



Minimal residual disease

# Measurable residual disease as a biomarker in acute myeloid leukemia: theoretical and practical considerations

Roland B. Walter<sup>1,2,3,4</sup> · Yishai Ofran<sup>5</sup> · Agnieszka Wierzbowska<sup>6</sup> · Farhad Ravandi<sup>7</sup> · Christopher S. Hourigan<sup>8</sup> · Lok Lam Ngai<sup>9</sup> · Adriano Venditti<sup>10</sup> · Francesco Buccisano<sup>10</sup> · Gert J. Ossenkoppele<sup>9</sup> · Gail J. Roboz<sup>11</sup>

Received: 25 December 2020 / Revised: 19 February 2021 / Accepted: 11 March 2021 / Published online: 23 March 2021  
© The Author(s), under exclusive licence to Springer Nature Limited 2021

## Abstract

Several methodologies that rely on the detection of immunophenotypic or molecular abnormalities of the neoplastic cells are now available to quantify measurable (“minimal”) residual disease (MRD) in acute myeloid leukemia (AML). Although the perfect MRD test does not (yet) exist, the strong association between MRD and adverse patient outcomes has provided the impetus to use measures of MRD as biomarker in the routine care of AML patients and during clinical trials. MRD test results may inform the selection of postremission therapy in some patients but evidence supporting the use of MRD as predictive biomarker is still limited. Several retrospective studies have shown that conversion from undetectable to detectable MRD or increasing MRD over time is associated with overt disease recurrence, and MRD testing may therefore be valuable as a monitoring biomarker for early detection of relapse. Interpreting serial MRD data is complex, with open questions regarding the optimal timing and frequency of testing, as well as the identification of test-specific thresholds to define relapse. Importantly, it is unknown whether intervening at the time of MRD detection, rather than at overt disease recurrence, improves outcomes. Finally, using MRD as a surrogate efficacy-response biomarker to accelerate drug development/approval has already been accepted by regulatory authorities in other diseases and is of great interest as a potential strategy in AML. While the prognostic value of MRD in AML is well established, data from prospective clinical trials confirming that treatment effects on MRD directly relate to clinical outcomes are needed to further establish the role of MRD as a surrogate endpoint in AML.

## Introduction

Interest in identifying “minimal” residual disease—the population of leukemia cells that survives despite

morphological remission and causes disease relapse—dates back 40 years in acute myeloid leukemia (AML) [1–3]. More recently, technological advances in multiparameter flow cytometry (MFC), polymerase chain reaction (PCR),

✉ Roland B. Walter  
rwalter@fredhutch.org

<sup>1</sup> Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

<sup>2</sup> Department of Medicine, Division of Hematology, University of Washington, Seattle, WA, USA

<sup>3</sup> Department of Laboratory Medicine & Pathology, University of Washington, Seattle, WA, USA

<sup>4</sup> Department of Epidemiology, University of Washington, Seattle, WA, USA

<sup>5</sup> Department of Hematology, Sharee Zedeq Medical Center, Jerusalem, Israel

<sup>6</sup> Department of Hematology, Medical University of Lodz,

Lodz, Poland

<sup>7</sup> Department of Leukemia, University of Texas MD Anderson Cancer Center, Houston, TX, USA

<sup>8</sup> Laboratory of Myeloid Malignancies, Hematology Branch, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, USA

<sup>9</sup> Department of Hematology, Amsterdam UMC, VU University Medical Center, Amsterdam, The Netherlands

<sup>10</sup> Department of Biomedicine and Prevention, Hematology, University Tor Vergata, Rome, Italy

<sup>11</sup> Hematology and Oncology, Weill Cornell Medicine and NewYork-Presbyterian Hospital, Cornell University, New York, NY, USA

and next-generation sequencing (NGS) have allowed multimodality detection and quantification of measurable residual disease (MRD) and tracking of immunophenotypic and/or genetic/molecular abnormalities in AML cells [4]. Numerous retrospective studies have consistently shown a strong association between detection of MRD and adverse outcomes in AML patients [4–16]. Furthermore, recent meta-analyses have confirmed that the presence of MRD, by any measure, has a negative impact on overall survival (OS), even in the setting of allogeneic hematopoietic cell transplantation (HCT) [17, 18]. Thus, clinicians are eager to use MRD test results as a biological marker (“biomarker”) in the routine care of patients with AML, and investigators are hopeful to use such results as a surrogate endpoint in clinical trials to expedite drug testing and facilitate earlier access to novel therapeutics.

Biomarkers can be operationally defined as characteristics that are measured as indicators of normal biological processes, pathogenic processes, or biological responses to an exposure or intervention, including therapeutic interventions [19]. They play important roles in various aspects of clinical medicine, with applications in disease risk/susceptibility, disease diagnosis, and staging; as indicators of disease prognosis; to detect or monitor changes in the degree or extent of the disease over time; as sensitive predictors of impending clinical relapse; as measures of safety of exposure to a medical product or environmental agent; and as tools for the prediction and assessment of clinical responses to an intervention [19, 20]. Here, we will summarize current limitations of MRD assessments in AML and review theoretical and practical aspects for using MRD as a biomarker in AML clinical care and drug development. Since different considerations apply for different biomarker purposes, we will discuss potential uses of MRD as a prognostic, predictive, monitoring, and/or efficacy-response biomarker separately. Of note, MRD testing is relevant only after administration of antileukemia therapy and, thus, has no role as a diagnostic biomarker in AML.

### Limitations of MRD assessments: general and AML-specific considerations

A perfect MRD assay should accurately and precisely identify the population(s) of leukemia cells which, if left untreated, would cause disease recurrence, while being indifferent toward the other residual leukemia cells, including AML precursor or progeny cells, that do not cause relapse [5]. Although the technologies to detect MRD have improved over time and continue to evolve rapidly [4], it is clear that the perfect MRD test does not (yet) exist in AML (or any other disease) for biological and methodological reasons, an important limitation to keep in mind.

