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
Food and Drug Administration
Oncology Center of Excellence (OCE)
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

January 2020
Clinical/Medical

Hematologic Malignancies: Regulatory Considerations for Use of Minimal Residual Disease in Development of Drug and Biological Products for Treatment

Guidance for Industry

Additional copies are available from:

Office of Communications, Division of Drug Information
Center for Drug Evaluation and Research
Food and Drug Administration
10001 New Hampshire Ave., Hillandale Bldg., 4th Floor
Silver Spring, MD 20993-0002
Phone: 855-543-3784 or 301-796-3400; Fax: 301-431-6353; Email: druginfo@fda.hhs.gov
https://www.fda.gov/drugs/guidance-compliance-regulatory-information/guidances-drugs
and/or
Office of Communication, Outreach, and Development
Center for Biologics Evaluation and Research
Food and Drug Administration
10903 New Hampshire Ave., Bldg. 71, Room 3128
Silver Spring, MD 20993-0002
Phone: 800-835-4709 or 240-402-8010; Email: ocod@fda.hhs.gov
https://www.fda.gov/vaccines-blood-biologics/guidance-compliance-regulatory-information-biologics/biologics-guidances

U.S. Department of Health and Human Services
Food and Drug Administration
Oncology Center of Excellence (OCE)
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

January 2020
Clinical/Medical

TABLE OF CONTENTS

Hematologic Malignancies: Regulatory Considerations for Use of Minimal Residual Disease in Development of Drug and Biological Products for Treatment
Guidance for Industry[1]

I. INTRODUCTION

This guidance is intended to help sponsors planning to use minimal residual disease (MRD) as a biomarker in clinical trials conducted under an investigational new drug application (IND) or to support marketing approval of drugs and biological products[2] for treating specific hematologic malignancies.

The use of MRD as a biomarker in drug development is distinct from FDA's requirement for investigating, clearing, or approving an in vitro diagnostic device for clinical use in measuring MRD. Manufacturers interested in developing a specific MRD assay for clinical use should consult the Office of In Vitro Diagnostics and Radiological Health in the Center for Devices and Radiological Health (CDRH).

In general, FDA's guidance documents do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidances means that something is suggested or recommended, but not required.

[1] This guidance has been prepared by the Oncology Center of Excellence in cooperation with the Center for Drug Evaluation and Research and the Center for Biologics Evaluation and Research at the Food and Drug Administration.

[2] For the purposes of this guidance, all references to *drug products* include both human drugs and biological products.

II. BACKGROUND

Despite development of treatments that eliminate morphologically detectable malignant cells, some patients with hematologic malignancies who have achieved complete remission or complete response (CR), even of considerable duration, will experience relapse. Conventional morphologic detection for hematologic malignancies has a threshold limit of one tumor cell in 100 cells. Technology exists that can detect the persistence of malignancy at orders of magnitude below the limit of conventional morphologic detection, a level of disease burden known as MRD. These technologies measure cell characteristics such as genetic mutations, cell surface markers, or specific DNA gene rearrangements.

MRD as a general measure of tumor burden has multiple potential regulatory and clinical uses as a biomarker. Depending upon the clinical setting, MRD may be used to reflect a patient's response to treatment or as a prognostic tool to assess a patient's risk of future relapse. As such, MRD can be used to enrich clinical trial populations or guide allocation into specific treatment arms in clinical trials. There are challenges within each context of use that need to be addressed, such as the underlying disease, patient heterogeneity, therapeutic context, target of therapy, or a combination of disease parameters, to allow effective use of MRD in regulatory decision-making.

MRD assessments can vary among laboratories and technologies, which can result in discrepant results. Many clinical laboratories develop their own protocols that can affect MRD measurements. Technologies can have different performance characteristics. Sample collection procedures can also differ. However, standardized methodologies can ensure that results obtained between technologies and laboratories are consistent. This includes standardized posttreatment timing for when a bone marrow (BM) or blood sample is collected, standardized sample processing, predetermined MRD thresholds, and accurate reporting of the performance characteristics of the test (e.g., accuracy, precision, specificity, sensitivity). For example, reporting MRD negative results without information regarding limit of detection is not meaningful.

The evidence to support the clinical validity of MRD as a biomarker varies across hematologic cancer types and patient populations. To gain a better understanding of the state of the science of MRD, FDA cosponsored public workshops on MRD in acute lymphoblastic leukemia (ALL), chronic lymphocytic leukemia (CLL), and acute myeloid leukemia (AML) as well as a symposium on MRD in multiple myeloma (MM) from 2012 through 2014. In addition, a public workshop, Minimal Residual Disease as a Surrogate Endpoint in Hematologic Cancer Trials,[3] was held on September 7, 2016, under a cooperative agreement with FDA to discuss the clinical, statistical, and technical barriers to implementing use of MRD in clinical trials. As a result of these workshops and an analysis[4] of marketing applications showing inconsistent quality of

[3] See https://healthpolicy.duke.edu/events/minimal-residual-disease-surrogate-endpoint-hematologic-cancer-trials.

