developing leukemia.[24] Allelic burdens of *DNMT3A*, *TET2* and *ASXL1* mutations, including as circulating tumour DNA, may also have potential utility post allogenic stem cell transplantation to track ablation of patient clonal hematopoesis.[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 NGS mutation profiling) to $10^{-5\ to\ -6}$. 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 (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/CRi after their first course of induction in the NCRI AML17 trial had a 5 year survival of 52%; if also categorised as CR_{MRD-} by flow cytometry (sensitivity of 10^{-4}) survival increased to 63% overall and 70% when excluding poor risk patients.[5] Observed clinical false negatives (by relapse frequency) from single MRD assessment time points are 20-30% 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 Figure 1B/C 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 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 bone marrow 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 input into limiting the sensitivity / lower level of quantification for any bone marrow sample, independently of the theoretical assay sensitivity. Flow cytometric assays can assess and should incorporate information on all 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% 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 pre-sorted 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 low cell numbers.

Is Blood informative for AML MRD?

Despite blood (PB) providing lower sensitivity than BM (1 log less for mutated-*NPM1* [3]), measuring mutated-*NPM1* transcripts in PB post induction is highly prognostic for the mutated-*NPM1* subgroup.[3,4] Interestingly

the 'false negative' relapse risk is not increased for MRD negativity in blood versus BM. [4] This implies that in the ~25% of mutated-*NPM1* patients with only BM positivity post induction, either the MRD is at a level concomitant with clearance by consolidation or is from non-leukemic initiating, more mature BM cells. Bone marrow is recommended however for maximal sensitivity during later sequential monitoring since bone marrow positivity by qPCR targets usually precedes that of blood providing an increased time window for any interventions. [7]

Reduced sensitivity may be compensated by a differential increase in specificity when testing blood for flow cytometric MRD [30] (due to fewer normal progenitors/precursors and therefore less 'noise'), as observed when measuring *WT1* levels.[31] Relapse free [32] and overall survival [30,33] appear significantly better for patients with MRD negative blood samples post induction and consolidation 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 measurements of VAF frequencies (at >5%) for mutation profiles in PBs with <60% lymphocytes correlated well with paired BM results.[35]

How we would monitor: Core-binding-factor 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 (InDel) 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 at least 2 years at 3 monthly intervals when prior result was negative or low copy numbers. Average kinetics from molecular relapse to clinical relapse range from more than 3-4 months for *CBFB-MYH11* to 2- 3 months for *RUNX1-RUNX1T1* or mutated-*NPM1* when *FLT3 ITD* positive.[37] There is no new data as yet to support a survival benefit from pre-emptive intensification in CBF AMLs despite the clear association between MRD status (post induction/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 as yet 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* rearrangements. Although these can track response (Figure 1), there is insufficient

data to provide guidance for their use hence it is recommended post induction to also assess response by flow cytometric MRD.[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-ITD). *FLT3*-ITD mutations represent late events in leukemic development and hence are not always detectable at relapse, particularly after targeted therapy, often being "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 as 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 one universal approach to low-level quantitative assessment by conventional PCR and bioinformatic mapping of data from NGS challenging. This technical constraint has however been mitigated both by novel PCR methods[44] and sequencing approaches.[39,41,43,45] In approximately 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]; it is therefore recommended to also track such markers[7] particularly for patients receiving *FLT3*-ITD directed therapy. For those without a second mutation in addition to *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 not only for new drug development but also potentially for leukemic disease burden tracking. While such approaches are highly publishable several limitations prevent direct translation to the clinic at present. Digital PCR (dPCR), developed in the 1990s, allows absolute quantification of abnormal DNA sequence by, unlike conventional quantitative PCR (qPCR, Figure 2A), partitioning each template molecule for the PCR reaction into an individual compartment (Figure 2B). Advantages of this technology include high sensitivity due to low background error rate, highly accurate quantification due to elimination of template competition, limited bioinformatic requirements and a rapid "sample-to-result" turn-around 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, "hot-spot" 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 increasing become part of the standard initial diagnostic workup given the importance of such mutations in both risk stratification and therapy selection[27] (Figure 2C). Given detected mutations are thought to origin from the leukemic clone and may persist in remission at those at increased risk of 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 due to limited sample input, insufficient read-depth and lack of appropriate error correction (see Table). NGS "AML panels" at diagnosis are used to determine likely somatic variants typically with a VAF (variant allele frequency, also known as MAF for mutant allele frequency) of equal to or greater than 5% (ie: 5 variant reads per every 100 at that position). Accordingly, those tests use genomic DNA inputs as low as 10ng which represents less than 8000 total cells (typically range of such tests: 10ng - 100ng). Given losses during NGS library preparation, a requirement to detect 3-5 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 ten-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 200-500ng range (ie: at least 1ml of marrow aspirate or blood). Similarly, most AML MRD approaches using NGS, in recognition of massive number of false positive results particularly for single nucleotide variants when considering potential variants less than 2% VAF, have used some form of error correction in the form of laboratory techniques such as 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 both PCR-based library preparation and sequencing itself [17,49] (Figure 2D). The sequencing depth requirement is often substantially higher for AML MRD than that routinely done for diagnostic testing, as in addition to the increased sample input, typically each unique DNA molecule will be sequenced 5 or 10 times in UMI-based error corrected sequencing. Finally, the prognostic significance of AML "MRD" as detected by NGS compared with that detected by qPCR or flow cytometry may not be equivalent, 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 a RNA-seq assay capable of simultaneously detecting all of these molecular targets in a single standardized NGS workflow while maintaining a limit of detection comparable with qPCR.[52] Future NGS AML MRD assays may combine RNA and DNA based approaches.

Update for immunophenotypic MRD

The ELN paper provides a framework for the harmonisation / standardisation 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 colour panels (acceptable in clinical laboratories to increase single-cell information), the importance of harmonising cytometer settings for comparable antibody stained cell profiles and some technical recommendations for sample processing. Knowledge of the diagnostic LAIP prevents reporting MRD negativity in the 5-10% of marker negative patients (no identifiable LAIP at diagnosis as well as at follow-up) while 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 in order to minimise false negatives. Defining MRD positivity by ≥0.1% of leukocytes (1 log greater than limit of detection) is proposed as this threshold is relevant in most studies. The proviso for this however is that some patients with quantifiable MRD at <0.1% may have residual AML that is prognostic, for example due to MRD under-representation 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 Harmonisation / Standardisation

Experience from acute lymphoblastic leukemia (ALL) flow cytometric MRD work has shown that muticenter standardisation[56] can be achieved with interpretive discordance reduced amongst 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 etc. Figure 3 shows an example of an implemented harmonisation strategy (Personal Communication, Plesa A. and Roumier C., on behalf of Acute Leukaemia French Association Intergroup). The ELN are testing a consensus standardised tube as a template to simplify inter-laboratory 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 HSC activity in normal bone marrow. 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 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 standardised 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- 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, mainly due to increased specificity for very poor outcomes in the ~10% of patients with positivity for both LSC and standard MRD. [65] Although increasing assay sensitivity up to 1 in a million by testing more cells could reduce the false negative relapse frequency observed in this study (3yr 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 rationalises the repertoire of informative markers that can inform future routine AML MRD antibody panels [68] as well as for immunotherapy targets. For example IL1RAP has been identified by both proteomic [66] and single cell RNA sequencing [29] strategies as a discriminatory marker particularly in *FLT3-ITD* AML.