For an assay to successfully discriminate between cells that can vs. cannot cause relapse, both disease biology and technical aspects of the selected marker(s), measurements, and data analysis need to be aligned. The established clinical value of molecular MRD assays in acute promyelocytic leukemia (APL) [21] and chronic myeloid leukemia (CML) [22] reflect the existence of canonical genetic translocations that are essential for the pathogenesis of the leukemia and are present almost uniformly in all leukemia cells and subclones in these malignancies. Other disease-relevant canonical molecular aberrations, such as *nucleophosmin-1* (*NPM1*) mutations or core-binding factor (CBF) translocations, have also been identified as useful targets for MRD detection in non-APL AML, but are limited to specific patient subgroups. In general, however, the genetic heterogeneity of leukemia cells, both within an individual patient and between different patients, has substantially complicated the development of MRD assays for non-APL AML. Importantly, although we have detailed insights into the molecular complexity of AML, we are still unable to distinguish specific characteristics of the AML cells that cause relapse, and do not understand how we could separate such cells from the many other leukemia cells that do not have that potential [23]. The apparent diversity of cells capable of causing relapse is one important challenge for MRD assay development in non-APL AML. For example, data suggest that relapses can originate not only from rare AML stem cells but also from larger subclones of immunophenotypically committed leukemia cells that retain stem cell-like properties [24]. This observation may identify one of many potential mechanisms leading to the well-recognized phenomenon of clonal shift, and explains the finding that relapsing AML subclones may not express the original immunophenotypic and/or molecular abnormalities identified at the time of diagnosis.

There are a number of additional reasons why MRD tests in AML are imperfect [23]. First, normal and regenerating cells in the bone marrow and mutations associated with nonmalignant clonal hematopoiesis can provide immunophenotypic and molecular background “noise” that may interfere with the ability of an MRD assay to detect residual AML cells. In contrast, molecular MRD testing for specific gene rearrangements in acute lymphoblastic leukemia (ALL) or chronic lymphocytic leukemia (CLL) is relatively straightforward because malignant lymphoid cells are easily distinguished from normal ones in these leukemias with this strategy. Second, MRD testing is limited by methodological considerations related to the sensitivity, specificity, reproducibility, repeatability, and replicability of different assays. These concerns are currently being addressed by the European LeukemiaNet (ELN) MRD Working Party, with initial work presented in a first consensus document [4]. Further efforts are underway to reduce methodological

differences between laboratories by developing more detailed guidelines for assay standardization or, at the minimum, assay harmonization. A third problem with MRD testing in AML relates to sample procurement. The few milliliters of bone marrow in routine clinical samples may not be representative of the heterogeneous distribution of AML cells across the body and may limit the accuracy of MRD tests [25]. This problem, combined with variable skill levels in bone marrow aspiration, almost certainly contributes to false negative results. Quantifying MRD in the blood rather than bone marrow [26] may be one strategy to address this limitation, and additional studies to assess the value of peripheral blood MRD testing are underway. It should be noted that AML relapses can take place in body compartments, e.g., the central nervous system and skin, for which neither sampling of bone marrow nor peripheral blood is likely informative.

There are statistical limitations in AML MRD testing. For example, the true value of a detectable MRD reading could be obscured by patients dying from unrelated causes prior to AML relapse and/or delayed time of relapse beyond the designated observation period. Also, we typically reduce an MRD test to a binary read-out, such as detectable/undetectable, present/absent, or positive/negative, and this practice may result in reduced test performance. Such statistical reasons are not specific to MRD testing in AML but apply to MRD testing in general. Finally, therapeutic interventions may impact the relationship between an MRD test result and relapse in AML. In this regard of particular importance are immunologic “graft-versus-leukemia” effects conferred by allogeneic HCT, which may act on MRD to reduce the likelihood, or delay the occurrence, of AML relapse, thus potentially interfering with the predictive power of an MRD test.

## MRD as a prognostic biomarker in AML

A prognostic biomarker provides information about the natural history and outcomes of specific diseases by identifying the likelihood of a clinical event, e.g., disease recurrence or progression, in patients who have the disease or medical condition of interest. Specifically, a prognostic biomarker informs about the natural history of the disease in a particular patient in the absence of a therapeutic intervention [19, 27]. In practice, this situation is rarely encountered in the care of patients with AML, as treatment plans routinely include repeated courses of therapy, and cytotoxic, immunologic, and other antileukemic effects may persist well beyond the actual administration of the therapeutic.

The vast majority of investigations conducted to date to assess the role of MRD as possible biomarker in AML have

focused on prognostic information provided by results of MRD assays. Numerous studies have shown that, at the cohort level, results from MRD tests can risk-stratify patients in morphologic remission: those with MRD have higher cumulative incidence rates of relapse and, often, shorter relapse-free survival (RFS) and/or OS than similarly treated individuals without MRD. The strong association between detectable MRD and inferior patient outcomes has been confirmed at several timepoints throughout the course of intensive AML therapy: as early as several days after the start of induction chemotherapy; after completion of one or two courses of induction chemotherapy; after postremission therapy; both before and after HCT; and after salvage chemotherapy for relapsed/refractory disease. Furthermore, the negative prognostic impact of a positive MRD test on outcomes has been demonstrated irrespective of MRD testing methodologies, e.g., MFC, quantitative PCR for single genetic abnormalities and NGS for multiple molecular abnormalities [4–14]. MRD test results were routinely found to be the most important adverse factor in univariate analyses and, often, the only significant one remaining as an independent factor in multivariable models. Overall, the available data indicate that patients who test positive for MRD at any given timepoint, regardless of the detection methodology used, have a high but not guaranteed likelihood of experiencing relapse. On the other hand, not all patients without MRD will remain in remission. These observations may be related to assay performance, disease biology, patient selection, therapeutic intervention, or a combination of these factors, as discussed above. On an individual patient level, results from MRD assessments refine the prediction of RFS and OS to some degree, but the ability to predict these outcomes accurately remains limited [28]. Of note, the vast majority of studies conducted to date have assessed patients who received intensive therapies for remission induction. The prognostic role of MRD after administration of lower-intensity therapies is less well established, but emerging data suggest that detectable MRD is associated with inferior outcomes after such therapies as well [29, 30].