[4] Gormley N, V Bhatnagar, LA Ehrlich, B Kanapuru, H-Z Lee, AE McKee, A Farrell, and R Pazdur., 2017, FDA Analysis of MRD Data in Hematologic Malignancy Applications, J Clin Oncol, 35:2541.

MRD data, FDA identified a need to provide sponsors with guidance on use of MRD as a biomarker in regulatory submissions.

III. DEVELOPMENT OF MRD AS A BIOMARKER FOR REGULATORY USE

A. Regulatory Uses of Biomarkers

The term *biomarker* is commonly understood as referring to a characteristic that is measured as an indicator of normal biologic processes, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions.[5] MRD can be used as a biomarker. The terminology listed below is derived from the BEST Resource[6] definitions and the draft guidance for industry and FDA staff *Qualification Process for Drug Development Tools* (December 2019)[7] but has been slightly modified to reflect applicability to MRD. Sponsors can potentially use MRD status as any of the following types of biomarkers:

- **Diagnostic biomarker:** A biomarker used to detect or confirm presence of a disease or condition of interest or to identify individuals with a subtype of the disease.

- **Prognostic biomarker:** A biomarker used to identify likelihood of a clinical event, disease recurrence or progression in patients who have the disease or medical condition of interest. A prognostic biomarker informs about the natural history of the disease in that particular patient in the absence of a therapeutic intervention.

- **Predictive biomarker:** A biomarker 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 drug product.

- **Efficacy-response biomarker:** A biomarker that is used to show that a response has occurred in an individual who has been exposed to a drug product.

- **Monitoring biomarker:** A biomarker measured serially and used to detect a change in degree or extent of the disease.

[5] See FDA-NIH Biomarker Working Group, 2018, BEST (Biomarkers, EndpointS, and other Tools) Resource, accessed September 9, 2019, https://www.ncbi.nlm.nih.gov/books/NBK338448/. See also Section 507 of the Federal Food, Drug, and Cosmetic Act, which defines biomarker for purposes of that section, in relevant part, as "a characteristic (such as a physiologic, pathologic, or anatomic characteristic or measurement) that is objectively measured and evaluated as an indicator of normal biologic processes, pathologic processes, or biological responses to a therapeutic intervention."

[6] FDA-NIH Biomarker Working Group, 2018, BEST (Biomarkers, Endpoints, and other Tools) Resource.

[7] When final, this guidance will represent the FDA's current thinking on this topic. For the most recent version of a guidance, check the FDA guidance web page at https://www.fda.gov/regulatory-information/search-fda-guidance-documents.

An efficacy-response biomarker could be a surrogate endpoint. A surrogate endpoint does not measure the clinical benefit of primary interest; instead, it predicts the clinical benefit based on epidemiologic, therapeutic, pathophysiologic, or other scientific evidence. Specifically, a surrogate endpoint predicts a specific clinical outcome of the patient at some later time and can be used as the basis of marketing application approval decisions.

Understanding which of these biomarker attributes applies to the proposed use of MRD is important to consider when validating MRD for that proposed use and for the trial design. There are challenges within each MRD context of use that should be adequately justified, such as underlying disease, patient heterogeneity, therapeutic context, target of therapy, or a combination of disease parameters.

B. Mechanisms for Novel Surrogate Endpoint Acceptance or Qualification

Two mechanisms exist to obtain the Agency's feedback on the use of a novel surrogate endpoint to support approval. One mechanism is through the formal drug development tool (DDT) qualification process, specifically the biomarker qualification process. The DDT qualification process is an initiative undertaken in response to FDA's Critical Path Initiative and updated under the 21st Century Cures Act, adding section 507 to the Federal Food, Drug, and Cosmetic Act. The purpose of the DDT qualification process is to qualify a DDT for a specific context of use, such that a sponsor and FDA can rely on the DDT to have a specific interpretation and application in drug development and regulatory review. FDA will make information about a DDT that has been formally qualified for a specific context of use publicly available to expedite drug development and review of regulatory applications. A qualified DDT can be included in submissions of INDs, new drug applications (NDAs), or biologics license applications (BLAs) without the need for FDA to reconsider and reconfirm the suitability of the DDT. Qualifying a biomarker requires robust scientific evidence, and there is a higher evidentiary standard if the biomarker is to be used as a surrogate endpoint.[8]

A second mechanism to obtain the Agency's feedback on the use of a novel surrogate endpoint to support approval is through discussions with the specific Center for Drug Evaluation and Research (CDER) or Center for Biologics Evaluation and Research (CBER) review division. In this setting, the pharmaceutical sponsor or interested group meets with the FDA review division to present scientific data in support of the proposed surrogate endpoint. These data may be from previous clinical trials conducted by the sponsor, a meta-analysis of several trials conducted in the disease area, or other data, including product-nonspecific data, that support the use of the proposed surrogate endpoint. An example of this mechanism for a surrogate endpoint reasonably likely to predict clinical benefit is pathologic complete response in neoadjuvant treatment of breast cancer.[9] An example of a validated surrogate endpoint that used this mechanism is the

[8] For additional information on the DDT qualification process, see the DDT Qualification Programs web page at www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/default.htm and the draft guidance for industry and FDA staff *Qualification Process for Drug Development Tools*. When final, this guidance will represent the FDA's current thinking on this topic.

