Critical Reviews in Oncology / Hematology 133 (2019) 142–148

Contents lists available at ScienceDirect

Critical Reviews in Oncology / Hematology

journal homepage: www.elsevier.com/locate/critrevonc

Clinical implications of measurable residual disease in AML: Review of current evidence

Ine Moors[a,*], Karl Vandepoele[b], Jan Philippé[c], Dries Deeren[d], Dominik Selleslag[e], Dimitri Breems[f], Nicole Straetmans[g], Tessa Kerre[h], Barbara Denys[b]

[a] Ghent University Hospital, Dpt. Haematology, C. Heymanslaan 10, 9000 Ghent, Belgium
[b] Ghent University Hospital, Laboratory of Molecular Diagnostics, C. Heymanslaan 10, 9000 Ghent, Belgium
[c] Ghent University Hospital, Dpt. Laboratory Medicine, C. Heymanslaan 10, 9000 Ghent, Belgium
[d] AZ Delta, Dpt. Haematology, Wilgenstraat 2, 8800 Roeselare, Belgium
[e] AZ Sint-Jan, Dpt. Haematology, Ruddershove 10, 8000 Brugge, Belgium
[f] ZNA Stuivenberg, Dpt. Haematology, Lange Beeldekensstraat 267, 2060 Antwerpen, Belgium
[g] University Hospital Saint-Luc, Dpt. Haematology, Avenue Hippocrate 10, 1200, Brussels, Belgium
[h] Ghent University Hospital, Dpt. Haematology, C. Heymanslaan 10, 9000, Ghent, Belgium

ARTICLE INFO

Keywords:
Measurable or minimal residual disease
Acute myeloid leukaemia
Flow cytometry
Next generation sequencing
Prognosis
MRD-directed treatment

ABSTRACT

Despite the fact that 80% of adult acute myeloid leukaemia patients reach complete morphological remission after induction chemotherapy, many of them relapse. Many studies have shown that detection of minimal residual disease (defined as 'any detectable evidence of persistent leukaemic cells during complete morphological remission') has an added value in prediction of relapse and survival, and is more than just a surrogate marker for already known risk factors in AML. As such, the behaviour of the disease during treatment might become equally or even more important to decide whether or not an upgrade of treatment (such as an allogeneic stem cell transplantation) is necessary to improve outcome. However, there are still many open issues as to what the ideal time point is to measure MRD, which threshold is clinically significant, what sample (peripheral blood or bone marrow) should be used and how we can standardize tests so that results from different labs become comparable. This review gives an overview of currently available evidence regarding technical issues, prognostic impact and MRD-directed treatment in AML.

1. Introduction

Despite the fact that 80% of adult acute myeloid leukaemia (AML) patients reach complete morphological remission (CR) after induction chemotherapy, many of them relapse, with a 5-year overall survival (OS) of only 30–40%. Relapse is the result of incomplete eradication of disease, clonal evolution from preleukaemic clones, or therapy-related mutagenesis. To date, remission in clinical studies has been defined as purely morphological remission (< 5% blasts on bone marrow), but during the past decades, new much more sensitive techniques have been developed to detect minimal or measurable residual disease (MRD), which is defined as 'any detectable evidence of persistent leukaemic cells below the threshold of complete morphological remission' (Terwijn et al., 2013; Benton and Ravandi, 2017).

MRD already has an established role in several haematological malignancies such as chronic myeloid leukaemia (CML), acute lymphoblastic leukemia (ALL) and acute promyelocytic leukemia (APL), where it impacts on treatment decisions, such as therapy switch, consolidation with/without allogeneic stem cell transplantation (alloSCT) and administration of pre-emptive therapy. In non-APL AML however, although extensive data is available on the prognostic impact of MRD during the course of intensive chemotherapy, it does not yet routinely determine post-induction treatment strategies (i.e. chemotherapy/autologous stem cell transplantation versus allogeneic stem cell transplantation), which is nowadays mainly based on pre-treatment leukaemia-specific characteristics, such as cytogenetics and a limited number of molecular aberrations. Nor is there evidence to guide pre-emptive therapy in case of molecular relapse, as is the case for APL.