Studies of MRD in AML have differed not only in the methodology for MRD assessment, but also with regard to patient characteristics, disease status, and type of treatment. A literature-based meta-analysis was recently conducted to further delineate the prognostic role of MRD in AML [18]. Using Bayesian hierarchical modeling, this meta-analysis included 11,151 patients described in 81 studies published between January 1, 2000 and October 1, 2018 that reported on OS (17 publications; 3118 patients), RFS (20 publications; 1783 patients), or both (44 publications; 6250 patients). At 5 years, the estimated OS was 68% (95% Bayesian credible interval: 63–73%) for patients without MRD vs. 34% (28–40%) for those with MRD. Similar

5-year RFS estimates were 64% (59–70%) and 25% (20–32%), respectively. The relative benefit of not having MRD was comparable for both OS and DFS (average hazard ratio: 0.36 [0.33–0.39] for OS and 0.37 [0.34–0.40] for DFS). Absence of MRD was associated with superior RFS and OS across all age groups (adult or pediatric), MRD assessment timepoint (induction, during consolidation or after consolidation), AML subgroup (CBF or non-CBF), and specimen source (bone marrow or peripheral blood). The effect of MRD on survival was more profound in studies reporting outcomes of CBF AML compared to non-CBF AML. Overall, multivariable analyses, performed to control for possible confounding factors, were consistent with the univariate results [18]. Together, these data strongly support the use of MRD as a prognostic biomarker in AML. While different MRD assay methodologies can be used to provide prognostic information, it is important to note that the concordance between these assays is currently not absolute and, thus, it may be most valuable to use different MRD assessments in a complementary, rather than isolated manner. For example, retrospective studies have shown that when both MFC and NGS assays are used, patients without MRD by both methodologies have particularly good outcomes, patients with MRD by both methodologies have particularly poor outcomes, and patients with MRD by one methodology but without MRD by the other have intermediate outcomes [31, 32]. It is an open question whether further refinements in MRD detection methodologies will allow optimal prognostic information to be obtained from a single assay.

### MRD as a predictive biomarker in AML

A predictive biomarker is used to identify individuals who are more likely than similar individuals without the biomarker to experience a favorable or unfavorable effect from exposure to a medical product or an environmental agent [19]. Increasing data suggest that MRD may play a role as predictive biomarker in AML for some treatment situations and subsets of patients. For example, postremission therapies are generally divided into transplant and nontransplant strategies, which carry significantly different risks and toxicities. Since MRD assessments can, as shown above, stratify patients based on risk of disease recurrence, it is appealing to consider MRD as a marker to inform the allocation of a patient to a particular type of postremission therapy. Allogeneic HCT is associated with reduced likelihood of relapse compared to nontransplant postremission therapy but bears considerable risks of nonrelapse morbidity and mortality. While AML patients with MRD prior to allogeneic HCT have inferior outcomes to those without, retrospective analyses have shown that the relative

reduction in the risk of post-HCT AML recurrence is similar for adults with AML who have vs. do not have MRD at the time of allografting [33]. Because of the higher absolute risk of relapse for patients with pre-HCT MRD, the absolute reduction of relapse risk is greater for these patients. Since the risks for nonrelapse mortality (NRM) are relatively similar for patients with and without pre-HCT MRD, the absolute benefit of allografting may be greater for patients with pre-HCT MRD. Consistent with this notion, retrospective analyses of patients with CBF AML nonrandomly assigned to either allogeneic HCT or chemotherapy-based postremission therapy have indeed suggested better outcomes when allogeneic HCT was used in individuals with pre-HCT MRD; in those without pre-HCT MRD, outcomes with postremission chemotherapy were superior because NRM with allografting more than offset the reduced risk of relapse [34]. Likewise, data from the GIMEMA AML1310 trial have indicated that MRD-directed selection of postremission treatment strategy (autologous vs. allogeneic HCT) for AML with intermediate-risk AML in first morphologic remission might optimize treatment outcomes [35].

Some studies suggest MRD may also have value as predictive biomarker to inform selection of the optimal conditioning intensity before allogeneic HCT. Several retrospective studies of patients nonrandomly assigned to receive higher- or lower-intensity conditioning regimens suggested lower relapse rates with myeloablative conditioning (MAC) compared to reduced-intensity conditioning (RIC) or nonmyeloablative conditioning [36–40]. Concordant with these findings, data from the randomized phase 3 BMT CTN 0901 trial showed that, for adults age 18–65 years with AML transplanted in morphologic remission, MAC was associated with lower relapse rates and longer survival compared to RIC [41]. In a recent post hoc analysis of a subset of 190 AML patients (>70% older than age 50) transplanted on the BMT CTN 0901 trial for whom pretransplant peripheral blood specimens were archived, Hourigan et al. used ultra-deep, error-corrected sequencing of 13 commonly mutated genes in AML as an approach to test for mutation-defined MRD before HCT [42]. Results showed that there was a statistically significantly lower incidence of relapse, as well as longer RFS and OS, with MAC in the 66% of patients entering transplantation with genomic evidence of residual AML (or, more specifically, the 41% of patients with mutations present in genes other than *DNMT3A*, *TET2*, and *ASXL1*) compared with those randomized to RIC. On the other hand, in the patients without detectable genomic MRD, MAC was associated with only a statistically nonsignificant improvement in relapse incidence, higher NRM, and similar OS compared to RIC [42]. Slightly different from these data are findings from FIGARO, an open label phase 2



randomized trial of 244 patients with high-risk AML or myelodysplastic syndrome randomly assigned 1:1 to a fludarabine-based RIC regimen or an “augmented” RIC regimen with FLAMSA-busulfan [43]. This study confirmed the poor prognosis associated with the presence of pre-HCT MRD, in this case measured by MFC, in patients receiving RIC. In this study, unlike MAC in BMT CTN 0901, “augmented” RIC could not overcome the negative prognosis associated with pretransplant MRD. Furthermore, ~50% of patients on the FIGARO trial with evidence of pretransplant MRD did not relapse within the study follow-up period, underscoring the observation from both clinical trials and clinical practice that not all patients with MRD are destined to relapse quickly [39]. The BMT CTN 0910 and FIGARO trials highlight the need for additional studies to delineate further how MRD before HCT should inform the selection of specific conditioning regimens and to determine the mechanism of any benefit associated with conditioning intensification.

## MRD as a monitoring biomarker in AML

A monitoring biomarker is measured serially and used to assess the status of a disease or medical condition or as evidence of exposure to (or effect of) a medical product or an environmental agent [19]. Data showing that conversion from a negative to a positive MRD test or an increase in MRD over time is associated with overt disease recurrence provides the rationale to consider MRD as a monitoring biomarker for routine surveillance and care of patients following completion of AML therapy or, perhaps, during maintenance treatment [4]. Conversely, demonstration of serially negative MRD tests could provide the basis to withhold further therapy or change treatment strategy, somewhat analogous to patients with CML who test negative or minimally positive for the *breakpoint cluster region1/Abelson murine leukemia viral oncogene-1* fusion gene by sensitive quantitative RT-PCR assay. This MRD test result informs on a very low risk of disease recurrence or progression, at least with continued use of tyrosine kinase inhibitor (TKI) therapy [22], and has been used to help decision-making regarding TKI discontinuation.