[9] See the guidance for industry *Pathological Complete Response in Neoadjuvant Treatment of High-Risk Early-Stage Breast Cancer: Use as an Endpoint to Support Accelerated Approval* (October 2014). We update guidances

surrogate of complete response at 30 months in follicular lymphoma. A surrogate endpoint that is reasonably likely to predict clinical benefit can be used to support accelerated approval, and a validated surrogate endpoint can support traditional approval.[10] To explore this approach further, sponsors should request a meeting with the relevant review division.

With either approach, the strength of evidence to support surrogacy depends on (1) biological plausibility of the relationship, (2) demonstration in epidemiological studies of the prognostic value of the surrogate endpoint for the clinical outcome, and (3) evidence from clinical trials that treatment effects on the surrogate endpoint correspond to effects on the clinical outcome.[11]

C. Meta-Analyses for Validation of MRD as a Surrogate Endpoint

Various statistical criteria have been proposed for validating a surrogate endpoint; often, meta-analytical approaches have been used. The terminology and definitions below provide further detail about statistical principles relevant to the validation of a surrogate endpoint.

- *Individual-level association* is the strength of the association between the surrogate and the true clinical endpoint.

- *Trial-level association* is the strength of the association between the effects of treatment on the surrogate and the true endpoint.

Although single-arm trial data may be used to demonstrate individual-level association and assess efficacy outcome of interest in subgroups by MRD level for the purposes of hypothesis generation, the meta-analysis to validate MRD at the trial level should include only randomized trials. The issues pertinent to meta-analyses in general have been discussed in another guidance.[12]

Sponsors should discuss with the Agency and provide details of the meta-analysis plan, which should include but not be limited to considering the following aspects:

- Details of trial designs, inclusion and exclusion criteria, MRD assessment, and disease setting. Sponsors should justify poolability of the data.

periodically. To make sure you have the most recent version of a guidance, check the FDA guidance web page at https://www.fda.gov/regulatory-information/search-fda-guidance-documents.

[10] For additional information on expedited programs and surrogate endpoints used to support accelerated or traditional approval, see the guidance for industry *Expedited Programs for Serious Conditions—Drugs and Biologics* (May 2014).

[11] See the ICH guidance for industry *E9 Statistical Principles for Clinical Trials* (September 1998).

[12] For additional information on meta-analyses, see the draft guidance for industry *Meta-Analyses of Randomized Controlled Clinical Trials to Evaluate the Safety of Human Drugs or Biological Products* (November 2018). When final, this guidance will represent FDA's current thinking on this topic.

- Inclusion of trials that include a patient population representative of the population in which the surrogate endpoint ultimately will be used.

- Inclusion of an adequate number of randomized trials with sufficient follow-up time. Sponsors should justify the number of trials to be included in the meta-analysis.

- Inclusion of trials that demonstrate both positive and negative results. For example, sponsors should present randomized trials that failed to meet their primary endpoint, and trials that had divergent MRD and event-free survival/progression-free survival/overall survival (EFS/PFS/OS) results, if available, should also be presented. Sponsors should explain the divergent results if possible.

- Analysis based on individual patient-level data to allow an assessment of individual-level surrogacy.

- Prespecified criteria for concluding surrogacy based on both trial-level and patient-level association measures, including prespecified timing and window of the MRD assessment. If a fixed time point is not feasible, the MRD assessments in a window of the trial should be prespecified. Sponsors should explore sensitivity analyses based on different time windows. Sponsors should discuss with the Agency the time window chosen. For example, sponsors can prespecify for patients with newly diagnosed ALL the MRD assessment at the time of the first complete response (CR1), 28 days plus or minus a window of a specific number of days.

- Inclusion of long-term clinical endpoints, such as EFS/PFS and OS that have been clearly and consistently defined across studies. Sponsors should explore alternative event definitions for EFS/PFS or alternative censoring schemes for EFS/PFS/OS as sensitivity analyses.

- Discussion of missing MRD assessments and reasons for missing data (e.g., caused by sample collection issues, loss to follow-up). Sponsors should explore the effects of missing data on the results.

- Consideration of the statistical handling of unevaluable samples.

- Potential confounding factors, which sponsors should incorporate into the planned validation analyses.