In this review we discuss some technical aspects of MRD and provide an overview of what is currently known about the prognostic

* Corresponding author.
E-mail addresses: Ine.Moors@ugent.be (I. Moors), Karl.Vandepoele@uzgent.be (K. Vandepoele), Jan.Philippé@uzgent.be (J. Philippé), Dries.Deeren@azdelta.be (D. Deeren), Dominik.Selleslag@azsintjan.be (D. Selleslag), Dimitri.Breems@zna.be (D. Breems), Nicole.Straetmans@uclouvain.be (N. Straetmans), Tessa.Kerre@Ugent.be (T. Kerre), Barbara.Denys@uzgent.be (B. Denys).

https://doi.org/10.1016/j.critrevonc.2018.11.010
Received 16 May 2018; Received in revised form 30 August 2018; Accepted 23 November 2018

impact of MRD during AML treatment, and its possible impact on treatment decisions.

2. Technical aspects of MRD detection

Different techniques for detecting low amounts of residual leukaemic cells are available: *multiparameter flow cytometry* (MFC) following a known 'leukaemia-associated immunophenotype' (LAIP) during therapy, or searching for 'different-from-normal' (DFN) populations (sensitivity up to 10E-4 to 10E-5, available in ~80% of patients); *(q)RT-PCR* detecting specific molecular abnormalities and quantifying abnormal leukaemic gene transcripts in comparison to normal genes in the same sample (sensitivity up to 10E-4 to 10E-5, available for 60–70% of patients); *mutational analysis* e.g. by next-generation sequencing (NGS); *FISH and karyotyping*; as well as *gene expression profiling* (e.g. WT1 overexpression).

Each technique has its own advantages and limitations, which makes the choice for the most informative approach for each AML subtype, challenging. They all differ in terms of sensitivity, applicability, complexity and the ability to evaluate clonal evolution. To keep in mind, MRD is never 100% reliable: false-negative MRD could be the result of immunophenotypic changes, residual leukaemic cells hiding in difficult-to-sample micro-environments, bone marrow sample hemodilution, the small number of residual leukaemic cells, clonal evolution, the loss of a specific mutation. On the other hand, some positive MRD-values need to be interpreted with caution too: residual *CBFB-MYH11*-transcripts might represent a preleukaemic clone, flow cytometry might detect a regenerating clone with a different-from-normal phenotype, residual mutations such as DNMT3A, ASLX1, TET2 could reflect CHIP (clonal haematopoiesis of indetermined potential, which is not always associated with a higher risk of relapse and can be detected in long-term remission) (Jongen-Lavrencic et al., 2018; Mosna and Gottardi, 2016).

A very important discussion related to MRD, concerns the questions about the *optimal cut-off* and *the ideal time point* to measure MRD. For MFC and different leukaemia-specific RT-qPCR assays (fusion transcripts and mutations like NPM1), many studies have defined particular threshold levels or 'cut-off' values that are predictive for outcome at various time points. However, these cut-off values are not universally applicable. A general problem is the fact that many of the techniques are not yet standardized between laboratories. Therefore, it is difficult to compare results between different studies, or to implement the results from studies in the daily practice in one's own routine lab.

For RT-qPCR assays, this issue could be solved by reporting *relative* values (compared to diagnostic samples), as this will correct for methodological differences between laboratories. Moreover, *at diagnosis, absolute* values of several markers repeatedly failed to correlate with outcome (Schnittger et al., 2009; Krönke et al., 2011; Yin et al., 2012; Corbacioglu et al., 2010; Perea et al., 2006; Pigazzi et al., 2015; Zhang et al., 2013). Also *during therapy*, Hubmann et al. found that NPM1 *kinetics*, but not absolute levels, were significant prognostic indicators for relapse risk (Hubmann et al., 2014). Similar results were found for t(8;21)-AML (Yin et al., 2012; Pigazzi et al., 2015; Jourdan et al., 2013): for example, in the MRC-AML15 trial, where log reduction of BM *RUNX1-RUNX1T1*-transcripts was more prognostic than absolute copy numbers in an adjusted regression analysis.

In MFC as well, different methods are used to quantify the MRD burden (relative to CD45+ leukocytes or CD34+ cells or total viable cells; correction for percentage LAIP-positive blasts at diagnosis). To harmonize reporting, the ELN guidelines recommend to relate LAIP events to the CD45+ leukocytes (excluding CD45- erythroblasts) (Schuurhuis et al., 2018).