Hi-dimensional immunophenotyping using Cytometry by Time of flight (CyTOF) when available allows deeper and/or broader population coverage due to increased parameters (>40) than 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. While CyTOF has proven to be a powerful discovery tool, its slow acquisition rates (500 events/sec) preclude applying CyTOF technology to clinical laboratory assays that require high cell numbers, MRD assays included. However, as 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 personalised 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 versus Molecular MRD detection: *Friend or Foe*

Initial evidence points to molecular and flow cytometric MRD providing complimentary independent information at the same time-points during the active treatment, particularly when the molecular assays have lower sensitivity and specificity.[17,18,21,72,73] About 30% of patients in two studies[17,18] had discordant results after induction between flow cytometric and standard NGS MRD (excluding the main CHIP mutations) and these had an intermediate risk of relapse (~50% at 4yr) in the larger cohort.[17] In order to understand how best to improve and deploy these assays it will be helpful in the research setting to define if their respective false positive / negative results cluster in certain AML subtypes and 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 versus Morphology to assess remission

Flow cytometric assays continue to provide the fastest turn-around time for evaluating blast percentages and MRD so that routine results may be available concurrently if not before morphology reporting. Inter-observer variability for morphological BM blast counts has been recently reported as greatest for blast percentages from 2-10% (and only moderate agreement for <2% and >10%) while 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 as well as sensitivity for discriminating leukemic 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 cohorts75), equivalent to those in CRMRD-.[75] Also of note is that in a series of 87 patients morphologically categorised as relapsed, none were MRD negative by flow.[79] Together these results support refining the present criteria of refractory and relapsed AML to incorporate flow cytometric MRD negative results when a validated assay is available.

Since MRD positivity in CR and morphologic refractory disease appear equivalent for outcomes pre allogeneic 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 IWG criteria) or MRD positive CR after a first course of induction had an equivalent intermediate prognosis (~40% 5yr survivals) with the exception of MRD positive patients with incomplete count recovery who had a much poorer prognosis (19% 5yr survival). Patients with resistant disease had a similar outcome to the latter.[5] Thus when flow cytometric MRD is incorporated to response assessment post first induction, the prognostic effect from blast percentages of ≥5% by morphology appears 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 in published studies support directing allogeneic transplant in first remission only to those patients testing positive for MRD post induction chemotherapy (by qPCR for mutated-*NPM1* or flow cytometry for wild type *NPM1*). However when ELN2017 adverse risk patients are included, the benefit from allogeneic transplant as consolidation appears 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 post-induction time-point will be further informed by forthcoming data in trials implementing this strategy in younger adults (such as NCRI AML17 /19 and HOVON 132 for ELN intermediate risk) and 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 post-consolidation MRD levels in addition to genetic risk. Initial reports from these[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 post-consolidation MRD reinforces the difficulty of testing non-biased 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 monthly qPCR MRD monitoring is recommended to allow a greater time-window for treatment decisions as 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 as well as MRD status after induction and consolidation, pre- and potentially post-transplant MRD testing can inform risk of poorer outcomes.

MRD Status pre Transplant

While 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, due to the lack of randomized clinical trials it is unproven if additional intervention can improve clinical outcomes rather than simply increase treatment toxicity. Results from ultra-deep NGS on pre-transplant blood of AML patients on the BMT CTN phase III randomized trial 0901 were recently presented demonstrating, in those with an AML

associated variant detected, increased post-transplant relapse risk and inferior overall survival in those randomized to reduced intensity conditioning compared with those randomized to myeloablative conditioning.[86] This study provides the best current evidence so far that intervention on the MRD state in AML may improve clinical outcomes.

MRD post Transplant

For those patients who develop MRD positivity off treatment, pre-emptive treatment may be feasible even after allogeneic transplant. In the RELAZA2 phase 2 trial, azacytidine converted 36% of 53 patients back to CR_{MRD-} with 20% maintaining this during the 2yr follow-up. While such non-intensive interventions directed to younger patients progressing to MRD positivity post-transplant are of interest for further evaluation, optimising delivery will be challenging due to the variables of relapse kinetics and tolerability.

Comment on Case Presentation

A 62yr old man with a normal WBC was diagnosed with cytogenetically normal AML and subsequent molecular studies for ELN 2017 good and adverse risk mutations including mutated-NPM1 were negative. He was fit enough to be treated with two 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. Due to deterioration in performance status, his co-morbidities and preference, further treatment options did not include transplant. He was known to have an IDH1 R132 *and DNMT3A mutations and was considered for novel regimens or azacytidine. Should he be monitored if he has further treatment and how?*

This patient even if he were younger has a very high risk of relapse by his post induction MRD status even when treated with 3-4 courses of standard induction / consolidation chemotherapy (3yr CIR 89%, data from NCRI AML17 trial for younger adults[5]). Since he could not proceed to allogeneic transplant other treatment options include azacytidine maintenance (although it is uncertain if 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. As 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 $2–4\times10^{-4}$) 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. While 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 ALL, CML, CLL 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 UK, where MRD-informed treatment for non-APMLs 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. As there is no uniform marker of either leukemia or clonality in AML unlike CML 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 intra-tumoral genetic heterogeneity (and consequently also increased mechanisms for treatment escape) are particularly challenging for tracking relevant MRD by mutations.

Acknowledgements

We thank Richard Dillon (UK NCRI Group), Adriana Plesa, & 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. Also Naeem Khan (University of Birmingham, UK) for assistance regarding mass cytometry. We gratefully acknowledge our EuropeanLeukemiaNet colleagues for their major contributions to information in this review, including a continuing tribute to David Grimwade.

SDF with RD would like to thank Nigel Russell, Alan Burnett, Robert Hills, Charlie Craddock, Paresh Vyas, Brian Huntly and other members of the National Cancer Research Institute (NCRI) AML Working Group for advice and enabling evaluation of MRD in the NCRI AML trials. SDF 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 is supported in part by the Intramural Research Program of the National Heart, Lung, and Blood Institute of the National Institutes of Health.

Conflict of interest statement

CSH has received laboratory research funding from Merck and Sellas.