- Sensitivity analyses to demonstrate robustness of the surrogate (e.g., alternative statistical methods for evaluation of association,[13] cross validation) and subgroup analyses.

[13] Shi Q, CR Flowers, W Hiddemann, R Marcus, M Herold, A Hagenbeek, E Kimby, H Hochster, U Vitolo, BA Peterson, E Gyan, M Ghielmini, T Nielsen, S De Bedout, T Fu, N Valente, NH Fowler, E Hoster, M Ladetto, F Morschhauser, E Zucca, G Salles, and DJ Sargent., 2017, Thirty-Month Complete Response as a Surrogate End Point in First-Line Follicular Lymphoma Therapy: An Individual Patient-Level Analysis of Multiple Randomized Trials, J Clin Oncol, 35(5):552–560.

- Discussion and analyses using different MRD cutoffs (e.g., 10^{-4}, 10^{-5}). For assisting in the interpretation of the results, sponsors can present analyses such as surrogate threshold effect.[14]

Even if a meta-analysis supports validation of MRD as a surrogate endpoint, applying these results to a new trial requires a certain amount of extrapolation. Some caveats regarding the use of MRD as a surrogate endpoint include the following:

- Even if MRD can be validated as a surrogate endpoint, the use of MRD may not be applicable to subgroups of the patient population or future trial populations if there are important differences (e.g., prior therapy, disease status, line of treatment) between the population evaluated in the meta-analysis and the to-be-enrolled population. This may represent a different context of use, and as such, any differences should be justified. Sponsors should perform sensitivity and subgroup analyses to evaluate the strength of the surrogate endpoint in different disease settings or patient characteristics.

- When a new drug product is under investigation, it may not be reasonable to assume that the quantitative relationship between the drug product's effects on the surrogate endpoint and the clinical benefit endpoint will be the same as previously studied drug products' effects. This is especially true for drug products that have a markedly different mechanism of action (e.g., cytotoxic therapy versus immunotherapy). Although the credibility of this extrapolation will be primarily based on biological considerations, the meta-analyses mentioned above can provide supportive evidence. To obtain best estimates of the relationship between the surrogate and clinical benefit endpoints, the meta-analysis should include drug products with varying mechanisms of action and evaluate the relationship in mechanistic subtypes.

D. MRD as an Endpoint in Clinical Trials

MRD analyses should be based on the intent-to-treat (ITT) population. A patient may not have an MRD assessment because of a missed assessment, test failure, or not meeting clinical criteria for assessment (i.e., lack of CR). For ITT-based analyses, sponsors should consider any patient without an MRD assessment as not responsive to treatment. Analyses based on the MRD evaluable population are appropriate for sensitivity analyses.

Missing and unevaluable assessments of MRD should be kept to a minimum. Sponsors should collect and summarize reasons for missing MRD assessments and consult with the Agency before finalizing the statistical analysis plan. Sponsors also should perform further exploratory or sensitivity analyses to evaluate comparability of the results using different evaluation populations.

[14] Burzykowski T and Buyse M, 2006, Surrogate Threshold Effect: An Alternative Measure for Meta-Analytic Surrogate Endpoint Validation, Pharm Stat, 5(3):173–186.

Sponsors can also include MRD in the clinical trial as a secondary or an exploratory endpoint. If MRD-negative response (e.g., MRD-negative CR) is used as a secondary endpoint and is planned for inclusion in the prescribing information, it should be included as a key secondary endpoint with appropriate control for multiplicity.[15]

E. MRD for Patient Selection or Enrichment

Many clinical risk classifications may not be able to accurately predict relapse in patients with hematologic malignancies, which may result in inappropriate use or timing of treatments. MRD has been regarded as an important prognostic factor for predicting disease recurrence, which may improve risk classification. Sponsors can use MRD level to serve as a stratification factor, select patients at high risk, or enrich the trial population.[16]

FDA recommends that sponsors consult the Agency about incorporating any MRD assay into a trial before submitting the protocol for trials that include MRD for patient selection or as an endpoint (primary or key secondary).

IV. TECHNOLOGY CONSIDERATIONS

A. Assay Considerations

Currently, four general technologies are used for MRD assessment in hematologic malignancies: multiparametric flow cytometry (MPFC), next-generation sequencing (NGS), quantitative reverse transcription polymerase chain reaction (RT-qPCR) of specific gene fusions, and allele-specific oligonucleotide polymerase chain reaction (ASO-PCR). These cellular (MPFC) and molecular (NGS, RT-qPCR, and ASO-PCR) technology platforms have different advantages and limitations in terms of sample input, cost, robustness, and reproducibility.

FDA is agnostic as to which technology platform is used in clinical trials assessing MRD. However, sponsors should fully prespecify the selected platform (in terms of assay procedure, reagents, and analysis) and analytically validate the platform for its context of use. Also, in the context of a clinical trial, ideally sponsors should use a single technology to assess MRD to compare results directly. Although use of multiple technologies is discouraged, if the sponsor includes multiple technologies in the trial and plans for the primary analysis to be based on data from multiple technologies, the sponsor should prespecify the methodology for combining these technologies into a single MRD determination and discuss this with the Agency.

