SUMMARY OF DISCUSSION

Cancer Drug Development Forum (CDDF) Multi-Stakeholder Workshop on

Minimal Residual Disease in Acute Myeloid Leukaemia (AML) and Chronic Lymphocytic Leukaemia (CLL)

8-9 November 2018
London, UK

Prepared by Excerpta Medica

TABLE OF CONTENTS

PROGRAMME

Day 1

SESSION 1: INTRODUCTORY SESSION

Chairs: *Ralf Herold (EMA, UK) and John Smyth (CDDF/University of Edinburgh, UK)*

12:30	**Welcome and introduction to the topic**
	Axel Glasmacher (CDDF/University of Bonn, DE)
13:00	**Regulatory aspects – AML & CLL**
	Beatriz Flores (EMA/MHRA), UK

SESSION 2: WORKSHOP ON MRD IN AML

Chairs: *Konstanze Döhner (University Hospital of Ulm, DE) & Nicole Gormley (FDA, USA)*

13:30	**New developments in AML**
	Christoph Röllig (Technische Universität Dresden, DE)
14:00	**Clinical overview: MRD in AML**
	Konstanze Döhner (University Hospital of Ulm, DE)
14:45	**European LeukemiaNet guidelines**
	Arjan van de Loosdrecht (VU University Medical Center, NL)
15:25	**Coffee break**
16:00	**Methodological overview: MRD in AML**
	Chris Hourigan (National Institutes of Health, USA)
16:45	**Standardization of MRD measurement**
	Christian Thiede (Technische Universität Dresden, DE)
17:30	**Consortium on MRD in AML and industry perspective**
	Sharon McBain (Johnson & Johnson, USA)

Day 2

8:00	**Regulatory aspects: AML & CLL**
	Nicole Gormley (FDA, USA)

SESSION 3: PROGRESS ON MRD IN CLL

Chairs: *Robert Gale (Celgene, USA) & Andy Rawstron (Leeds Teaching Hospital NHS Trust, UK)*

8:30	**Clinical overview: MRD in CLL**
	Mathias Ritgen (University Medical Center Schleswig-Holstein, DE)

9:15	**Methodological overview**
	Andy Rawstron (Leeds Teaching Hospital NHS Trust, UK)

10:00	**Industry perspective**
	Davy Chiodin (Acerta, AstraZeneca, USA)

10:30	**Coffee break**

SESSION 4 (ROUNDTABLE): LESSONS LEARNED AND OPEN RESEARCH QUESTIONS FOR AML AND CLL, NEXT STEPS

Moderators: *John Smyth (CDDF/University of Edinburgh, UK) & Axel Glasmacher (CDDF/University of Bonn, DE)*

11:00	**Regulators** *(B. Flores, N. Gormley)*
	Academia *(K. Döhner, A. Rawstron)*
	Industry *(C. Pallaud, I. Radtke)*

12:30	**End of workshop**

BACKGROUND AND OBJECTIVES

Drug development have traditionally relied on standard endpoints for haematologic malignancies, e.g. response rate (RR), PFS, event-free survival (EFS), overall survival (OS). The advancement of novel therapies leading to improved clinical outcomes has created the need to introduce surrogate endpoints to accelerate regulatory approvals. With the availability of highly sensitive assays able to detect residual tumour at the end of treatment, MRD is emerging as a relevant and objective novel endpoint and clinical decision-making tool.

In a recent analysis, nearly 40% of applications submitted to the Food and Drug Administration (FDA)'s Division of Hematology Products between 2014 and 2016 included MRD data. The data submitted was deemed adequate for inclusion in the prescribing information (PI) in only 46% of cases; the remainder was excluded because it was considered inadequate (31%), or was not proposed for inclusion (23%) because of missing or disparate data, high amounts of test failure rate, incomplete test characteristics data, and incomplete planned statistical analysis.

Despite the recent approval of novel agents for acute myeloid leukaemia (AML) and chronic lymphocytic leukaemia (CLL), there is still a significant unmet need. New data and scientific and regulatory guidelines have moved the field forward and warrant an updated discussion. This workshop focused on status and next steps for the use of MRD in AML and CLL clinical trials as well as on open exchange, learning, and collaborative search for agreements from regulatory, academic, and industry perspectives.

REGULATORY ASPECTS

> *"Medicine is a science of uncertainty and an art of probability "*
> *William Osler*

The CHMP/EMA view

Beatriz Flores, Medicines and Healthcare Products Regulatory Agency (MHRA), UK

MRD as an endpoint in CLL (EMA/CHMP/703715/2012 Rev. 2)

A condition-specific guidance was published in 2015, 19 months after the initial Cancer Drug Development Forum (CDDF) meeting discussing the endpoint. The guideline postulates that undetectable MRD (defined as $<10^{-4}$ residual cells in PB, confirmed by BM assessment) in patients with CLL in clinical CR after induction therapy may be used as an intermediate endpoint for licensure in randomized well-controlled studies designed to show superiority in terms of PFS.