References

1. Hourigan CS, Gale RP, Gormley NJ, et al: Measurable residual disease testing in acute myeloid leukaemia. Leukemia 31:1482-1490, 2017
2. Epstein-Peterson ZD, Devlin SM, Stein EM, et al: Widespread use of measurable residual disease in acute myeloid leukemia practice. Leuk Res 67:92-98, 2018
3. Ivey A, Hills RK, Simpson MA, et al: Assessment of Minimal Residual Disease in Standard-Risk AML. N Engl J Med 374:422-33, 2016
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 35:185-193, 2017
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 36:1486-1497, 2018
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 31:4123-31, 2013
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 131:1275-1291, 2018
8. Food and Drug Administration H: Hematologic Malignancies: Regulatory Considerations for Use of Minimal Residual Disease in Development of Drug and Biological Products for Treatment; Draft Guidance for Industry; Availability in Services USDoHaH, Administration FaD, Research CfDEaRCfBEa (eds). Federal Register, 2018, pp 52225
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 19:1668-1679, 2018
10. Selim AG, Moore AS: Molecular Minimal Residual Disease Monitoring in Acute Myeloid Leukemia: Challenges and Future Directions. J Mol Diagn 20:389-397, 2018
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 16:1959-73, 2002
12. Hoffmann H, Thiede C, Glauche I, et al: The prognostic potential of monitoring disease dynamics in NPM1-positive acute myeloid leukemia. Leukemia, 2019
13. Hollein 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 32:2270-2274, 2018
14. Martinez-Losada C, Serrano-Lopez J, Serrano-Lopez 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 103:e400-e403, 2018
15. Hollein A, Meggendorfer M, Dicker F, et al: NPM1 mutated AML can relapse with wild-type NPM1: persistent clonal hematopoiesis can drive relapse. Blood Adv 2:3118-3125, 2018
16. Gaidzik VI, Weber D, Paschka P, et al: DNMT3A mutant transcript levels persist in remission and do not predict outcome in patients with acute myeloid leukemia. Leukemia 32:30-37, 2018
17. Jongen-Lavrencic M, Grob T, Hanekamp D, et al: Molecular Minimal Residual Disease in Acute Myeloid Leukemia. N Engl J Med 378:1189-1199, 2018
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 36:1788-1797, 2018
19. Ottone T, Alfonso V, Iaccarino L, et al: Longitudinal detection of DNMT3A(R882H) transcripts in patients with acute myeloid leukemia. Am J Hematol 93:E120-E123, 2018
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 103:822-829, 2018

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 104:305-311, 2019
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 378:2386-2398, 2018
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 133:676-687, 2019
24. Desai P, Hassane D, Roboz GJ: Clonal Hematopoiesis and risk of Acute Myeloid Leukemia. Best Pract Res Clin Haematol 32:177-185, 2019
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 132:1604-1613, 2018
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
27. Dohner H, Estey E, Grimwade D, et al: Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood 129:424-447, 2017
28. Parkin B, Londono-Joshi A, Kang Q, et al: Ultrasensitive mutation detection identifies rare residual cells causing acute myelogenous leukemia relapse. J Clin Invest 127:3484-3495, 2017
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 176:1265-1281 e24, 2019
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 30:708-15, 2016
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 27:5195-201, 2009
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 92:605-11, 2007
33. Guenot C, Lacombe F, Allou K, et al: Peripheral blood minimal/measurable residual disease assessed in flow cytometry in acute myeloblastic leukemia. Leukemia, 2019
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 8:669, 2018
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 129:1397-1401, 2017
36. Buccisano F, Dillon R, Freeman SD, et al: Role of Minimal (Measurable) Residual Disease Assessment in Older Patients with Acute Myeloid Leukemia. Cancers (Basel) 10, 2018
37. Ommen HB: Monitoring minimal residual disease in acute myeloid leukaemia: a review of the current evolving strategies. Ther Adv Hematol 7:3-16, 2016
38. Boddu P, Gurguis C, Sanford D, et al: Response kinetics and factors predicting survival in core-binding factor leukemia. Leukemia 32:2698-2701, 2018
39. Bibault JE, Figeac M, Helevaut N, et al: Next-generation sequencing of FLT3 internal tandem duplications for minimal residual disease monitoring in acute myeloid leukemia. Oncotarget 6:22812-21, 2015
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. Cancer Discov, 2019
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 2:825-831, 2018
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 100:59-66, 2002

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 6:31284-94, 2015
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 19:409-17, 2015
45. Blatte TJ, Schmalbrock LK, Skambraks S, et al: getITD for FLT3-ITD-based MRD monitoring in AML. Leukemia, 2019
46. Papaemmanuil E, Gerstung M, Bullinger L, et al: Genomic Classification and Prognosis in Acute Myeloid Leukemia. N Engl J Med 374:2209-2221, 2016
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 49:25-33, 2017
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 19:537-548, 2017
49. Thol F, Gabdoulline R, Liebich A, et al: Measurable residual disease monitoring by NGS before allogeneic hematopoietic cell transplantation in AML. Blood 132:1703-1713, 2018
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 175:226-236, 2016
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 6:42345-53, 2015
52. Dillon LW, Hayati S, Roloff GW, et al: Targeted RNA-sequencing for the quantification of measurable residual disease in acute myeloid leukemia. Haematologica 104:297-304, 2019
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 34:329-36, 2016
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 116:2295-303, 2010
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 123:426-435, 2017
56. Bruggemann M, Schrauder A, Raff T, et al: Standardized MRD quantification in European ALL trials: proceedings of the Second International Symposium on MRD assessment in Kiel, Germany, 18-20 September 2008. Leukemia 24:521-35, 2010
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 94:239-249, 2018
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 76:39-47, 2019
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 92:388-396, 2017
60. Khan N, Freeman SD, Virgo P, et al: An immunophenotypic pre-treatment predictor for poor response to induction chemotherapy in older acute myeloid leukaemia patients: blood frequency of CD34+ CD38 low blasts. Br J Haematol 170:80-4, 2015
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 9:e107587, 2014

62. Pollyea DA, Jordan CT: Therapeutic targeting of acute myeloid leukemia stem cells. Blood 129:1627-1635, 2017
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 30:439-46, 2016
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 29:988-91, 2015
65. Zeijlemaker W, Grob T, Meijer R, et al: CD34(+)CD38(-) leukemic stem cell frequency to predict outcome in acute myeloid leukemia. Leukemia, 2018
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 34:483-498 e5, 2018
67. Ho TC, LaMere M, Stevens BM, et al: Evolution of acute myelogenous leukemia stem cell properties after treatment and progression. Blood 128:1671-8, 2016
68. Coustan-Smith E, Song G, Shurtleff S, et al: Universal monitoring of minimal residual disease in acute myeloid leukemia. JCI Insight 3, 2018
69. Kimball AK, Oko LM, Bullock BL, et al: A Beginner's Guide to Analyzing and Visualizing Mass Cytometry Data. J Immunol 200:3-22, 2018
70. Hu Z, Jujjavarapu C, Hughey JJ, et al: MetaCyto: A Tool for Automated Meta-analysis of Mass and Flow Cytometry Data. Cell Rep 24:1377-1388, 2018
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, Am Soc Hematology, 2017
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 37:1606-11, 2013
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 23:1064-1071, 2017
74. Hodes A, Calvo KR, Dulau A, et al: The challenging task of enumerating blasts in the bone marrow. Semin Hematol 56:58-64, 2019
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, 2018
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 90:504-10, 2015
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 30:3625-32, 2012
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 120:1581-8, 2012
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 31:2536-2537, 2017
80. Hourigan CS, Goswami M, Battiwalla M, et al: When the Minimal Becomes Measurable. J Clin Oncol 34:2557-8, 2016
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 1:1-13, 2017