Analytical validation ensures that the assay measures the analyte or analytes that it is intended to measure in the intended tissue type. The process of analytical validation is defined as establishing that the performance characteristics of the assay are acceptable in terms of the

[15] See the draft guidance for industry *Multiple Endpoints in Clinical Trials* (January 2017). When final, this guidance will represent the FDA's current thinking on this topic.

[16] See the guidance for industry *Enrichment Strategies for Clinical Trials to Support Determination of Effectiveness of Human Drugs and Biological Products* (March 2019).

assay's sensitivity, specificity, accuracy, precision, and other relevant performance characteristics using a specified technical protocol, which may include specimen collection, handling, and storage procedures. Analytical validation is concerned with the assay's technical performance and does not address clinical utility.

MRD assay validation should encompass the entire assay system from sample collection (e.g., BM aspirate versus blood) to system output (e.g., decision-making threshold for MRD positive versus negative) and should use relevant clinical samples. Where technically feasible, the detection threshold of the MRD assay should be at least 10-fold below the clinical decision-making threshold (the definition of MRD). For example, if MRD positive or negative is defined as detection of greater or less than 1×10^{-5} cells, respectively, then the assay should be optimized and validated to have an analytical sensitivity of at least 1×10^{-6}. If this level of detection is not feasible with the proposed assay, sponsors should provide appropriate justification that the assay is adequate to fulfill its intent in the trial. Additionally, to ensure that the assay performance achieved in validation testing is replicated in the clinical trial, sponsors should strictly adhere to the assay protocol in all clinical trial laboratory sites. The following sections detail specific considerations for the different technology platforms.

1. Cellular Technology Platforms

When using cellular technology platforms for MRD assessments in clinical trials, sponsors should do the following:

- Prespecify the total number of events to be collected to support the quantitative assessment of MRD
- Use a consistent panel of antibodies and fluorochromes, as no single antigen is specific for any neoplasm
- Consider sample stability, which may limit the utility of flow cytometry
- Use a consistent analysis template (e.g., gating strategy)
- Determine whether the therapy affects the expression and therefore detectability of the specific antigens targeted by the antibody panels of the flow cytometry assay
- Evaluate the potential for the flow assay to detect normal BM cells that are regenerating after chemotherapy to reduce the likelihood that those cells are misinterpreted as abnormal cells

2. Molecular Technology Platforms

When using molecular technology platforms for MRD assessments in clinical trials, sponsors should do the following:

- Prespecify nucleic acid quantity (e.g., micrograms) and quality metrics

- Consider the need for an internal control when a cell number is derived from DNA content calculations because poor DNA quality may cause artificially low MRD levels

- Store diagnostic samples to be used for clone identification in case of assay changes

- Consider how to account for shifts in clonality as assessed by molecular markers (i.e., the specific molecular marker may be lost as a result of treatment while the disease remains present)

- Track assay failures (i.e., failures of the assay to identify the relevant clone for a patient) and consider this failure rate for clinical endpoint calculations

3. All Technology Platforms

When using any technology platform for MRD assessments in clinical trials, sponsors should do the following:

- Prespecify preanalytical procedures and ensure that the sample collection and storage procedures used are appropriate to obtain the desired cell population

- Take hemodilution into account (specifically, the amount of blood needed for the procedure to obtain the required number of events or amount of nucleic acid) and request that investigators use the first BM pull for MRD assessments

- For all testing, especially if centralized testing is not used, assay protocols and result interpretation should be standardized to ensure MRD measurements are comparable between laboratories

B. Sampling Considerations

Target levels of MRD for use in a regulatory setting are disease-specific and dependent upon the proposed use of the biomarker. In a clinical trial, the protocol should prespecify the measurement of MRD, which sponsors should conduct at prespecified times using a consistent and validated assay. The MRD assessment at a prespecified postinduction therapy time point is anticipated to be a sensitive measure of CR to induction therapy in either a frontline or relapsed/refractory setting. Consistent time-point specification would provide an opportunity to assess the kinetics of an MRD response and its duration, which may provide supportive evidence of drug activity. The timing of MRD assessment also is important when considering using MRD before allogeneic hematopoietic cell transplantation to predict transplant outcomes.

V. DISEASE-SPECIFIC CONSIDERATIONS

A. Acute Lymphoblastic Leukemia

MRD has emerged as one of the most significant prognostic factors in ALL, independent of patient age, B- or T-cell origin, or genetic subtype. Additional considerations for using MRD in ALL treatment trials include the following:

- BM is the preferred substrate for measuring MRD. If blood samples are used for assessing MRD in the clinical trial, sponsors should include justification for using blood rather than BM.