In the case of APL, conversion to detectable MRD, using a sensitive quantitative RT-PCR assay for the *promyelocytic leukemia/retinoic acid receptor alpha* fusion protein, is almost always followed by hematologic relapse, although the interval between conversion of the MRD test and overt relapse can span more than 1 year [44]. Likewise, conversion to a positive RT-PCR test for *RUNX1/RUNX1T1* transcripts in patients with t(8;21)(q22;q22.1) leukemia is strongly indicative of overt disease recurrence, often with a very short latency from molecular to morphologic relapse,

which necessitates MRD assessments at short intervals [45]. Both of these examples refer to serial monitoring using highly sensitive and specific RT-PCR assays in molecularly defined AML subgroups. The interpretation of serial MRD data in other AML subgroups is much more complex and leads to a number of largely unresolved challenges and questions. For example, more data are required regarding the need and timing for confirmatory testing if a positive result is obtained, the thresholds best suited to define relapse, and how to approach patients with molecular MRD persistence at low copy numbers. Data are also needed to define the optimal timing and interval between tests, which may differ based on the cytogenetic and/or molecular characteristics of the leukemia [46–49] and, possibly, on the interval since completion of therapy. Monitoring MRD serially every 3 months, as has been recommended [4], may not be ideal for some patients and may provide insufficient lead time to identify MRD-level relapse in patients with rapid relapse kinetics [45, 50]. Finally, while the value of MRD conversion as indicator of impending relapse is established, at least for some molecularly defined AML subgroups, not all patients with conversion of a negative to a positive MRD test ultimately relapse even in the absence of further AML therapy. Most importantly, clinical benefit from early intervention based on MRD data has yet to be shown for most AML types and treatment scenarios. Data from NCRI AML17 trial, in which patients were allocated to MRD monitoring or not after completion of chemotherapy, will provide insight into the practical use of MRD monitoring in AML and help clarify what impact such a strategy has on clinical decision-making and, by offering an early (“preemptive”) treatment opportunity, on treatment outcomes.

In the era before all-*trans* retinoic acid (ATRA)/arsenic trioxide (ATO)-based upfront therapy of APL, data from a prospective, nonrandomized study as well as retrospective analyses suggested that treatment at the time of molecular relapse can prevent overt relapse in the majority of patients and may lead to longer survival [51–53]. Because of the very low risk of relapse following ATRA/ATO-based upfront therapy, however, serial MRD monitoring is no longer recommended in low-risk APL [21]. It is likely that serial MRD measurements in non-APL AML patients will identify many with increasing disease burden before overt disease recurrence. Early relapse detection may allow early therapeutic intervention, and many efforts are ongoing to develop drugs that effectively “eradicate” MRD as a novel therapeutic strategy in frontline and maintenance settings. However, while various interventions at the MRD level of disease recurrence are being explored [54–56], it is currently unknown whether an early treatment strategy (which would potentially deliver unnecessary additional therapy to some patients not destined to relapse) will lead to better

outcomes compared to the strategy of treating only at the time of overt relapse which, by definition, results in treatment of only those patients who need it. MRD-based intervention trials that include serial MRD monitoring are underway.

## MRD as an efficacy-response biomarker in AML

An efficacy-response biomarker is used to show that a response has occurred in an individual who has been exposed to a medical product or an environmental agent [19, 27]. The consistently observed, strong association between MRD assessments in blood and/or bone marrow at different timepoints (in some [57] but not all [58] studies, as early as after 3 days of treatment) and relapse risk and/or survival has raised interest in using MRD as a surrogate efficacy-response biomarker, i.e., a marker thought to predict a clinical outcome that is not itself a measure of clinical benefit [59], to accelerate drug development/testing and regulatory approval. Regulatory drug approval requires demonstration of clinical benefit, most typically an improvement in either survival or disease-associated symptoms and quality of life. Even in a disease such as AML where survival is relatively short in many patient subsets, demonstration of improved survival may require long follow-up, and the effects of subsequent treatments, which are rapidly evolving, can confound the effect of the therapy of interest. A validated, early post-therapy surrogate endpoint for clinical benefit would address these limitations and could accelerate and simplify drug testing and approval. Such a surrogate endpoint could also lead to shorter clinical trials, reduced costs, and exposure of fewer patients to potentially toxic and/or ineffective treatments. There are now examples of disease areas, notably ALL, in which MRD is already accepted by some regulatory authorities as a surrogate endpoint.

The U.S. Food & Drug Administration (FDA) has issued a guidance document regarding the regulatory considerations for the use of MRD in the development of therapeutic drugs and biological products [27]. This document provides a conceptual framework how MRD data could serve as the basis for accelerated or even traditional approval, depending on the strength of the evidence supporting surrogacy. The guidance document adheres closely to principles described previously in the statistical literature regarding which factors have to be considered for the strength of this evidence. One such factor is biological plausibility. Fundamentally, an MRD assay provides a quantitative assessment of the number of residual leukemia cells. Therefore, we assume that becoming MRD test negative is a biologically plausible surrogate for longer survival or is at least one of the