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, Am Soc Hematology, 2017
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 130:567-567, 2017
84. Venditti A, Piciocchi, A., Candoni, A., Melillo, L., Calafiore, V., Cairoli, R., De Fabritiis, P., Storti, G., Salutari, P., Lanza, F, Martinelli, G.: Risk-adapted, MRD-directed therapy for young adults with newly diagnosed acute myeloid leukemia: results of the AML1310 trial of the GIMEMA group. . Presented at the EHA, 2017
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 102:865-873, 2017
86. Hourigan CS DL, Logan B, Scott B, Ghannam J, Gui G, Li Y, Licon A, Fei M, Alyea E, Bashey, AH. Deeg J, Devine S, Fernandez H, Giralt S, Hamadani M, Horowitz M, Howard A, Maziarz R, Porter D, Warlick E, Pasquini M, Horwitz M: Impact of Conditioning Intensity of Allogeneic Transplantation for Acute Myeloid Leukemia with Genomic Evidence of Residual Disease Presented at the EHA, 2019
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 132:284-284, 2018
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 32:241-244, 2018
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 132:2739-2739, 2018

Figure Legends

Figure 1 legend

A. Suggested management algorithm for patients with AML with a molecular MRD target.

[1] ELN favourable risk patients with <4 log reduction in *NPM1* mutant transcripts after first induction shown to benefit from a CR1 allograft [4] and any positivity in the PB after second induction is associated with a very high risk of relapse.[3] **[2]** MRD positivity >200 copies / 10^5 *ABL* (i.e. molecular persistence) and serially rising transcript levels after treatment (i.e molecular progression) reliably predicts relapse[7].

[3] At the end of treatment, patients with core binding factor AML with high or serially rising 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 BM or >10 copies in the PB)[7]. Salvage according to panel C should be considered for these patients although there is currently no evidence that this improves outcome. Conversely patients with low copy numbers below these thresholds can be safely monitored according to panel B. **[4]** Although CBF patients with an early unfavourable MRD response have a higher risk of 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 rise while on treatment (i.e. molecular progression).

[5] Although there is currently no evidence that standard risk patients who remain MRD positive benefit from transplant, this is a reasonable approach and adopted in the current NCRI AML19 and MyeChild01 protocols.

B. Suggested algorithm for sequential monitoring after treatment. Patients conversion to MRD positivity, confirmed on a second sample with >1 log rise should be diagnosed with molecular relapse and treated according to panel B [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 DLI and those to whom these standard therapies cannot be given should be considered for investigational approaches.

Figure by Richard Dillon, NCRI Group.

Figure 2 Legend

A) Quantitative polymerase chain reaction (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 cross).

Advantages include ubiquitous in most clinical laboratories, fast turn-around 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 due to sample impurity, 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 template of interest into individual compartments (top) improving performance compared with qPCR due to 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) Next generation sequencing (NGS) has revolutionized 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 (Variant Allele Frequency, VAF). Not all variants detected will be pathogenic somatic mutations however, 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 due to lack of sensitivity and high false positive rates.

D) NGS for AML MRD performed in recent high-quality research studies has typically included error correction, by (upper) incorporation of unique molecular identifiers followed by consensus determination of true (red) variants versus false positives introduced by the technique (green) and/or (lower) bioinformatic approaches to model background error rates at each nucleotide position in those not having a variant and determine probability that observed variant is a true positive (red).

Figure by Erina He, NIH Medical Arts.

Figure 3 legend

A. Strategy applied for Flow Cytometric AML MRD multi-center harmonisation by Acute Leukaemia French Association (ALFA) Intergroup.

[1] Rationale of AML MRD Flow panel design was based on simplicity, reproducibility and cost. Tube 1 was a core combination for Leukemic Associated ImmunoPhenotypes (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 harmonised ('mirrored') between Canto vs Navios cytometer platforms. Voltages were set to reach target 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 shown for FL1 and FL8 fluorescent channels with an example of resulting comparable antibody profiles between cytometer platforms.

[3] Quality controls for reproducibility of staining profiles from harmonised 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 vs Navios . Intensity profiles are similar between cytometer platforms for each sample.

[4] External quality assessment for all harmonised steps from pre-analytical to final gating analyses by distribution of a normal bone marrow 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.

Figure 4 legend

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-, in orange) compared to normal BM. (CN, normal cytogenetics).

Figure by Adriana Plesa and Christophe Roumier.

Table 1: Limitations of use of next generation sequencing to detect residual disease in AML

Problem	Significance	Solution
NGS AML panels used at diagnosis are unfit for purpose of AML MRD	Insufficient sample input/sequencing depth limits assay sensitivity. Insufficient error correction increases false positive rate at low VAF.	NGS for research not clinical use in AML MRD currently. 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 (less than 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 D) helpful. Low VAF variants seen in diagnostic sample or multiple surveillance samples more likely true variants (but see above)
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 "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 suited for research not clinical use in AML MRD currently. FDA guidance for MRD in Hematological Malignancies published in draft form. ELN guidelines on use of NGS in AML MRD coming.