- CR with recovery of blood counts is the preferred time point to assess MRD. For regimens for which the efficacy-response evaluation is based on a calendar-driven time point rather than waiting for blood count recovery, at least an M1 marrow (marrow with leukemic blasts less than 5%) should be documented for patients being assessed for MRD.

- When MRD is used as an efficacy endpoint for ALL, the absence of extramedullary disease should be documented concurrently with assessment of BM and blood counts. However, FDA does not expect the conduct of invasive procedures to test for extramedullary disease if the procedures are not within the clinical standard of care at the time of the efficacy evaluation.

- FDA has accepted an MRD level of 0.1% or more to define patients with ALL in first or second CR with high risk of relapse. For trials that use MRD levels of less than 0.1% with CR for patient selection, the submission should include information to justify use of the lower MRD level.

- For new drugs that have a demonstrated durable CR in patients with relapsed or refractory ALL, FDA has accepted MRD of less than 0.01% as supporting evidence of efficacy. As technologies improve and new clinical findings emerge, the level of MRD needed to support an efficacy claim may change.

B. Acute Myeloid Leukemia

The molecular heterogeneity of AML poses substantial challenges to the use of MRD as a biomarker. Additional considerations for use of MRD in AML treatment trials include the following:

- BM is the preferred substrate for measuring MRD. If blood samples are used for response assessment of MRD in the clinical trial, sponsors should include justification for using blood rather than BM.

- CR with recovery of blood counts is the preferred time point to assess MRD. If assessments are made at CR without count recovery or at lesser responses (e.g., complete

remission with incomplete hematologic recovery), sponsors should include data to justify the plan.

- For the marker (e.g., cell surface or genetic mutation) selected to assess MRD, sponsors should provide data showing that the marker reflects the leukemia and not underlying clonal hematopoiesis (false-positive result). Sponsors should also describe the false-negative rate that might result from relapse from a marker-negative clone. If multiple markers and/or multiple platforms are used, sponsors should provide an analysis of the risk of false-positive and false-negative results for each marker individually and for the panel as a whole.

- For studies of targeted therapies (e.g., IDH1, IDH2, or FLT3 inhibitors) for which the MRD marker is the target of the therapy, sponsors can use nonclinical data to identify the mutations in the marker that are known to be sensitive to the therapy and those that are known to be resistant to the therapy. If using only the target of therapy as the MRD marker, sponsors should provide justification for not using other MRD markers to avoid false-negative results when clonality changes.

C. Acute Promyelocytic Leukemia

The standard-of-care use of MRD testing and monitoring is established for the initial treatment of patients with acute promyelocytic leukemia (APL) using tretinoin with arsenic and/or anthracycline. Whether the same guidelines for use of MRD apply to other drug classes needs to be confirmed as new drugs are evaluated for initial or salvage therapy. Additional specific considerations include the following:

- BM is the preferred substrate for measuring MRD. If blood samples are used for response assessment of MRD in the clinical trial, sponsors should include justification for using blood rather than BM.

- CR following recovery of blood counts is the preferred time point to assess MRD. If assessments are made at CR without count recovery or at lesser responses, sponsors should include data to justify the plan.

- MRD should be assessed at the end of consolidation rather than at the end of induction, when differentiating agents are used, to avoid false-positive results. For new drug products for treatment of APL, sponsors should use data from early-phase trials to establish the optimal timing for MRD assessment in the pivotal trials.

- Patients with low-risk APL who achieve confirmed MRD negativity after arsenic/tretinoin-based therapy are generally considered cured and require no further monitoring. For new drug products for treatment of APL, long-term monitoring may be required in the pivotal trial if data from early-phase trials are not sufficient to confirm that MRD negativity is also durable with the new drug product.

- An MRD level less than 0.01% is generally considered negative after first-line arsenic/tretinoin- or idarubicin/tretinoin-based induction. For new drug products for treatment of APL, sponsors should use data from early-phase trials to confirm this threshold for defining MRD negativity for the new drug product.

- Although an MRD level less than 0.01% is generally considered negative after first-line treatment, marketing applications for treatment of molecular relapse may need clinical outcomes (i.e., EFS) if data are not available to support a proposed MRD threshold as the sole criterion for response to salvage therapy.

D. Chronic Lymphocytic Leukemia

Current literature suggests that there is an association between MRD negativity and OS in patients with CLL treated with chemoimmunotherapy. The therapeutic paradigm with small molecule inhibitors of the B-cell receptor signaling pathway and other novel products continue to rapidly evolve in this area. Additional specific considerations include the following:

- MRD status should be measured by a standardized method with a quantitative lower limit of detection sufficient to evaluate the prospective cutoff in the trial and at least less than 10^{-4} (0.01%). Currently, MRD is most commonly assessed using RT-qPCR and flow cytometric methods, but NGS can also reliably assess MRD in CLL.