MRD as an endpoint in AML

AML is a complex disease characterized by multiple driver mutations, competing co-occurring clones, and disease evolution leading to a phenotype shift at progression. With only 3 drugs obtaining European Medicines Agency (EMA) regulatory approval in recent years, i.e. gemtuzumab ozogamicin (Mylotarg), daunorubicin and cytarabine (Vyxeos), and midostaurin (Rydapt), the Committee for Medicinal Products for Human Use (CHMP) has limited experience compared to other haematological malignancies. There are several uncertainties:

- What mutations are associated with relapse?

- What is the best timing for MRD testing?

- What is the appropriate threshold?

- Do multicentre clinical trials need a centralized laboratory?

- There are no validated tests

- Can results be extrapolated across risk groups within a subtype?

- Can results be extrapolated across different treatments (transplant, nonintensive treatment, etc.)?

- How to achieve statistical power and demonstrate an effect in AML molecular subtypes with low prevalence?

It is expected that new phase 3 trials are adopting current European LeukemiaNet (ELN) recommendations for assessment of MRD and the results are eagerly awaited.

Conclusions

MRD holds the promise to become an intermediate clinical trial endpoint; it may also serve as a stratification factor and a prognostic marker guiding treatment decisions. However, the considerable heterogeneity across haematologic malignancies requires that MRD is validated per disease. Moreover, if considerable intra-disease heterogeneity exists, MRD may need to be validated per disease subtype.

The FDA view

Nicole Gormley, Clinical Team Leader, Division of Hematology Products, FDA, USA

US regulatory framework

In the USA, there are two regulatory pathways for approval of new therapeutics: regular approval, and accelerated approval for therapies of serious or life-threatening illness. Both approvals can be based on a surrogate endpoint, either on an established surrogate with regular approval or one that is reasonably likely to predict clinical benefit in the case of accelerated approval. Accelerated approval may require post-approval trials to verify the anticipated clinical benefit.

Development of MRD as a biomarker for regulatory use

Subsequently, Dr Gormley discussed the draft industry guidance on the regulatory consideration for use of MRD in drug development for haematologic malignancies. The guidance is currently open for comments.

There are two ways by which surrogate endpoints can be qualified or accepted by the FDA:

1. FDA's Biomarker Qualification Program, also known as Drug Development Tool (DDT) qualification. DDT is qualified for a specific context of use. Higher evidentiary standard is required if a biomarker is to be used as a surrogate endpoint

2. Discussion with the Review Division, where the pharmaceutical company or group meet to discuss scientific data from previous clinical trials or meta-analysis.

Examples include the use of a pathologic CR in neoadjuvant breast cancer, or CR at 30 months in follicular lymphoma.

The guidance also includes meta-analysis considerations and technology considerations. It is important to note that the use of a surrogate may not be appropriate for subpopulations or future trials if patients/disease characteristics differ substantially from those studied. In addition, the use of a surrogate may not be appropriate for therapeutic modalities that have substantially different mechanisms of action (MOAs) (e.g. cytotoxic vs immunotherapies).

MRD in CLL

Undetectable MRD (LOD 10^{-4}) in patients with CLL is associated with prolonged PFS and OS. However, CLL is a multicompartmental disease and the multiple reservoirs of residual disease should be considered in addition to MRD assessments.

Recommendations are provided in the FDA draft guidance and are generally similar to the EMA guidance. MRD results have already been included in the PI of venetoclax and obinutuzumab. The PI for venetoclax was updated based on the MURANO trial and includes adequate information on MRD definition and timing of assessment. MRD was detected in PB in patients who achieved at least a partial response (PR). Information about the CR/CRi rate (where CRi = CR with incomplete haematologic recovery) was also included and justification was provided for inclusion of patients in PR.

MRD in AML

Recommendations are provided in the FDA draft guidance. CR with recovery of blood counts is the preferred time for measurement of MRD, with BM being the preferred substrate.

The molecular heterogeneity of AML poses substantial challenges to the use of MRD as a biomarker. The sponsor should provide data showing that the marker selected for MRD assessment reflects leukaemia, and not underlying clonal haematopoiesis. If multiple markers and/or platforms are used, an analysis of the risk of false-positives and false-negatives for each marker individually, and for the panel, should also be provided.

MRD results have not been included in PIs for AML therapies to date.

Conclusions

Dr Gormley concluded that the FDA is committed to working with the community on the development of MRD in haematologic malignancies and encourages assessment of MRD in clinical trials.

Discussion

- *Endpoints:* The objective of a trial is critical, as it determines the choice of endpoints and whether MRD is your intermediate or exploratory endpoint.

- *Technology considerations*: Quality assurance and information regarding the quality of the sample is critical for interpretation.

- *Validated tests*: In September 2018, the FDA authorized the first NGS-based test (ClonoSEQ® assay, Adaptive Biotechnologies) to detect MRD in patients with acute lymphoblastic leukaemia (ALL) or multiple myeloma (MM) in an attempt to ensure that these evolving tests are accurate and reliable. If there is no approved test, a discussion with the regulators is encouraged to ensure it is validated for its purposes.

- *Number of assays*: The FDA does not currently recommend the use of two assays unless a correlation between the assays is demonstrated. This is in contradiction with the EMA guidelines for MM recommending use of two assays. A concern was raised that if not reconciled, this may represent a challenge for drug development on a global scale.