Box: Eight “S’s of AML MRD

Specificity of test

Sensitivity of test

Stability of MRD marker

Subtype of AML

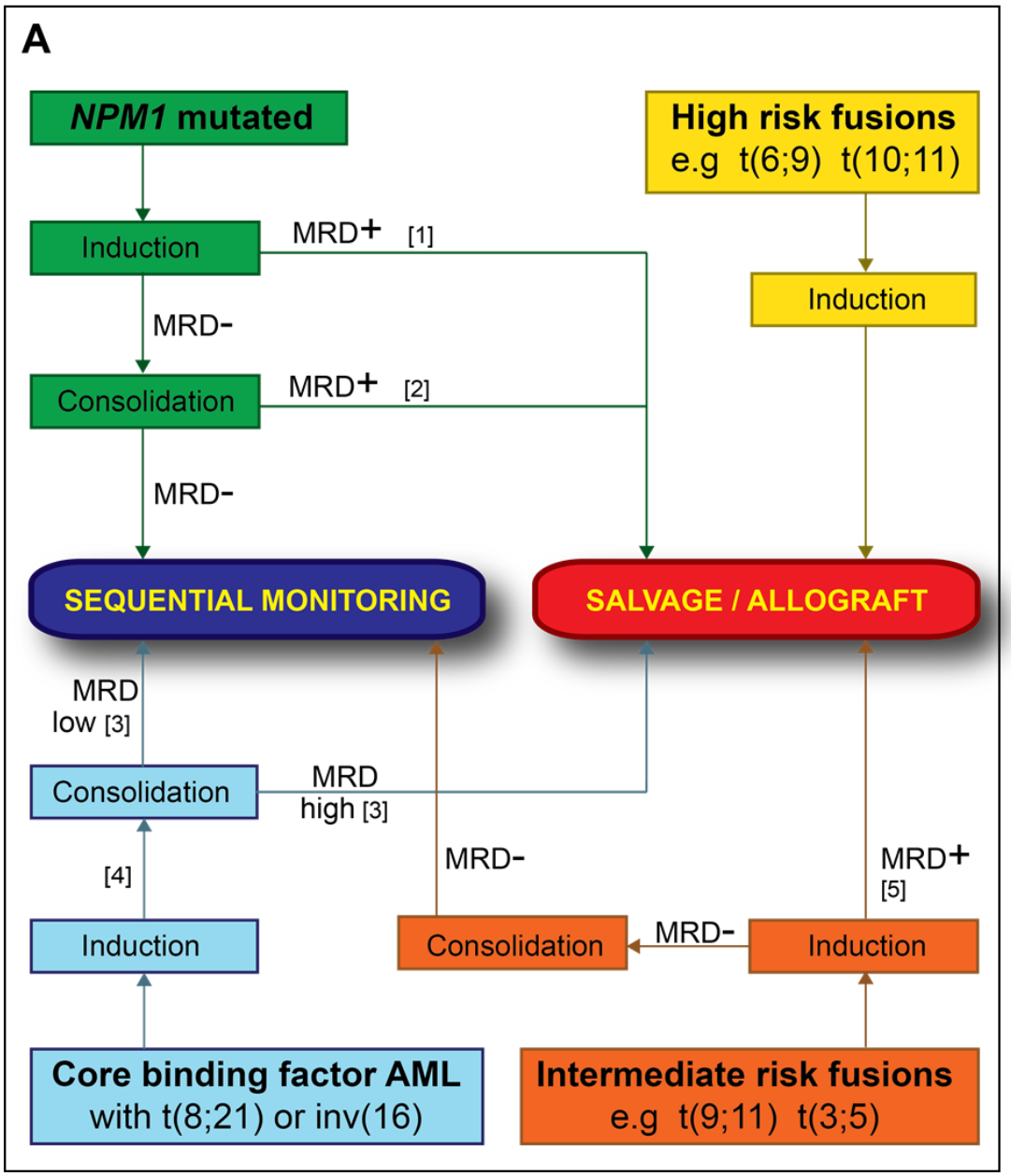
Sample type

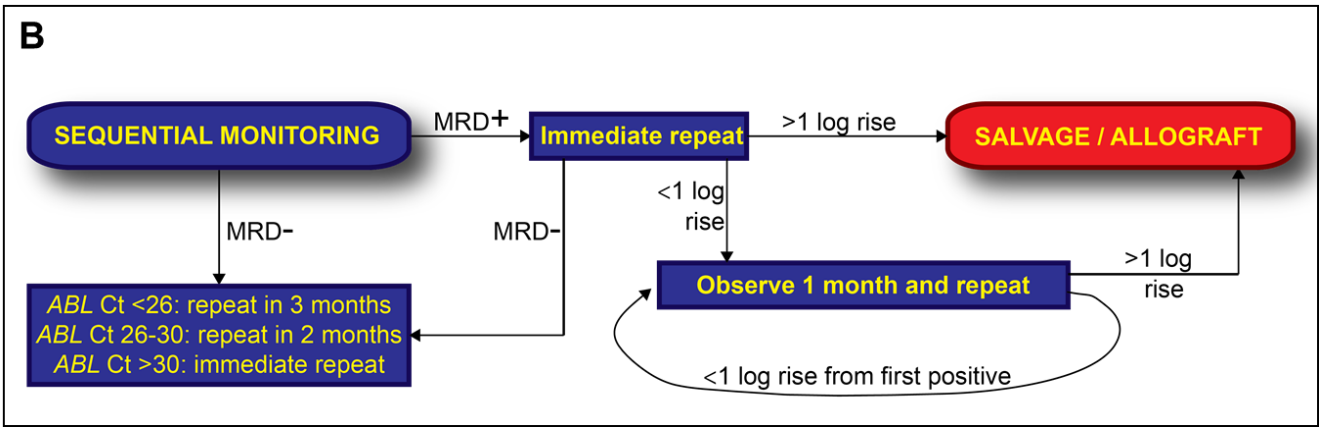
Stage during treatment sample taken

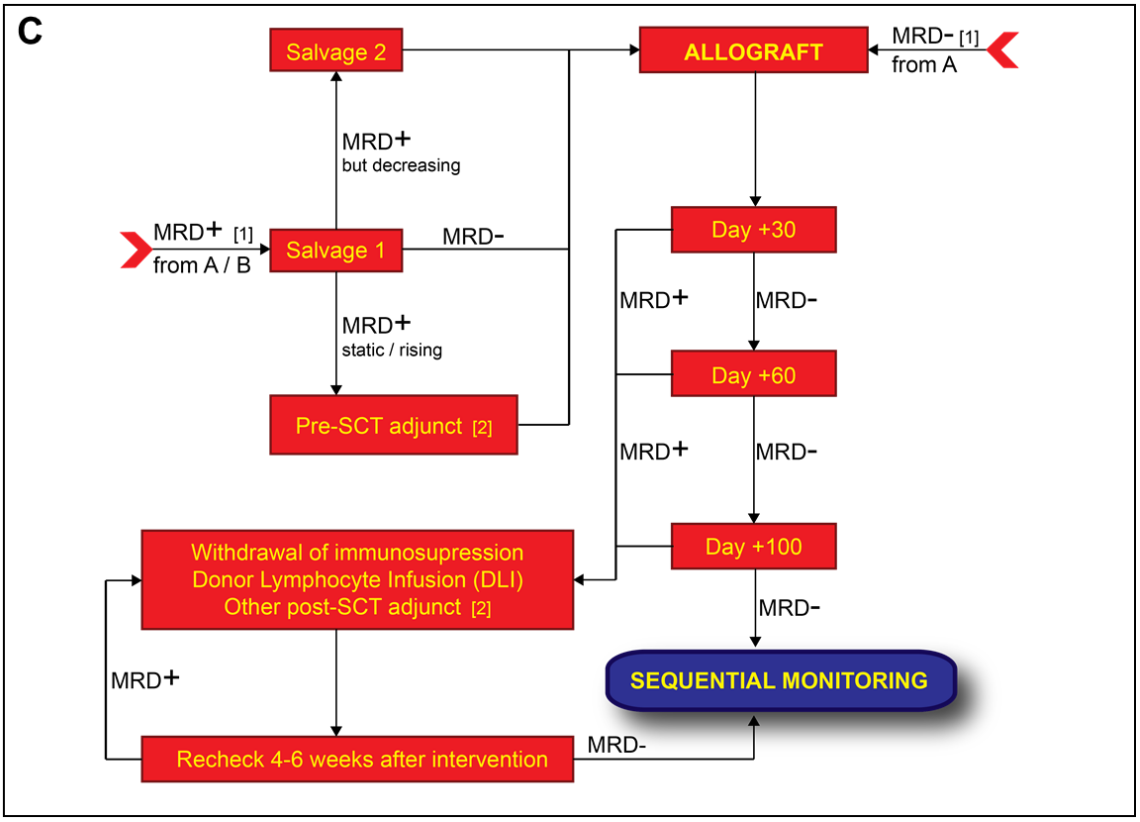
Standardisation of Assay

Price (**$**)