- A challenge in MRD testing is that CLL is a multicompartmental disease involving the BM, blood, lymph nodes, liver, and spleen; after treatment, one or more of these sites may serve as a reservoir for residual disease. Sponsors should carefully consider for assessment the sample source, which should be the same throughout the trial. This is especially important as therapeutic intervention differentially affects MRD measurement in peripheral blood and BM, as has been demonstrated with certain therapeutics (e.g., anti-CD20 monoclonal antibodies, alemtuzumab).

- The timing of when to test for MRD has yet to be standardized and the time to response and response durations may vary by type of therapeutic regimen. Sponsors should prespecify the timing and method of MRD testing and provide adequate justification in the protocol. MRD should also be measured at the end-of-treatment response assessment to fully capture the treatment effect.

- MRD should be assessed in patients who are in CR. If MRD assessments are to be made in patients in other response categories (e.g., partial response), sponsors should include data to justify the plan.

E. Chronic Myeloid Leukemia

There have been dramatic improvements in clinical outcomes in patients with chronic myeloid leukemia (CML) from targeting the BCR-ABL1 oncoprotein. The detection and monitoring of MRD has become standard of care in patients with CML. Specific considerations include the following:

- Monitoring MRD in CML should use assays with results based on the International Scale (IS) with the standardized baseline set to 100%. Molecular response is expressed as log reduction from 100%.

- Currently, RT-qPCR(IS) is the preferred assay to monitor response to therapy. In general, RT-qPCR assays with a sensitivity of more than 4.5-log reduction from the standardized baseline are recommended for measuring BCR-ABL1 transcripts.

- Major molecular response (MMR) is defined as BCR-ABL(IS) of less than 0.1% or more than 3-log reduction in BCR-ABL1 mRNA from the standardized baseline if RT-qPCR(IS) is not available.

- There is evidence that achieving an MMR predicts superior long-term clinical outcomes (PFS/EFS).

- Achieving MMR has become a consensus goal of CML therapy, and durable MMR can be a measure of clinical benefit.

- In addition, MRD can be used to select and monitor patients who are eligible for treatment discontinuation of tyrosine kinase inhibitor therapy.

F. Multiple Myeloma

Significant improvements in clinical outcomes of MM have spurred interest in the use of MRD as a potential surrogate endpoint to expedite drug development. Multiple trials have evaluated the relationship between MRD status and PFS/OS.

Additional specific considerations for use of MRD in trials of new drug products for treatment of MM include the following:

- Most published literature to date has evaluated MRD in the newly diagnosed posttransplant setting. Fewer studies have evaluated MRD in the setting of relapsed/refractory disease or newly diagnosed patients with myeloma who are not eligible for transplant. The relationship between MRD and clinical benefit and the test performance characteristics should be demonstrated in each disease setting (e.g., relapsed refractory, newly diagnosed, nontransplant eligible, smoldering MM) to validate MRD as a surrogate endpoint in MM. This is especially important in disease settings, such as smoldering myeloma, in which there is a lower disease burden and the potential for toxicity or other nondisease-related factors influence long-term outcomes.

- MRD should be assessed only in patients who are in CR. If MRD assessments are to be made in patients in other response categories (e.g., partial response, very good partial response), sponsors should include data to justify the plan.

- MRD is currently assessed using MPFC and NGS methods in the BM. These methodologies are not able to detect extramedullary disease. There has been interest in using imaging techniques (e.g., positron emission tomography/computed tomography, magnetic resonance imaging) in combination with MRD to assess response. When considering using MRD in MM clinical trials, sponsors should discuss with FDA how extramedullary disease will be assessed and whether imaging should be incorporated into the assessment of response.

- At this time, the relationship between MRD and clinical benefit in patients with different cytogenetic abnormalities and their associated risks is unclear. When considering using MRD in clinical trials, sponsors should consider the patient's cytogenetic risk. For example, given the prognostic effect of cytogenetics, the trial may benefit from stratification to ensure that there is no imbalance between the arms that would affect the MRD assessment. Alternatively, trials may be designed to intervene in patients who are MRD positive and have poor risk cytogenetics because this may represent a group at risk for particularly poor outcomes.

VI. REGULATORY SUBMISSIONS THAT USE MRD

As indicated above, FDA views MRD as a biomarker that is a reliable quantitation of tumor burden, independent of assay. As such, FDA does not foresee the need to codevelop an MRD assay with a drug product.[17] However, for FDA to adequately assess the safety of a proposed clinical trial that uses MRD (e.g., for patient selection) or to determine the credibility of a clinical trial outcome based in part on MRD, submissions that use MRD for regulatory purposes or for critical treatment purposes should include sufficient information to address the following two main issues:

- Is MRD, as assessed (sample, timing, threshold, etc.), a clinically valid biomarker for the context of use (disease, disease status, type of therapy, etc.)?

- Is the MRD assay used (or to be used) in the clinical trial analytically valid for the range of results that are important to the trial?