prerequisites. A second factor is the availability of epidemiological studies demonstrating the prognostic value of the surrogate endpoint for the clinical outcome, e.g., achieving a complete remission without MRD must correlate with longer survival compared to achieving a complete remission with MRD. This is measured at the individual patient level. Single arm trial data can be used for the purpose of hypothesis generation, but ultimately data from meta-analyses, such as those discussed above [18], are required for formal assessment. Further efforts are currently ongoing to generate additional meta-analysis data to support the prognostic value of results from MRD assessments. A third important factor supporting surrogacy is the availability of clinical trial evidence demonstrating that treatment effects on the surrogate endpoint correspond to treatment effects on the clinical outcome, i.e., the experimental treatment needs to increase both the rate of complete remissions without MRD and survival compared to the control treatment. This effect is measured at the trial level, and the FDA guidance document suggests that meta-analyses conducted to assess this correlation should only include data from randomized trials [27]. Currently, while some data from mostly non-randomized trials show a treatment effect on both MRD responses and survival [60–62], data from randomized trials that may support this requirement are currently extremely limited. One existing example is the AMLSG 09-09 trial, in which 588 patients with newly diagnosed *NPM1*-mutated AML were randomized to intensive chemotherapy plus ATRA with or without the CD33 antibody-drug conjugate gemtuzumab ozogamicin (GO) [63]. In the GO arm, *NPM1* transcript levels were significantly lower, and a significantly higher proportion of patients achieved a remission without MRD than in the control arm. This was associated with a lower incidence of relapse and better RFS with GO. Prospective clinical trials aimed at demonstrating a treatment effect on both MRD and survival in AML are ongoing in several cooperative study groups (e.g., PALG, HOVON, AMLSG, NCRI). Recent data from the randomized phase 3 QUAZAR AML-001 trial of maintenance therapy with oral azacitidine (CC-486) vs. placebo showed a significant benefit in OS for oral azacitidine, regardless of whether MRD was detectable or not at baseline. Of note, almost 20% of patients with detectable MRD at baseline who were assigned to the placebo arm still converted to MRD negativity during follow-up, highlighting the challenge of using MRD as a possible efficacy-response biomarker in AML [64]. Still, data supporting the use of MRD assessments for this purpose are already available from other hematologic malignancies. In CLL, a meta-analysis of three large randomized chemoimmunotherapy trials showed a statistically significant relationship between treatment effect on peripheral blood MRD and treatment effect on progression-free survival (PFS), supporting the use of peripheral blood MRD

as a surrogate for PFS [65]. Likewise, a meta-analysis of six randomized trials in newly diagnosed multiple myeloma should show a strong correlation between the treatment effect on the odds ratio for responses without MRD and the hazard ratio for PFS [66]. In ALL, MRD has robust prognostic significance and has been successfully used as a primary endpoint in a clinical trial leading to the FDA approval of blinatumomab for relapsed disease [67]. However, a meta-analysis of two large, randomized phase 3 trials investigating the effects of different corticosteroids on disease outcome in children with ALL found that MRD at the end of induction was a poor surrogate for the treatment effect on event-free survival (EFS) at the trial level [68]. Thus, study population (e.g., newly diagnosed vs. relapsed disease), age, and timepoint of assessment may be important factors in determining the suitability of MRD as a biomarker as surrogate for treatment efficacy.

### Conduct of clinical trials that assess value of MRD measures as surrogate biomarkers

There is an ongoing need for well-designed, prospective, randomized trials with sufficient statistical power to assess the role of MRD as a surrogate biomarker in AML. To date, MRD negativity has not been used as a primary endpoint, but clinical trials investigating MRD as a secondary and/or exploratory endpoint are underway. The FDA guidance document on the regulatory considerations for the use of MRD [27] outlines the general framework for the conduct of such trials. It is recommended that intention-to-treat analyses be used to evaluate MRD as an endpoint as an unbiased estimate of treatment effect, and that any patient without an MRD assessment (e.g., because of failed sample collection, methodological/technical issues, failed sample shipment, etc.) be considered as not responsive to treatment. Since the latter issue will certainly dilute the treatment difference of interest, special efforts should be made to minimize the number of missing and/or unevaluable MRD assessments. To help with these efforts, the study protocol should provide explicit and complete information regarding MRD sample collection, including site (e.g., bone marrow or peripheral blood), sample volume and, ideally, sequence of bone marrow pulls for the different analyses to reduce the risk of inadequate sample quality (e.g., due to bone marrow dilution) for MRD testing. Prespecified timepoints or, if fixed timepoints are not feasible, time windows for MRD sample collection are essential. The ELN MRD Working Party recommends that molecular and/or flow cytometric assessments of MRD should be performed whenever a treatment response is evaluated [4].

As no single approach to detect or quantify MRD has been proven superior in heterogeneous AML patient

populations, multimodality MRD testing should be considered, with the use of preferred detection methodologies for specific AML subtypes (i.e., molecular assays to detect translocations in APL or CBF leukemias and canonical mutations in *NPM1*-mutated AML) [4]. A prespecified, standardized, and validated threshold to distinguish detectable MRD from undetectable MRD, which optimizes the positive and negative predictive values for each technique, should be used. To reduce methodology- or staff-related test result variabilities that complicate the interpretation of MRD results, it is highly recommended to perform MRD testing in experienced central laboratories. Availability of a network of harmonized laboratories with standardized procedures could facilitate the conduct of larger, international trials, improve the quality of samples being tested, and ultimately optimize the MRD testing itself. This approach would also facilitate comparisons between clinical studies [5]. Finally, attention is required in the choice of long-term clinical endpoints, e.g., EFS, DFS, or OS. These should be prespecified during the study planning period and should be precisely defined, with explicit descriptions as to whether/how disease recurrence at the MRD level impacts the outcomes defined.

### Conclusions and future perspective

Clinicians treating AML have always known that morphological complete remission belies the dangerous leftover leukemia cells that hide after treatment and eventually emerge to cause relapse and bone marrow failure. With nine new drug approvals by the FDA in the last 3 years, more AML patients than ever are achieving complete remission, and OS is improving. Over the same period of time, substantial data have emerged unequivocally supporting the use of MRD as a prognostic biomarker in AML. The assessment of AML MRD is undeniably complicated, with many unanswered questions related to optimization of technologies, methodologies, laboratory procedures, timing of sample procurement, and other significant issues (Table 1). Still, there has been substantial headway in advancing the utility of MRD as an essential biomarker in AML. While MRD remains an exploratory endpoint in clinical trials, consensus guidelines have been issued and are being updated; prospective, randomized clinical trials are underway with multimodality assessment of MRD at different timepoints; clinical trials are ongoing with MRD-directed interventions; and guidance from regulatory authorities has been issued. Also, both clinicians and patients are paying increasing attention to the potential value of MRD assessments as a tool to help individual treatment decision-making. While much more work remains, it is anticipated that the challenges limiting routine