- *Data-driven approach*: Science evolves, and the regulators are taking this into account. Scientific advice should be sought early in the process. The regulators are interested in innovation and drug development but would always take a data-driven approach.

MRD DEFINITION

- MRD indicates the persistence of leukaemic cells after treatment measured either by multiparameter flow cytometry (MFC) or a molecular technique in numbers that are well below the sensitivity detection level of routine morphology.

- **CDDF workshop consensus:** MRD should be referred to as <u>measurable</u> (instead of minimal) residual disease.

Hourigan CS, et al. Leukemia 2017;31:1482–90.

WORKSHOP ON MRD IN AML

New developments in AML

Christoph Röllig (Technische Universität Dresden, DE)

There is a considerable unmet need in AML, with only approximately 30% of patients surviving at 5 years, despite >50% of patients achieving CR. Old age and the presence of adverse cytogenetic lesions confers poor prognosis. Moreover, 50% of patients relapse from a morphologic R. Despite being an early endpoint and a prerequisite for cure, CR is not sensitive enough as a marker of long-term outcomes. OS remains the most relevant endpoint; however, this may slow down approvals and bringing new, much-needed therapies to patients as exemplified by the RATIFY trial. In addition, OS is likely to be confounded by post-relapse therapy, i.e. allotransplant or clinical trials. Several new drugs, including targeted therapies, have been approved by the FDA and/or EMA in 2017–2018 (midostaurin, gemtuzumab ozogamicin, ivosidenib, CPX-351, enasidenib, gilteritinib, venetoclax, glasdegib). In the RATIFY trial, the addition of midostaurin to conventional chemotherapy led to a significant OS advantage, although there was no difference between the arms in conventional CR rates, suggesting an effect at the MRD level. Other novel drugs, such as the anti-CD33 antibody, gemtuzumab, the second-generation FLT3 inhibitors (e.g. gilteritinib), bispecific antibodies (CD33/CD123), or therapies such as chimeric antigen receptor (CAR) T-cells are associated with high levels of CR and are also efficacious or expected to be efficacious in dimensions of MRD. Thus, MRD negativity may become an early clinical endpoint having the potential to provide quicker read-out of clinical trials and expedited drug development.

It was discussed that the association between MRD and outcome is likely to be dependent on the MOA of a drug. For instance, MRD may not account for immunologic effects or late drug effects, and other endpoints or markers may be relevant.

Clinical overview of MRD in AML

Konstanze Döhner (University Hospital of Ulm, DE)

Dr Döhner gave an overview of clinical trials in AML that have assessed MRD. Most of the studies were performed retrospectively, as post-hoc analyses. Data was driven by the

availability of BM or PB samples at defined time points. Most patients were in first CR. Patients were selected according to the presence of a molecular marker:

- **Core-binding factor (CBF) leukaemia/*RUNX1/RUNX1T1* (5 studies)**
- **NPM1 (5 studies)**

Overall, the studies were performed on heterogeneous patient populations with respect to age, treatment, cohort size, or type of material. MRD was monitored by real-time quantitative (RQ)-PCR or NGS, and the definition of MRD negativity was not consistent. However, in most studies, achievement of MRD negativity or significant reduction of transcript levels/mutations by RQ-PCR/NGS after 2 cycles of therapy and/or at the end of treatment was significantly associated with improved outcomes. This is illustrated by a retrospective analysis of 437 patients with ***NPM1*-mutated** standard-risk AML. In this study, PB assessment of MRD positivity (RQ-PCR; sensitivity 10^{-5}) after 2 treatment cycles was associated with a reduced risk of relapse (30% vs 82% for MRD negativity) and improved OS (24% vs 73%) at 3 years. Genetic complexity, e.g. concurrent *FLT3ITD/DNMT3A* mutations, had a negative impact on achieving MRD as well as long-term outcomes. However, patients with concurrent *FLT3ITD/DNMT3A* mutations who achieved MRD negativity in PB after 2 cycles of therapy had a significantly lower risk of relapse, which was independent of the *FLT3ITD/DNMT3A* mutations status.

NGS-based MRD monitoring has been shown to be useful in approximately 90% of AML patients. When combined with multiparameter flow cytometry (MFC), the 2 assays conferred an independent prognostic value with respect to RFS and OS. Therefore, Dr Döhner recommended that both assays should be validated further in clinical trials. As a conclusion, Dr Döhner highlighted the need to standardize MRD techniques (NGS, MFC) and establish guidelines for their consistent use in all clinical trials for AML.

During the discussion, Dr Döhner highlighted that a significant proportion of patients do not achieve MRD negativity in BM after 2 cycles of therapy; however, MRD responses deepen over time, beyond the early responses. Thus, *MRD kinetics allow monitoring of treatment effects*, and this is of interest in clinical studies evaluating novel drugs.