Figure 1


A
NPM1 mutated
High risk fusions
e.g t(6;9) t(10;11)
Induction
MRD+ [1]
MRD-
Induction
Consolidation
MRD+ [2]
MRD-
SEQUENTIAL MONITORING
SALVAGE / ALLOGRAFT
MRD low [3]
Consolidation
MRD high [3]
[4]
MRD-
MRD+ [5]
Induction
Consolidation
MRD-
Induction
Core binding factor AML
with t(8;21) or inv(16)
Intermediate risk fusions
e.g t(9;11) t(3;5)


B
SEQUENTIAL MONITORING
MRD+
Immediate repeat
>1 log rise
SALVAGE / ALLOGRAFT
MRD-
MRD-
<1 log rise
Observe 1 month and repeat
>1 log rise
ABL Ct <26: repeat in 3 months
ABL Ct 26-30: repeat in 2 months
ABL Ct >30: immediate repeat
<1 log rise from first positive


C
Salvage 2
ALLOGRAFT
MRD- [1]
from A
MRD+
but decreasing
MRD+ [1]
from A / B
Salvage 1
MRD-
Day +30
MRD+
MRD-
MRD+
static / rising
Day +60
Pre-SCT adjunct [2]
MRD+
MRD-
Withdrawal of immunosupression
Donor Lymphocyte Infusion (DLI)
Other post-SCT adjunct [2]
Day +100
MRD-
MRD+
SEQUENTIAL MONITORING
Recheck 4-6 weeks after intervention
MRD-

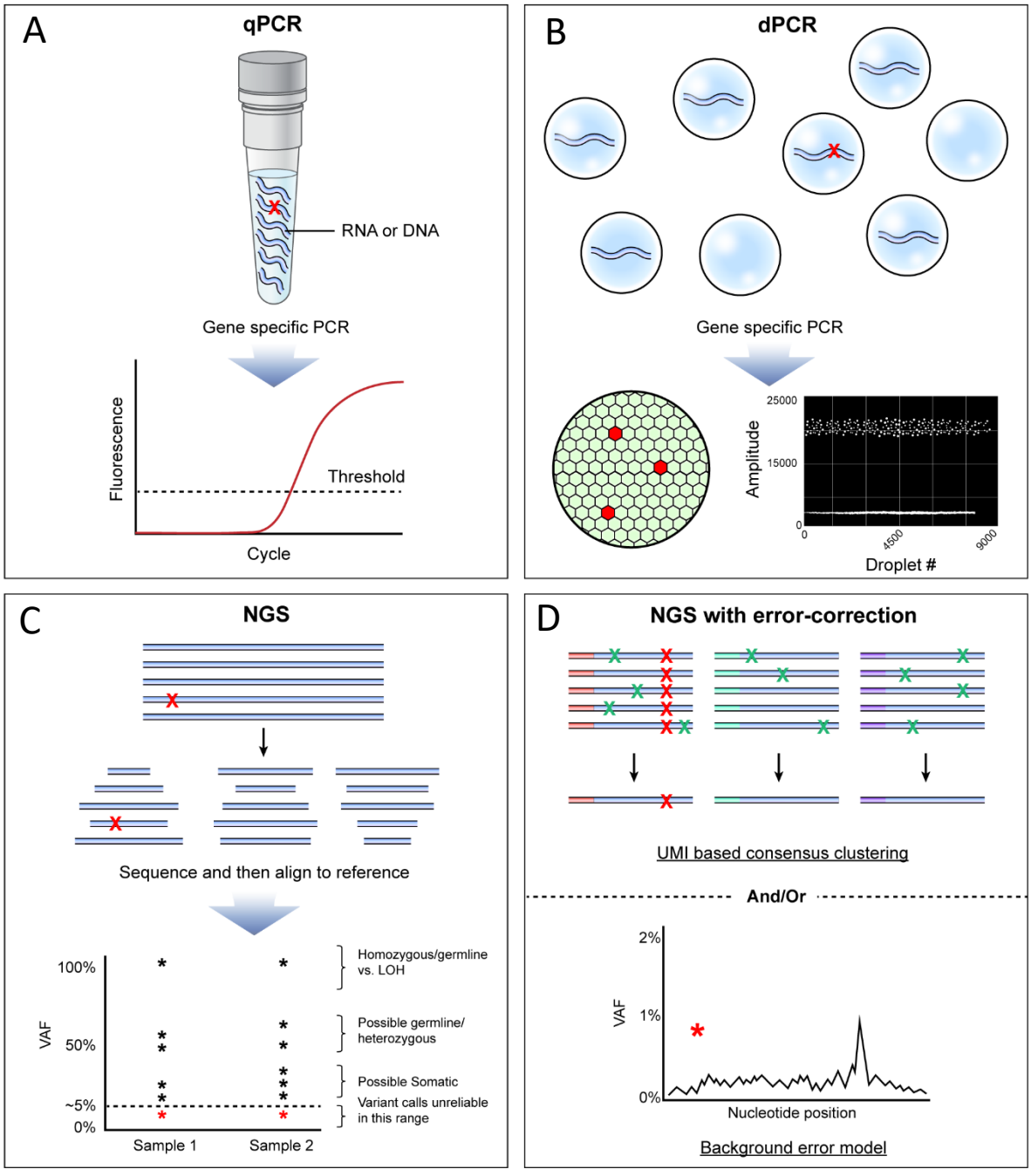
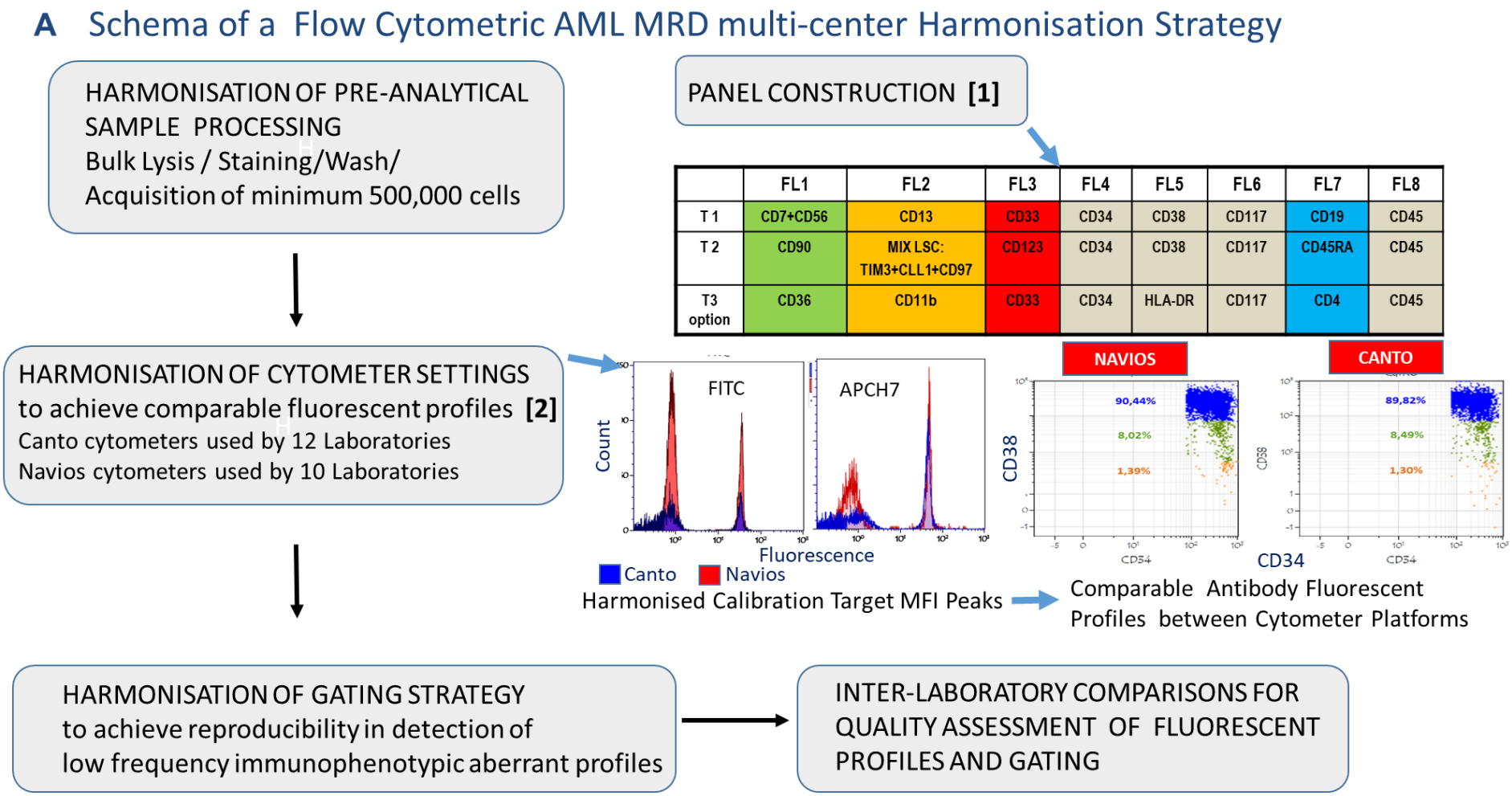
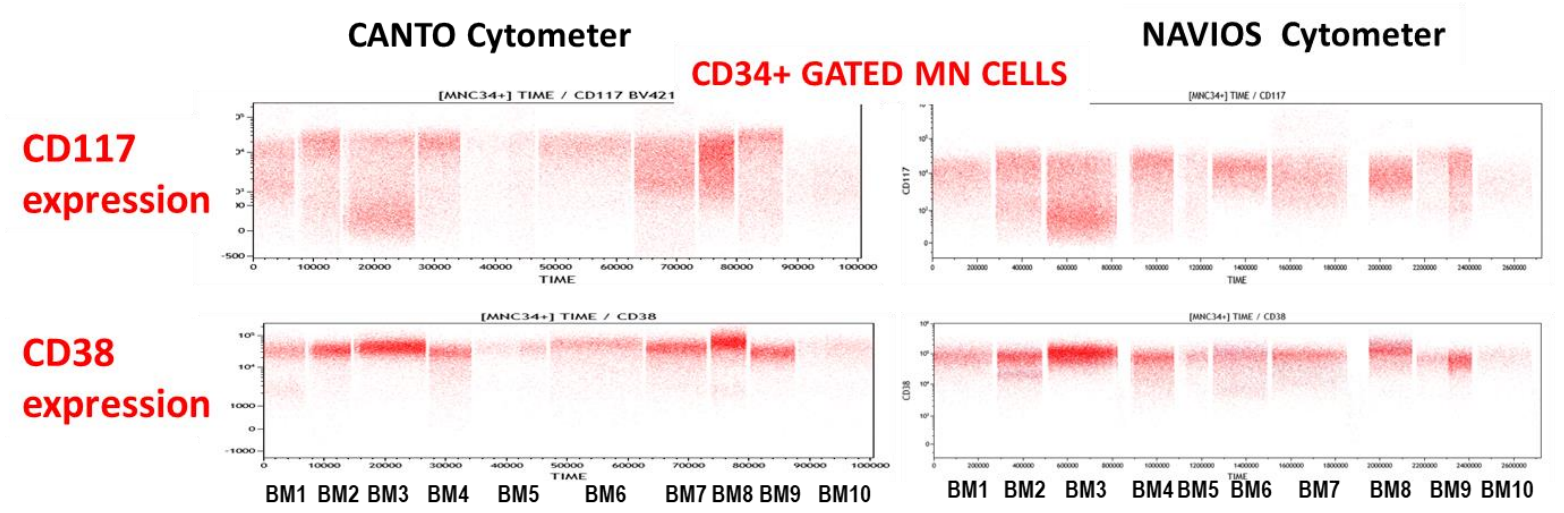