**Table 1** MRD as a biomarker in AML.

What we know	What we do not know
<ul style="list-style-type: none"> <li>• Several methodologies (e.g., MFC, PCR, and NGS) can be used to detect and quantify MRD.</li> <li>• The perfect MRD assays does not (yet) exist for biological, technical, and statistical reasons.</li> <li>• Retrospective studies, confirmed by meta-analysis, show a strong correlation between detectable MRD by any method and worse clinical outcomes, including higher rates of relapse, shorter relapse-free survival, and worse overall survival</li> <li>• Not all patients with MRD will relapse and not all patients without MRD will remain in remission.</li> </ul>	<ul style="list-style-type: none"> <li>• What are the optimal MRD assays and their most ideal assay and clinical conditions?</li> <li>• Should patients have multimodality assessments of MRD and what should be done if they are discordant?</li> <li>• Are certain assays better in certain treatment settings?</li> <li>• How should we track MRD if the relapsing cells change immunophenotypically and/or molecularly from baseline?</li> <li>• What is the optimal specimen (blood, bone marrow) and specimen procurement technique to ensure adequate material for MRD detection and quantification?</li> <li>• How should MRD data inform the selection of subsequent therapy?</li> <li>• What are the optimal timepoints and frequency to assess and monitor MRD and should these change according to the type of AML and treatment administered?</li> <li>• Does “preemptive” treatment at the MRD level lead to better outcomes than treatment at the time of morphologic relapse?</li> <li>• How should MRD serve as surrogate efficacy-response biomarker to accelerate and simplify drug testing and approval?</li> <li>• Does eliminating MRD improve clinical outcomes? For all AML subgroups? All ages? All treatment settings (e.g., intensive vs. nonintensive chemo; pre- versus postallogeic transplant)?</li> </ul>

application of MRD testing in AML will be surmounted, and that the role of MRD as a predictive, monitoring, and efficacy-response biomarker in AML will be determined in detail.

**Acknowledgements** 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.

### Compliance with ethical standards

**Conflict of interest** The laboratory of CSH has received research funding from Merck Sharp & Dohme and Sellas Life Sciences. The other authors declare no competing financial interests.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### References

1. Hagenbeek A, Martens ACM. Kinetics of minimal residual disease in a rat model for human acute myelocytic leukemia. In: Baum SJ, Ledney GD, van Bekkum DW, editors. *Experimental hematology today*. New York: Springer; 1980. p. 215–21.
2. Hagenbeek A, Martens ACM. Minimal residual disease in acute leukemia: from experimental models to man. *Bone Marrow Transplant*. 1989;4:68–9.
3. Freireich EJ, Cork A, Stass SA, McCredie KB, Keating MJ, Estey EH, et al. Cytogenetics for detection of minimal residual disease in acute myeloblastic leukemia. *Leukemia*. 1992;6:500–6.
4. Schuurhuis GJ, Heuser M, Freeman S, Béné MC, Buccisano F, Cloos J, et al. Minimal/measurable residual disease in AML: a consensus document from the European LeukemiaNet MRD Working Party. *Blood*. 2018;131:1275–91.
5. Hourigan CS, Gale RP, Gormley NJ, Ossenkoppele GJ, Walter RB. Measurable residual disease testing in acute myeloid leukaemia. *Leukemia*. 2017;31:1482–90.
6. Ravandi F, Walter RB, Freeman SD. Evaluating measurable residual disease in acute myeloid leukemia. *Blood Adv*. 2018;2:1356–66.
7. Buccisano F, Maurillo L, Schuurhuis GJ, Del Principe MI, Di Veroli A, Gurnari C, et al. The emerging role of measurable residual disease detection in AML in morphologic remission. *Semin Hematol*. 2019;56:125–30.
8. Ossenkoppele G, Schuurhuis GJ, van de Loosdrecht A, Cloos J. Can we incorporate MRD assessment into clinical practice in AML? *Best Pract Res Clin Haematol*. 2019;32:186–91.
9. Voso MT, Ottone T, Lavorgna S, Venditti A, Maurillo L, Lo-Coco F, et al. MRD in AML: the role of new techniques. *Front Oncol*. 2019;9:655.
10. Freeman SD, Hourigan CS. MRD evaluation of AML in clinical practice: are we there yet? *Hematol Am Soc Hematol Educ Program*. 2019;2019:557–69.
11. Ghannam J, Dillon LW, Hourigan CS. Next-generation sequencing for measurable residual disease detection in acute myeloid leukaemia. *Br J Haematol*. 2020;188:77–85.
12. Paterno G, Del Principe MI, Venditti A. Detection and management of acute myeloid leukemia measurable residual disease: is it standard of care? *Curr Opin Hematol*. 2020;27:81–7.
13. Schwind S, Jentzsch M, Bach E, Stasik S, Thiede C, Platzbecker U. Use of minimal residual disease in acute myeloid leukemia therapy. *Curr Treat Options Oncol*. 2020;21:8.