European LeukemiaNet guidelines: recommendations and techniques
Arjan van de Loosdrecht (VU University Medical Center, NL)

MRD is emerging as an important regulatory and prognostic tool; however, there are still challenges with standardization of the available techniques. The ELN guidelines aim to harmonize the approaches to MRD measurement and improve its overall quality. In brief, the guidelines summarize key areas of agreement in the measurement and practical application of MRD among AML experts, and provide guidelines for current and future use in clinical practice. **The guidelines postulate integrating molecular and/or MFC MRD into all clinical trials at all times of evaluation of response, using the technical ELN recommendations.** These include recommendations for:

- Discriminative marker panels
- Threshholds
- Time points for MRD assessments (e.g. after 2 treatment cycles, at the end of treatment; every 3 months for 24 months after the end of treatment)
- Definitions for molecular remission, molecular progression, and molecular relapse
- Technical requirements for specimen collection/processing (e.g. within 72 hours)
- Calculation of MRD burden

Next, Dr van de Loosdrecht discussed current approaches to MRD measurement in the context of the ELN guidelines.

MFC: "LAIP-based different-from-normal (DfN) approach"

Detection of MRD by MFC (≥8 colours) is based on detecting aberrant immunophenotypes present on the leukaemia cell. This is a multistep process, since AML patients all show different phenotypes characterized by antigen over- or underexpression, antigen cross-lineage expression, or antigen asynchronous expression. At diagnosis, a screening panel with a wide range of markers is used to define which markers are aberrantly expressed on the leukaemic blasts, i.e. *leukaemia-associated immunophenotype* (LAIP). Subsequently, the aberrantly expressed markers (e.g. CD45) are combined with known primitive, myeloid, and lymphoid markers (e.g. CD34, CD33, CD7). This aberrant combination (different-from-normal [DfN]) is often seen on AML progenitors but absent on normal progenitors. This phenomenon is determined by differences in differentiation patterns between the patient's BM and normal BM. However, the "LAIP-based DfN" approach can be reliably performed only in experienced laboratories and by well-trained operators. Although MFC is highly applicable (approximately 90% of cases), sensitivity may not be uniform across patients, and comments on the quality of the sample, absolute number of cells, etc. help interpretation. In addition, as not all leukaemic cells express an aberrant immunophenotype, a concomitant

assessment of leukaemia stem cell (LSC) load might improve our ability to predict long-term outcomes and risk of relapse after therapy, but data are still limited.

Molecular techniques/NGS

RQ-PCR is a highly sensitive and well standardized technique and therefore currently considered to be the gold standard. However, it is currently limited to approximately 40% of AML patients. A set of 5 targets, including mutant NPM1, RUNX1-RUNXT1, CBFB-MYH11, and PML-RARA, have been approved by the international consensus committee on AML MRD. Clonal haematopoiesis of indeterminate potential (CHIP) mutations such as DNMT3A, ASXL1, or TET2 (abbreviated as DAT mutations) are *not* usable as they do not confer prognostic value. For AML patients who do not express a suitable marker, MRD should be assessed by MFC.

NGS offers the opportunity for detection and follow-up of a large number of aberrations with 1 test. As virtually all AMLs harbour genetic mutations, NGS is a very attractive tool with potentially broad applicability. However, sensitivity is currently about 1%, which cannot compete with the other discussed MRD techniques. It is expected that this sensitivitty will improve and NGS is expected to be the technique of the future because of its high throughput. Future randomized studies should assess the prognostic impact of MRD as measured by NGS status on OS.

Conclusions

Dr van de Loosdrecht concluded that MRD is an important objective methodology to establish remission status. As demonstrated by the HOVON trials (**HOVON/SAKK 42a; HO132 AML; HOVON-SAKK AML-102)**, it refines our ability to predict outcomes. Clinical trials randomized by MRD status should be performed to inform MRD-directed therapy according to ELN risk group, i.e. post-remission treatment, early intervention for impedending relapse, or post-tranplant surveillance.

In response to questions, it was mentioned that it remains to be evaluated whether combining various MRD methodologies has a better predictive value for outcome over a single approach. Studies are ongoing.

Methodological overview of MRD in AML

Chris Hourigan (National Institutes of Health, Bethesda, Maryland, USA)

At the beginning of his talk, Dr Hourigan highlighted the limitations of conventional CR, which was first introduced in 1956 and relies on cytomorphologic assessment. AML MRD tests provide additional information about those in CR, allowing better estimates of current disease burden and allowing stratification into those with high and low risks of "relapse" following equivalent treatment. The red line in the figure below indicates early relapse, purple line = late relapse, and the blue interrupted line = cure. The sensitivity of the different MRD techniques is shown on the right.
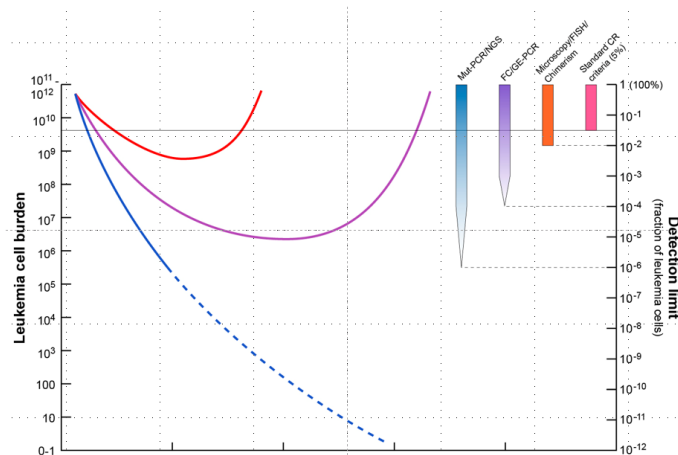
Why does MRD testing pose such a challenge in AML?