A
qPCR
RNA or DNA
Gene specific PCR
Fluorescence
Threshold
Cycle
B
dPCR
Gene specific PCR
Amplitude
25000
15000
0
0
4500
9000
Droplet #
C
NGS
Sequence and then align to reference
VAF
100%
50%
~5%
0%
Homozygous/germline vs. LOH
Possible germline/ heterozygous
Possible Somatic
Variant calls unreliable in this range
Sample 1
Sample 2
D
NGS with error-correction
UMI based consensus clustering
And/Or
VAF
2%
1%
0%
Nucleotide position
Background error model

Figure 3


A Schema of a Flow Cytometric AML MRD multi-center Harmonisation Strategy
HARMONISATION OF PRE-ANALYTICAL SAMPLE PROCESSING
Bulk Lysis / Staining/Wash/
Acquisition of minimum 500,000 cells
PANEL CONSTRUCTION [1]
FL1 FL2 FL3 FL4 FL5 FL6 FL7 FL8
T 1 CD7+CD56 CD13 CD33 CD34 CD38 CD117 CD19 CD45
T 2 CD90 MIX LSC: TIM3+CLL1+CD97 CD123 CD34 CD38 CD117 CD45RA CD45
T3 option CD36 CD11b CD33 CD34 HLA-DR CD117 CD4 CD45
HARMONISATION OF CYTOMETER SETTINGS
to achieve comparable fluorescent profiles [2]
Canto cytometers used by 12 Laboratories
Navios cytometers used by 10 Laboratories
FITC
APCH7
Count
Fluorescence
Canto
Navios
Harmonised Calibration Target MFI Peaks
NAVIOS
CANTO
CD38
CD34
90,44%
8,02%
1,39%
89,82%
8,49%
1,30%
Comparable Antibody Fluorescent Profiles between Cytometer Platforms
HARMONISATION OF GATING STRATEGY
to achieve reproducibility in detection of
low frequency immunophenotypic aberrant profiles
INTER-LABORATORY COMPARISONS FOR QUALITY ASSESSMENT OF FLUORESCENT PROFILES AND GATING