14. Yoest JM, Shirai CL, Duncavage EJ. Sequencing-based measurable residual disease testing in acute myeloid leukemia. *Front Cell Dev Biol.* 2020;8:249.
15. Short NJ, Rafei H, Daver N, Hwang H, Ning J, Jorgensen JL, et al. Prognostic impact of complete remission with MRD negativity in patients with relapsed or refractory AML. *Blood Adv.* 2020;4:6117–26.
16. Mushtaq MU, Harrington AM, Chaudhary SG, Michaelis LC, Carlson KB, Abedin S, et al. Comparison of salvage chemotherapy regimens and prognostic significance of minimal residual disease in relapsed/refractory acute myeloid leukemia. *Leuk Lymphoma.* 2021;62:158–66.
17. Buckley SA, Wood BL, Othus M, Hourigan CS, Ustun C, Linden MA, et al. Minimal residual disease prior to allogeneic hematopoietic cell transplantation in acute myeloid leukemia: a meta-analysis. *Haematologica.* 2017;102:865–73.
18. Short NJ, Zhou S, Fu C, Berry DA, Walter RB, Freeman SD, et al. Impact of measurable residual disease on survival outcomes in patients with acute myeloid leukemia: a meta-analysis. *JAMA Oncol.* 2020;6:1890–9.
19. FDA-NIH Biomarker Working Group. BEST (Biomarkers, Endpoints, and other Tools) resource. FDA-NIH Biomarker Working Group. 2021. <https://www.ncbi.nlm.nih.gov/books/NBK338448/>. Accessed 27 Jul 2020.
20. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharm Ther.* 2001;69:89–95.
21. Sanz MA, Fenau P, Tallman MS, Estey EH, Löwenberg B, Naoe T, et al. Management of acute promyelocytic leukemia: updated recommendations from an expert panel of the European LeukemiaNet. *Blood.* 2019;133:1630–43.
22. Hochhaus A, Baccarani M, Silver RT, Schiffer C, Apperley JF, Cervantes F, et al. European LeukemiaNet 2020 recommendations for treating chronic myeloid leukemia. *Leukemia.* 2020;34:966–84.
23. Othus M, Gale RP, Hourigan CS, Walter RB. Statistics and measurable residual disease (MRD) testing: uses and abuses in hematopoietic cell transplantation. *Bone Marrow Transplant.* 2020;55:843–50.
24. Shlush LI, Mitchell A, Heisler L, Abelson S, Ng SWK, Trotman-Grant A, et al. Tracing the origins of relapse in acute myeloid leukaemia to stem cells. *Nature.* 2017;547:104–8.
25. Latham S, Hughes E, Budgen B, Mechinaud F, Crock C, Ekert H, et al. Sources of error in measurement of minimal residual disease in childhood acute lymphoblastic leukemia. *PLoS ONE.* 2017;12:e0185556.
26. Godwin CD, Zhou Y, Othus M, Asmuth MM, Shaw CM, Gardner KM, et al. Acute myeloid leukemia measurable residual disease detection by flow cytometry in peripheral blood versus bone marrow. *Blood.* 2021;137:569–72.
27. 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—guidance for industry. U.S. Department of Health and Human Services. 2020. <https://www.fda.gov/vaccines-blood-biologics/guidance-compliance-regulatory-information-biologics/biologics-guidances>. Accessed 4 Jul 2020.
28. Othus M, Wood BL, Stirewalt DL, Estey EH, Petersdorf SH, Appelbaum FR, et al. Effect of measurable ('minimal') residual disease (MRD) information on prediction of relapse and survival in adult acute myeloid leukemia. *Leukemia.* 2016;30:2080–3.
29. Boddu P, Jorgensen J, Kantarjian H, Borthakur G, Kadia T, Daver N, 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;32:241–4.
30. Winters AC, Gutman JA, Purev E, Nakic M, Tobin J, Chase S, et al. Real-world experience of venetoclax with azacitidine for untreated patients with acute myeloid leukemia. *Blood Adv.* 2019;3:2911–9.
31. Getta BM, Devlin SM, Levine RL, Arcila ME, Mohanty AS, Zehir A, 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;23:1064–71.
32. Jongen-Lavrencic M, Grob T, Hanekamp D, Kavelaars FG, Al Hinai A, Zeilemaker A, et al. Molecular minimal residual disease in acute myeloid leukemia. *N Engl J Med.* 2018;378:1189–99.
33. Versluis J, Kalin B, Zeijlemaker W, Passweg J, Graux C, Manz MG, 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;1:1–13.
34. Zhu HH, Zhang XH, Qin YZ, Liu DH, Jiang H, Chen H, et al. MRD-directed risk stratification treatment may improve outcomes of t(8;21) AML in the first complete remission: results from the AML05 multicenter trial. *Blood.* 2013;121:4056–62.
35. Venditti A, Picocchi A, Candoni A, Melillo L, Calafiore V, Cairoli R, et al. GIMEMA AML1310 trial of risk-adapted, MRD-directed therapy for young adults with newly diagnosed acute myeloid leukemia. *Blood.* 2019;134:935–45.
36. Aoudjhane M, Labopin M, Gorin NC, Shimoni A, Ruutu T, Kolb HJ, et al. Comparative outcome of reduced intensity and myeloablative conditioning regimen in HLA identical sibling allogeneic haematopoietic stem cell transplantation for patients older than 50 years of age with acute myeloblastic leukaemia: a retrospective survey from the Acute Leukemia Working Party (ALWP) of the European group for Blood and Marrow Transplantation (EBMT). *Leukemia.* 2005;19:2304–12.
37. Shimoni A, Hardan I, Shem-Tov N, Yeshurun M, Yerushalmi R, Avigdor A, et al. Allogeneic hematopoietic stem-cell transplantation in AML and MDS using myeloablative versus reduced-intensity conditioning: the role of dose intensity. *Leukemia.* 2006;20:322–8.
38. Alyea EP, Kim HT, Ho V, Cutler C, DeAngelo DJ, Stone R, et al. Impact of conditioning regimen intensity on outcome of allogeneic hematopoietic cell transplantation for advanced acute myelogenous leukemia and myelodysplastic syndrome. *Biol Blood Marrow Transplant.* 2006;12:1047–55.
39. Ringdén O, Labopin M, Ehninger G, Niederwieser D, Olsson R, Basara N, et al. Reduced intensity conditioning compared with myeloablative conditioning using unrelated donor transplants in patients with acute myeloid leukemia. *J Clin Oncol.* 2009;27:4570–7.
40. Luger SM, Ringdén O, Zhang MJ, Pérez WS, Bishop MR, Bornhauser M, et al. Similar outcomes using myeloablative vs reduced-intensity allogeneic transplant preparative regimens for AML or MDS. *Bone Marrow Transplant.* 2012;47:203–11.
41. Scott BL, Pasquini MC, Logan BR, Wu J, Devine SM, Porter DL, et al. Myeloablative versus reduced-intensity hematopoietic cell transplantation for acute myeloid leukemia and myelodysplastic syndromes. *J Clin Oncol.* 2017;35:1154–61.
42. Hourigan CS, Dillon LW, Gui G, Logan BR, Fei M, Ghannam J, et al. Impact of conditioning intensity of allogeneic transplantation for acute myeloid leukemia with genomic evidence of residual disease. *J Clin Oncol.* 2020;38:1273–83.
43. Craddock C, Jackson A, Loke J, Siddique S, Hodgkinson A, Mason J, et al. Augmented reduced-intensity regimen does not improve postallogeneic transplant outcomes in acute myeloid leukemia. *J Clin Oncol.* 2021;39:768–78.
44. Diverio D, Rossi V, Avvisati G, De Santis S, Pistilli A, Pane F, et al. Early detection of relapse by prospective reverse