AML is a highly heterogeneous disease at various levels, including cytogenetics, somatic mutations, or the multiple subclones that can be present in a single patient or evolve over time. As AML is not one cancer, multiple MRD biomarkers might be relevant, requiring more than 1 test. Dr Hourigan went on to discuss the currently available MRD measurement tools and the current challenges. He highlighted issues with NGS that should be resolved before NGS can be become the standard, including **the stability of the marker, intrinsic errors of the test, false-positive rates,** and **uniform reporting requirements**, as well as our own **understanding of the genetic complexity of AML**.It was discussed that MRD in a patient is

---

The known unknowns:

- How do we move away from the idea of "thresholds" at "landmarks" to a more sophisticated sequential monitoring with evaluation of kinetics?
- How do we overcome the challenges of acquiring a diagnostic sample?
- How do techniques complement, or confuse, an assessment? Is orthogonal validation possible, and desirable?
- Which somatic mutations can be used for AML MRD tracking? How can analysis of large datasets help?

---

not necessarily the same as the MRD test result in a cohort, and we must expect false negatives which in an individual patient make treatment decisions difficult.

Standardization of MRD measurement

Christian Thiede (Technische Universität Dresden, DE)

Lessons can be learned from chronic myeloid leukaemia (CML), where molecular MRD tests are well standardized. Advancements were largely determined by the introduction of imatinib and real-time PCR. Sponsoring by the industry, especially Novartis, played a major role in the test standardization and generation of reference materials. Today, our ability to reliably identify patients with CML who achieve a deep response level of MR4, MR4.5, or MR5 in PB allows us to tailor tyrosine kinase inhibitor (TKI) treatment for the individual patient with "treatment-free remission" as the treatment goal.

While the pathophysiology of CML is relatively simple with one gene translocation as a target for monitoring, in AML we do not have a single molecular marker. In addition, the mutational spectrum changes with age. Thus, 68% of patients older than 60 years cannot be reliably monitored using a molecular marker. A reliable marker should also be stable in relapse, e.g. 91% of patients retain the NPM1 mutation, whereas post-onset driver mutations, e.g. FLT3-, RAS, PTPN11, and KIT, are frequently lost. However, not all persisting mutations are predictive of relapse. [Somatic CHIP mutations](#) (e.g. DNMT3A, TET2, and ASXL1, etc.) can be detected in healthy people without haematological abnormalities and can persist in persons in long-term remission and do not have prognostic significance on long-term outcomes.

Standardization of MRD measurements is at its early stages. Prerequisites for MRD standardization are use of consistent language, consensus on definitions (e.g. MR4, MR4.5), availability of reference material, and external quality controls. There is a lack of calibration material as well as availability of commercially available assays. Comparing data across the studies in AML is currently limited by choice and sensitivity of MRD tests, quality/quantity of the material, and target used. NGS is emerging as an attractive and versatile tool, but the methodology is still very expensive, laborious, and not feasible for most laboratories. Overall, our methodology needs to be improved as standardization for implementation of MRD in clinical trials of drug assessment as well as in clinical practice for individual treatment decisions.

Dr Thiede highlighted the results from the [RELAZA I](#) and [II trials,](#) which demonstrated that MRD-directed treatment (azacitidine) after allogeneic haematopoietic stem cell transplantation (HSCT) can substantially prevent or delay relapse. The trials were not randomized, and the forum discussed the ethical aspects of MRD-based randomization.

MRD STANDARDIZATION

Prerequisites for MRD standardization are use of consistent language, consensus on definitions (e.g. MR4, MR4.5), availability of reference material, and external quality controls.

Consortium on MRD in AML and industry perspective

Sharon McBain (Johnson & Johnson, USA)

Dr McBain presented the MRD in AML Industry Alliance, consisting of 4 pharmaceutical companies (Janssen, Celgene, Genentech-Roche, and Novartis) working together to help establish MRD as a surrogate for OS in AML.

Within Janssen, a dedicated AML working group has organized multiple advisory boards with important stakeholders, including global key opinion leaders (KOLs), ELN representatives, and regulatory agencies in multiple countries. Furthermore, a meta-analysis was performed to systematically assess the strength of evidence for using MRD as a surrogate endpoint in AML. In total, 32 publications were included. Preliminary results showed a 57% lower risk of death for patients achieving MRD negativity (n = 1742) compared with those who did not (n = 826) (HR 0.43; $P < 0.0001$).

The MRD in AML Industry Alliance is overseen by a joint steering committee. Specialist groups focus on 3 areas. The Methods team, co-led by Celgene and Genentech-Roche, is evaluating the potential for a core NGS platform as well as harmonization of current MFC methods. The Data group, led by Janssen, focuses on harmonizing the generation and read-out of MRD data across the alliance companies. The Stats team, co-led by Janssen and Genentech-Roche, is developing the statistical analysis plan for future meta-analyses, as well as establishing a working relationship with a third-party data gatekeeper.

In her concluding statements, Dr McBain highlighted the growing evidence of the prognostic utility of MRD status and OS in AML and the ongoing harmonization of MRD data generation. She also applauded the active collaboration between industry and academic partners in generating patient-level MRD and OS data that can be used in meta-analyses. Next steps include meetings with FDA and EMA, which are planned for next year.