When the MRD assay used is FDA-cleared or -approved for the specific malignancy and specimen type, identifying the assay with the required number of cells to be evaluated or the DNA input requirements will be sufficient to address the analytical validity in most cases. When the MRD assay is not FDA-cleared or -approved, FDA would expect additional information, such as those listed in Table 1, to be submitted for review.

[17] A potential exception might be when the MRD marker is the direct target of the drug product under study, such as for selecting patients for treatment in a clinical trial of an Fms-related tyrosine kinase 3 (FLT3) inhibitor when the MRD assay is for an FLT3 mutation. In such a circumstance, sponsors should consult with FDA about the need for a companion diagnostic early in clinical development.

Table 1. Information to Help Review of Regulatory Submissions That Use MRD

IND Clinical Trial Submission	NDA or BLA Submission
1. Justification that MRD as used is clinically valid for the proposed context **and**	1. Justification that MRD as used is clinically valid for the context of the proposed claim **and**
2. Letter of authorization to cross-reference the investigational device exemption or other device-related regulatory submission for information about the assay **or** • A statement of intended use; • The specific test method (including instruments, reagents, and specimen handling); • Confirmation that the lab has a process in place for reagent control; • A brief discussion of how the test method was validated analytically for each specimen type; **and** • A summary of the performance obtained for accuracy, precision, specificity, and sensitivity; **and**	2. Letter of authorization to cross-reference the investigational device exemption or other device-related regulatory submission for information about the assay **or** • A statement of intended use; • The specific test method (including instruments, reagents, and specimen handling); • Confirmation that the lab has a process in place for reagent control; • A brief discussion of how the test method was validated analytically for each specimen type; **and** • A summary of the performance obtained for accuracy, precision, specificity, and sensitivity; **and**
3. Indicate in the clinical trial informed consent document that the MRD assay is investigational.	3. A SAS XPORT file (xpt file extension) with results of MRD testing. For each result, specify the sample type, date of sample, assay used, input quantity, assay sensitivity, and assay result.

For an IND clinical trial submission, when using an MRD assay that is not FDA-cleared or -approved for the intended use and the trial is considered a significant risk device trial (e.g., eligibility criterion, allocation to a specific treatment, departures from standard of care, etc.), FDA may require an investigational device exemption to use the assay in the clinical trial.[18] Sponsors can submit a letter of authorization to cross-reference the investigational device exemption, which will then provide the necessary information regarding the assay. When the trial is considered a nonsignificant risk device study, the sponsor should submit abbreviated information about the assay (see Table 1) to the IND for review to allow FDA to confirm that results from the device will be interpretable. An NDA or BLA submission should include similar information about the assay (see Table 1) in addition to a data file with the results of MRD testing.

Although general principles outlined in this guidance should help sponsors with crucial questions about potential MRD use for marketing applications, FDA recommends that sponsors meet with

[18] As an alternative, sponsors can consider the streamlined approach to codeveloping the MRD assay with the drug product under the IND as described in the draft guidance *Investigational In Vitro Diagnostics in Oncology Trials: Streamlined Submission Process for Study Risk Determination* (April 2018). When final, this guidance will represent the FDA's current thinking on this topic. See 21 CFR 812; for information on risk determination for investigational use of devices, see the guidance for industry and FDA staff *Requests for Feedback and Meetings for Medical Device Submissions: The Q-Submission Program* (May 2019).

FDA before starting a drug development pathway incorporating MRD assessment intended to support an NDA or a BLA. FDA will ensure that these meetings include a multidisciplinary team of review staff from CBER, CDER, and CDRH, as needed. Sponsors can submit protocols using MRD after these meetings and request a special protocol assessment for eligible protocols, if they choose, that provides confirmation of the acceptability of assessments, endpoints, and protocol design to support drug marketing applications. Ultimately, marketing approval depends not only on the design of clinical trials but also on FDA review of the results and data from all studies in the drug marketing application.

GLOSSARY

ALL	acute lymphoblastic leukemia
AML	acute myeloid leukemia
APL	acute promyelocytic leukemia
ASO-PCR	allele-specific oligonucleotide polymerase chain reaction
BLA	biologics license application
BM	bone marrow
CBER	Center for Biologics Evaluation and Research
CDER	Center for Drug Evaluation and Research
CLL	chronic lymphocytic leukemia
CML	chronic myeloid leukemia
CR	complete response or complete remission
DDT	drug development tool
EFS	event-free survival
IDE	investigational device exemption
IND	investigational new drug application
IS	International Scale
ITT	intent-to-treat
MM	multiple myeloma
MMR	major molecular response
MPFC	multiparametric flow cytometry
MRD	minimal residual disease
NDA	new drug application
NGS	next-generation sequencing
OS	overall survival
PFS	progression-free survival
RT-qPCR	quantitative reverse transcription polymerase chain reaction