- transcriptase-polymerase chain reaction analysis of the PML/RAR $\alpha$  fusion gene in patients with acute promyelocytic leukemia enrolled in the GIMEMA-AIEOP multicenter "AIDA" trial. GIMEMA-AIEOP Multicenter "AIDA" Trial. *Blood*. 1998;92:784–9.
45. Rücker FG, Agrawal M, Corbacioglu A, Weber D, Kapp-Schwoerer S, Gaidzik VI, et al. Measurable residual disease monitoring in acute myeloid leukemia with t(8;21)(q22;q22.1): results from the AML Study Group. *Blood*. 2019;134:1608–18.
  46. Ommen HB, Nyvold CG, Brændstrup K, Andersen BL, Ommen IB, Hasle H, et al. Relapse prediction in acute myeloid leukaemia patients in complete remission using WT1 as a molecular marker: development of a mathematical model to predict time from molecular to clinical relapse and define optimal sampling intervals. *Br J Haematol*. 2008;141:782–91.
  47. Ommen HB, Schnittger S, Jovanovic JV, Ommen IB, Hasle H, Østergaard M, et al. Strikingly different molecular relapse kinetics in NPM1c, PML-RARA, RUNX1-RUNX1T1, and CBFβ-MYH11 acute myeloid leukemias. *Blood*. 2010;115:198–205.
  48. Hokland P, Ommen HB. Towards individualized follow-up in adult acute myeloid leukemia in remission. *Blood*. 2011;117:2577–84.
  49. Ommen HB, Touzart A, MacIntyre E, Kern W, Haferlach T, Haferlach C, et al. The kinetics of relapse in DEK-NUP214-positive acute myeloid leukemia patients. *Eur J Haematol*. 2015;95:436–41.
  50. Puckrin R, Atenafu EG, Claudio JO, Chan S, Gupta V, Maze D, et al. Measurable residual disease monitoring provides insufficient lead-time to prevent morphologic relapse in the majority of patients with core-binding factor acute myeloid leukemia. *Haematologica*. 2021;106:56–63.
  51. Lo Coco F, Diverio D, Avvisati G, Petti MC, Meloni G, Pogliani EM, et al. Therapy of molecular relapse in acute promyelocytic leukemia. *Blood*. 1999;94:2225–9.
  52. Esteve J, Escoda L, Martín G, Rubio V, Díaz-Mediavilla J, González M, et al. Outcome of patients with acute promyelocytic leukemia failing to front-line treatment with all-trans retinoic acid and anthracycline-based chemotherapy (PETHEMA protocols LPA96 and LPA99): benefit of an early intervention. *Leukemia*. 2007;21:446–52.
  53. Grimwade D, Jovanovic JV, Hills RK, Nugent EA, Patel Y, Flora R, et al. Prospective minimal residual disease monitoring to predict relapse of acute promyelocytic leukemia and to direct pre-emptive arsenic trioxide therapy. *J Clin Oncol*. 2009;27:3650–8.
  54. Platzbecker U, Middeke JM, Sockel K, Herbst R, Wolf D, Baldus CD, 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;19:1668–79.
  55. Tiong IS, Dillon R, Ivey A, Teh TC, Nguyen P, Cummings N, et al. Venetoclax induces rapid elimination of NPM1 mutant measurable residual disease in combination with low-intensity chemotherapy in acute myeloid leukaemia. *Br J Haematol*. 2021;192:1026–30.
  56. Bataller A, Oñate G, Diaz-Beyá M, Guijarro F, Garrido A, Vives S, et al. Acute myeloid leukemia with NPM1 mutation and favorable European LeukemiaNet category: outcome after pre-emptive intervention based on measurable residual disease. *Br J Haematol*. 2020;191:52–61.
  57. Gianfaldoni G, Mannelli F, Interemesoli T, Bencini S, Giupponi D, Farina G, et al. Early peripheral clearance of leukemia-associated immunophenotypes in AML: centralized analysis of a randomized trial. *Blood Adv*. 2020;4:301–11.
  58. Wong HY, Sung AD, Lindblad KE, Sheela S, Roloff GW, Rizzieri D, et al. Molecular measurable residual disease testing of blood during AML cytotoxic therapy for early prediction of clinical response. *Front Oncol*. 2018;8:669.
  59. U.S. Department of Health and Human Services, Food and Drug Administration. Guidance for industry: expedited programs for serious conditions—drugs and biologics. 2014. <https://www.fda.gov/media/86377/download>. Accessed 7 Oct 2020.
  60. Prebet T, Bertoli S, Delaunay J, Pigneux A, Delabesse E, Mozziconacci MJ, et al. Anthracycline dose intensification improves molecular response and outcome of patients treated for core binding factor acute myeloid leukemia. *Haematologica*. 2014;99:e185–7.
  61. Lambert J, Lambert J, Nibourel O, Pautas C, Hayette S, Cayuela JM, et al. MRD assessed by WT1 and NPM1 transcript levels identifies distinct outcomes in AML patients and is influenced by gemtuzumab ozogamicin. *Oncotarget*. 2014;5:6280–8.
  62. Löwenberg B, Pabst T, Maertens J, van Norden Y, Biemond BJ, Schouten HC, et al. Therapeutic value of clofarabine in younger and middle-aged (18–65 years) adults with newly diagnosed AML. *Blood*. 2017;129:1636–45.
  63. Kapp-Schwoerer S, Weber D, Corbacioglu A, Gaidzik VI, Paschka P, Krönke J, et al. Impact of gemtuzumab ozogamicin on MRD and relapse risk in patients with NPM1-mutated AML: results from the AMLSG 09-09 trial. *Blood*. 2020;136:3041–50.
  64. Wei AH, Döhner H, Pocock C, Montesinos P, Afanasyev B, Dombret H, et al. Oral azacitidine maintenance therapy for acute myeloid leukemia in first remission. *N Engl J Med*. 2020;383:2526–37.
  65. Dimier N, Delmar P, Ward C, Morariu-Zamfir R, Fingerle-Rowson G, Bahlo J, et al. A model for predicting effect of treatment on progression-free survival using MRD as a surrogate end point in CLL. *Blood*. 2018;131:955–62.
  66. Avet-Loiseau H, Ludwig H, Landgren O, Paiva B, Morris C, Yang H, et al. Minimal residual disease status as a surrogate endpoint for progression-free survival in newly diagnosed multiple myeloma studies: a meta-analysis. *Clin Lymphoma Myeloma Leuk*. 2020;20:e30–7.
  67. Berry DA, Zhou S, Higley H, Mukundan L, Fu S, Reaman GH, et al. Association of minimal residual disease with clinical outcome in pediatric and adult acute lymphoblastic leukemia: a meta-analysis. *JAMA Oncol*. 2017;3:e170580.
  68. Galimberti S, Devidas M, Lucenti A, Cazzaniga G, Mörücke A, Bartram CR, et al. Validation of minimal residual disease as surrogate endpoint for event-free survival in childhood acute lymphoblastic leukemia. *JNCI Cancer Spectr*. 2018;2:pk0069.