The discussion focused on how to increase acceptance of MRD as a surrogate endpoint by the diverse HTA bodies. HTA bodies may not consider such outcomes as a relevant endpoint for patients. In AML, OS and QOL are closely correlated, so it might help if it can be shown that patients achieving MRD negativity have an associated improvement in QOL. An economic model to predict OS benefit based on MRD outcomes may also be helpful in this regard. The fact that different therapies are used with different treatment goals further adds to the complexity of the discussion.

WORKSHOP ON MRD IN CLL

Clinical overview: MRD in CLL

Matthias Ritgen (University Medical Center Schleswig-Holstein, Germany)

The presentation by Dr Ritgen reviewed MRD in patients with CLL across different clinical scenarios and therapeutic modalities. Across trials (CLL8, CLL10, and CLL11), there was a correlation between MRD status and PFS, suggesting that MRD is suitable as surrogate endpoint marker. However, MRD negativity was an independent predictor of PFS regardless of treatment in the CLL8 trial, which evaluated addition of rituximab to conventional chemotherapy. In contrast, in the CLL11 trial, treatment was an independent prognostic marker of PFS: patients achieving MRD negativity had better outcomes if they were treated with the newer monoclonal antibody, obinutuzumab, than with rituximab.

The role of clinical stage was highlighted by a study which found that MRD-negative patients who achieved PR and MRD-negative patients who achieved CR had similar PFS (median PFS: 61.7 vs 68.9 months, respectively). However, for patients who remained MRD-positive, median PFS was 44.4 months for those with CR and 28.1 months for patients with PR. Other factors such as IGHV mutational status, line of therapy, presence of adverse cytogenetic abnormalities, and MRD status may also impact outcomes. These differences can be explained by different residual disease growth kinetics in clinically different subgroups of patients with CLL.

Recent studies have shown the superiority of BTK inhibitors over conventional chemotherapy for patients with CLL (RESONATE-2). Despite showing significant clinical benefit, ibrutinib alone does not lead to deep molecular remission. Combining BTK inhibitors with other agents may enhance the percentage of patients becoming MRD-negative and the speed of response. In the MURANO trial, a rapid MRD response was seen in the venetoclax plus rituximab arm, followed by a stable MRD level during maintenance treatment. In contrast, MRD response decreased after the end of treatment in the bendamustine plus rituximab arm, indicating that duration of treatment is another important factor.

The biology behind the improved clinical responses with novel agents was discussed at the meeting, as was the value of using MRD to direct treatment in younger versus older patients. Dr Ritgen emphasized that in addition to MRD negativity and the timing of achieving it, multiple risk factors need to be considered in prognostic models, given the heterogeneous nature of the disease.

Industry perspective: MRD in CLL

Davy Chiodin (Acerta, AstraZeneca, USA)

Dr Chiodin presented the current state of MRD analysis in CLL. Multiple challenges exist, especially the standardization of techniques, thresholds, and timing, as well as establishing the correlation between MRD and survival outcomes. These challenges can only be overcome by collaboration between all relevant stakeholders.

Next, Dr Chiodin discussed two recent trials in patients with CLL that include MRD data in different ways. The phase 2 CAPTIVATE study assesses both MRD-guided discontinuation and fixed duration therapy with the combination of ibrutinib and venetoclax in treatment-naïve CLL. In the phase 3 MURANO study, venetoclax plus rituximab showed significant superiority over bendamustine plus rituximab in terms of PFS (median PFS not reached vs 17 months; $P < 0.0001$) and deep MRD response (MRD-negativity $<10^{-4}$ over 18 months: 45–62% vs 5–13% across different timepoints) in patients with relapsed or refractory CLL. The MRD assessment timepoints from the MURANO study may serve as an example that could be used to ensure harmony in future trials.

The use of MRD as a response criterion, definitions of optimal MRD thresholds, and corresponding clinical consequences remain to be elucidated. Principles of success for an optimal use of MRD in CLL include the following:

- Consolidation: at least 60 trials are currently analysing MRD outcomes in patients with CLL
- Confirmation of responses and how these compare to long-term benefit in survival outcomes
- Alignment of trial costs with the relevance of the scientific questions
- Healthy collaboration instead of competition between stakeholders
- Regulatory partnership is needed from the start to ensure appropriateness of the plans and implementation of existing guidelines

The need for collaboration was further stressed during the discussion.

Methodological overview of MRD in CLL

Andy Rawstron (Leeds Teaching Hospital NHS Trust, UK)

MRD testing is highly relevant in CLL because >50% of patients achieve CR with the introduction of new therapies. Furthermore, the current tests can quantify the residual disease, and that burden is of prognostic significance. In terms of the methodology, there is still an enormous amount of debate as well as scientific progress achieved over the last 20 years. The European Research Initiative on CLL (ERIC) led by Dr Rawstron validated a core panel of 6 markers for flow-cytometric MRD detection (10^{-5}) that can be available to most laboratories. Allele-specific oligonucleotide (ASO) RQ-PCR for specific immunoglobulin rearrangements has a good concordance with MFC; there is a standard assay; however, it is of limited availability. Application of NGS commercial assays in CLL is of interest (e.g. CLonoSEQ®, Adaptative Biotechnologies). The sample size is critical for a reliable quantification; while quantitative PCR requires a sample size of 250,000 cells for a sensitivity of 10^{-5}, NGS requires 2.4 million cells to obtain a 10^{-6} result (molecular approaches need approximately 2.5 × LOD cells). The goal is to achieve harmonization of reporting, i.e. achieve comparable results with different assays.