B Examples of Inter-Laboratory Quality Assessments

Inter-laboratory QA - similar marker expression profiles for 10 distributed BMs **[3]**



Inter-laboratory QA (22 laboratories) - reproducibility for rare event detection **[4]**

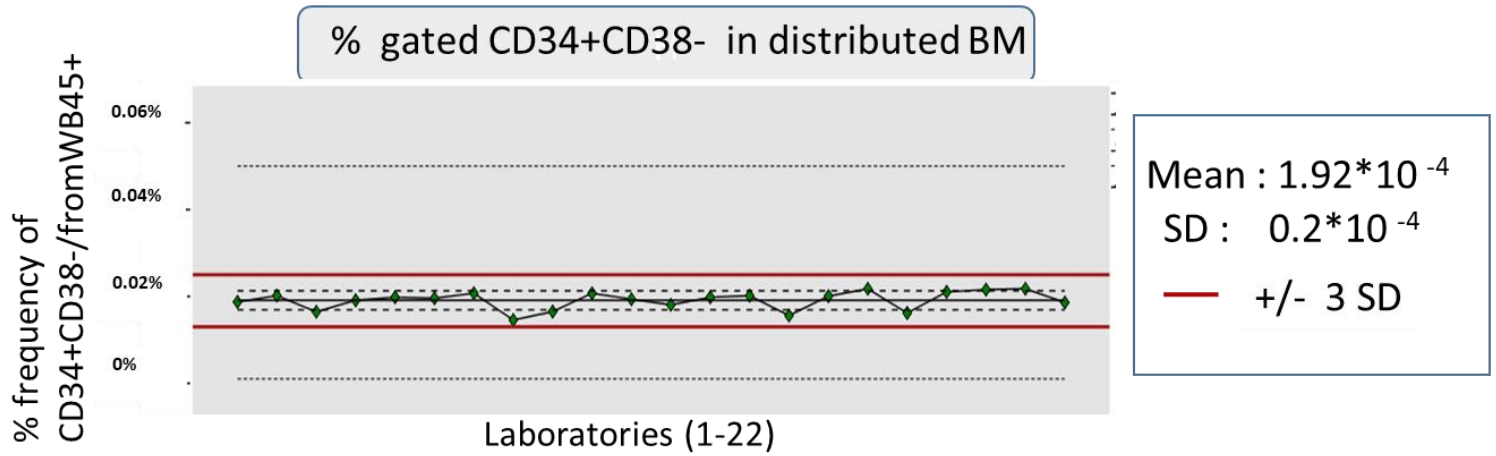
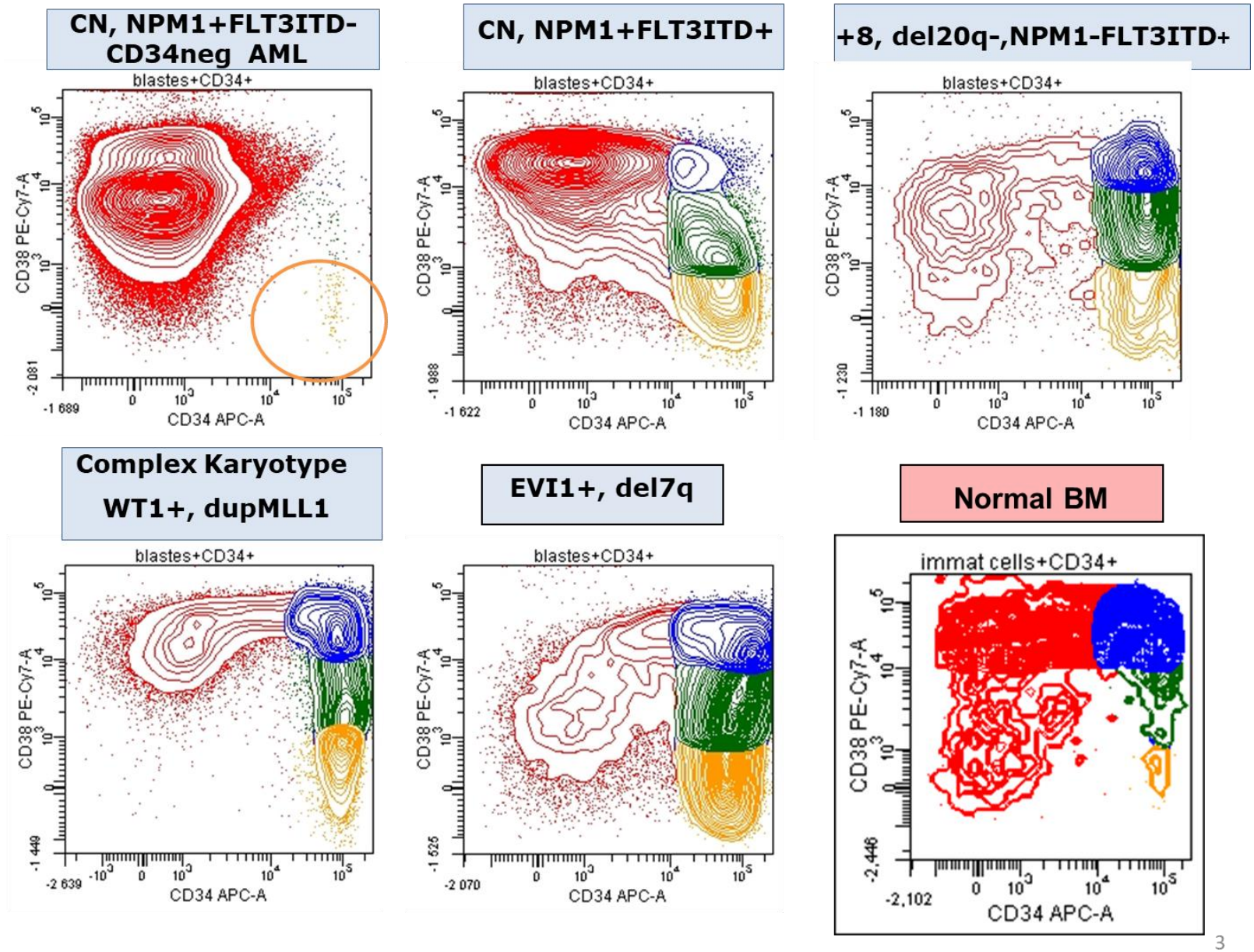


Figure 4

Figure 4


CN, NPM1+FLT3ITD-
CD34neg AML
blastes+CD34+
CD38 PE-Cy7-A
CD34 APC-A
CN, NPM1+FLT3ITD+
blastes+CD34+
CD38 PE-Cy7-A
CD34 APC-A
+8, del20q-,NPM1-FLT3ITD+
blastes+CD34+
CD38 PE-Cy7-A
CD34 APC-A
Complex Karyotype
WT1+, dupMLL1
blastes+CD34+
CD38 PE-Cy7-A
CD34 APC-A
EVI1+, del7q
blastes+CD34+
CD38 PE-Cy7-A
CD34 APC-A
Normal BM
immat cells+CD34+
CD38 PE-Cy7-A
CD34 APC-A