The relevant threshold of MRD has also been a long-standing matter of debate. Capturing MRD data at a single threshold can result in losing important information. Patients who are in CR but have 1% disease ("high MRD") are the ones who are going to progress within the 2–3 year window. Patients with MRD detectable at the 0.1%, 0.01%, and 0.001% level have similar PFS at 2 years, but at 5 years there is a linear improvement in PFS per log reduction in MRD. Thus, high-sensitivity MRD is predictive of long-term outcomes. Dr Rawston stressed that Identification of a single threshold as an intermediate endpoint does not preclude use of a more sensitive threshold for exploratory studies for evaluation of curative therapies, or prediction of early progression vs long-term outcomes.

There is a correlation between MRD in PB and BM and on average there is a log difference between the compartments. Different classes of drugs differ in their compartment effects. Chemotherapies in combination with antibodies deplete the blood and PB would give false-negative results. However, with ibrutinib, there is no difference between PB and BM. Information per treatment type can guide setting the appropriate threshold for MRD monitoring in PB.

Subsequent discussion focused on the challenges and opportunities of adopting a consistent MRD language.

MRD Language

- Measurable (instead of minimal) residual disease is the more accurate term for MRD

- MRD positive/negative can mean anything from <1% to <0.0001%, so "detectable" vs "undetectable" residual disease at a given threshold seem to be more appropriate terms, e.g.

  - If the assay detection rate is $1:10^{-n}$ and the sample/reagents are of sufficient quality to achieve a detection limit 10^{-n}, residual disease is not detected or measurable below 10^{-n} but above $10^{-(n-1)}$

- Lessons can be learned from CML. The CML community has defined different grades of molecular remission (MRD4, MRD5, MRD6)

- Identifying a reporting approach that is harmonized across different technologies would facilitate further advancements

---

KEY CONSIDERATIONS AND CONCLUSIONS

5 Considerations

Robert Gale (Celgene, USA)

- Prediction in a person is different from prediction in a cohort. The C-statistic, an assessment of reproducibility and variability, of current MRD tests is 0.76. This is relevant at a population level but not sufficient to guide individual treatment decisions.

- The sampling error is more important than MRD-test sensitivity and specificity at low frequencies of leukaemia cells. In low frequencies of cancer cells ($<10^{-4}$), the sample size more than the test sensitivity or specificity, determines MRD results.

- It is not necessary to get rid of every leukaemia cell to *cure* leukaemia. Approximately 40% of patients with CML and *BCRABL1* negative PCR test relapse after stopping the BTK inhibitor, and the remainder are "cured" although the genetic abnormality persists.

- It takes a long time to get CLL. Hiroshima atomic bomb survivors showed an increase in CLL incidence 50 years after the bomb, suggesting long lead time from the initial mutational event.

- MRD-testing in CLL is conceptually different than in AML or CML. In CLL, we are looking at the clone marker. In AML and CML, we are looking at the cancer marker, i.e. mutation, or mutations, that cause the disease. These are different targets: the clone versus the neoplasm.

Roundtable discussion

John Smyth (CDDF/University of Edinburgh, UK)
Axel Glasmacher (CDDF/University of Bonn, DE)

Consensus

- There was a general agreement that MRD is a unifying concept across haematologic malignancies.

- The forum agreed that MRD has the potential to become a surrogate marker of long-term outcomes in AML and adopting it as an endpoint is a matter of time. However, several issues are still to be resolved, e.g. establishing the appropriate markers and accumulating a body of data to substantiate the evidence required by the regulators.

- Collaboration is critical to resolving the issues.

- Standardized data collection and reporting is also essential.

- Lessons can be learnt from CLL, where MRD as an endpoint has evolved over 15 years and currently, MRD categories (MRD3, MRD4, MRD5) are being proposed.

Nomenclature

It was agreed that using universal MRD nomenclature would be helpful, use of "measurable" instead of "minimal", "detectable/undetectable" instead of "positive/negative", inclusion of MRD threshold (e.g. 10^{-5}). That would facilitate inclusion of this information in product labels in a consistent manner, allowing interpretation across different products.

Principles of success

The principles of "consolidate, confirm, align, healthy competition, and regulatory partnerships" were listed as a strategy for success in establishing MRD as a surrogate marker.

Challenges with AML disease biology and MRD

- Relapse in AML is a complex biologic event and it is dependent on the malignant potential of the residual leukaemic cells.

- It was mentioned that AML is not one cancer, so no reason why MRD biomarker would be one test.

- There's not necessarily a correlation between the number of residual leukaemia cells and the residual number of cells that have the biological capacity to cause relapse of leukaemia. In that regard, quantification of disease burden is a simplification, and we need to supplement it with disease biology channels information. NGS can potentially offer us information regarding residual clones.

- MRD will never explain 100% of relapses. For a regulatory endpoint we must define what is good enough. OS is a final endpoint, but neither is it 100% correlated with the treatment intervention.

Methodological considerations and quality assurance

- For AML there was yet no consensus regarding the best MRD test approach or whether one or two tests should be used.

- Standardized data collection and reporting is also an essential goal. There is a need to educate HCPs beyond the clinical trialists, including people who collect and transport the data, laboratory technicians, as well as the people writing the protocols.

- There was agreement that BM aspirations and biopsies should be drawn by well-trained individuals to ensure the right technique without increasing pain for the patient. Reporting should capture information on describing an MRD test (sensitivity, specificity, LOD, LOQ), the quality of the sample, etc.

- There was a consensus that MRD measurements should be performed in a central laboratory because of the technical skills and expertise required to perform the analysis.

- Laboratories should fulfil the requirements of the ELN accreditation programme. Shipment and handling of material is another important consideration.

- Biobanking is recommended to future-proof RCTs because the technology evolves. Considerations for high-quality sample acquisitions remain.

How good is good enough?

- FDA recommends that the sensitivity of the MRD assay should be at least 10-fold below the decision-making threshold (definition of detectable MRD), e.g. LOD. For example, if MRD positive or negative is defined as detection of greater or less than 1×10^{-5}, the assay should be optimized and validated to have an analytical sensitivity of at least 1×10^{-6}, e.g. limit of quantification (LOQ). Reporting this information is essential to interpreting the results. Moreover, the results are largely dependent on the quality of the sample (e.g. amount of DNA or number of cells analysed; haemodilution). The quality of the sample should be reported.

- While competition drives innovation, it is important to agree on the common ground in pre-competitive initiatives. Dr Rawstron highlighted that while diagnostic laboratories continue to search for more sensitive and specific assays, they present a united front on the minimum test requirements for regulatory assessment of new drugs in CLL. In that regard, consensus is more important than having the best MRD test.

- Strategies are needed to avoid a high amount of missing data and test failures. Missing data should be treated as non-responding patients.
- FDA and EMA encourage drug developers to seek early and frequent conversation with them, with the goal to bring the product in line with the data that is needed. Patient-level data might be critical to address questions on efficacy across subgroups based on various cytogenetics, lines of therapy, different subpopulations, etc. The possibility to seek simultaneous input from EMA and FDA was raised. Dr Gormley responded that this is a reasonable and feasible approach, and there is a process for joint advice.

Consortia

- The goal of the industry consortium is harmonization of MRD data generation in prospective trials. This can be achieved by standardized protocols and quality assurance, including study design, timing of sampling, sample requirements, instruments and technique of BM aspiration, material collection and transportation requirements, etc.
- Consortia (industry consortia; Harmony; EuroFlow Consortium) have the potential to resolve the unanswered questions associated with MRD.
- Data-sharing and openness of standards is going to be important in taking the concept of MRD further.
- Meta-analyses are required to demonstrate the prognostic association between MRD and long-term outcomes.

Future meetings

- The small format of the meeting stimulates productive discussion among industry, regulators, and academia. Input from patients and HTAs is currently lacking and should be considered for future meetings.
- It was recommended that future meetings poll information on top priorities to focus the debate prior to the meeting among attendees.

ABBREVIATIONS

ALL, acute lymphoblastic leukaemia

AML, acute myeloid leukaemia

ASO, allele-specific oligonucleotide

BM, bone marrow

BTK, Bruton's tyrosine kinase

CAR, chimeric antigen receptor

CBF, core binding factor

CDDF, Cancer Drug Development Forum

CHIP, clonal haematopoiesis of indeterminate potential

CHMP, Committee for Medicinal Products for Human Use

CLL, chronic lymphocytic leukaemia

CML, chronic myeloid leukaemia

CR, complete response

CRi, complete response with incomplete haematologic recovery

CR_{MRD-}, complete response without minimal/measurable residual disease

DDT, Drug Development Tool

DfN, different from normal

EFS, event-free survival

ELN, European LeukemiaNet

EMA, European Medicines Agency

ERIC, European Research Initiative on CLL

FDA, Food and Drug Administration

HSCT, haematopoietic stem cell transplantation

KOL, key opinion leader

LAIP, leukaemia-associated immunophenotype

LOD, limit of detection

LOQ, limit of quantification

LSC, leukaemia stem cell

MFC, multiparameter flow cytometry

MHRA, Medicines and Healthcare Products Regulatory Agency

MM, multiple myeloma

MOA, mechanism of action

MRD, minimal/measurable residual disease

NGS, next-generation sequencing

OS, overall survival

PB, peripheral blood

PFS, progression-free survival

PI, prescribing information

PR, partial response

QOL, quality of life

RCT, randomized controlled trial

RQ-PCR, real-time quantitative PCR

RR, response rate

TKI, tyrosine kinase inhibitor

UKNEQAS, United Kingdom National External Quality Assessment Service

If you have any inquiry about the CDDF multi-stakeholder workshops or publications, please contact the CDDF office via email (info@cddf.org) or by phone (+32 2 880 62 70). Thank you for your interest in CDDF activities and continued support.