In addition, the *ideal time point* to measure MRD is yet unclear. The time point at which MRD is most discriminatory might differ depending on the marker or AML subtype, and might be dependent on what it will be used for: MRD at early time points may impact post-remission treatment, MRD during follow up could trigger pre-emptive therapy.

Table 1
Key messages on clinical impact of MRD.

Clinical impact of MRD: take home messages
NPM1
During aplasia after the 1st induction, a 1-log reduction in BM-MRD is informative on relapse risk, independent of other risk factors.
There is convincing (retrospective) evidence that presence of PB-MRD after 2 cycles is an independent negative prognostic factor that can even overrule pre-treatment risk factors such as WBC count, DNMT3A or FLT3 status.
During follow up, persistence of NPM1 mutation is possible in long-term remission, albeit at low copy numbers, and most of the patients return negative on subsequent samples.
One log rise in MRD predicts a haematological relapse.
As BM-MRD is 1 log more sensitive than PB-MRD, BM-samples are preferred in follow-up phase after the end of treatment, although it has not been shown to improve outcome yet
NPM1 seems stable as MRD marker, but clonal evolution from an ancestral clone without NPM1 mutation is possible, as well as evolution to NPM1-negative MDS/MPN.
Inv(16)
High copy numbers after the 1st cycle are associated with a high relapse rate.
Two negative BM-MRD samples during consolidation and early follow-up (3 months) is associated with a high probability of continuing CR.
Persistent low-level MRD positivity is possible in long-term remission, reflecting a silent pre-leukaemic clone.
As there is more data on the value of *absolute* transcript levels, more efforts are necessary to standardize MRD testing, and/or more research is needed on MRD kinetics.
t(8;21)
More than 3-log reduction after the 2nd cycle is associated with very low risk of relapse.
BM-MRD negativity after treatment is not a prerequisite for continued remission. In contrast, PB-MRD is more predictive for relapse after treatment.
Median time to haematological relapse after MRD conversion is 4-5 months.
Multiparameter flow cytometry
Already at early time points during therapy, MFC-MRD is highly predictive for outcome, however, the most discriminatory time point has not been identified yet.
A cut-off of 0.1% can be used to define MRD negative versus positive.
Adding MFC- MRD to known risk factors can change prognostic categories.
Adding information on leukaemic stem cells to MFC-MRD after the 2nd cycle further improves prognostication.
Next generation sequencing
The use of special, more sensitive NGS techniques is a prerequisite to analyze MRD.
Not all genes have the same impact on prognosis: persistence of DTA-mutations (DNMT3A, TET2 and ASXL1) is not correlated with higher relapse rate. In contrast, persistence of non-DTA mutations during complete remission has negative prognostic impact on relapse and survival.
A clear threshold still has to be determined.
NGS MRD cannot be applied on patients with recurrent translocations, where MRD is still best detected using qPCR.

Finally, the preferred sample type is under discussion. Peripheral blood (PB) has certain practical advantages over bone marrow (BM), and some studies show significant concordance between the two sample types (Maurillo et al., 2007; Zeijlemaker et al., 2016). However, PB is characterized by a lower MRD burden, and correlation with BM results, as well as applicability for outcome prediction may depend on the target analyzed and the time point of the analysis.

A recent report by the MRD working party of the ELN provides an overview of the studies that have been published for several markers, with detailed recommendations regarding technical aspects and reporting of MRD results (Schuurhuis et al., 2018).

In conclusion, for every different marker, a clinically meaningful time point, threshold and sample type needs to be carefully determined before implementing MRD in daily clinical practice. This will be discussed in detail for several markers.

3. Clinical impact of MRD detection (Table 1)

3.1. NPM1 mutations

NPM1 mutation is present in 25–30% of AML, and 45–60% of normal karyotype AML, which makes it a very useful tumour marker in AML. This is reflected by the presence of many, although all retrospective, papers on the subject. The *absolute* NPM1 mutation level at diagnosis does not seem to predict outcome (Schnittger et al., 2009; Krönke et al., 2011), but already early in the induction phase, during aplasia (day 16–18) after the 1st cycle, MRD *kinetics* do. Despite the fact that virtually all patients remain MRD positive at this point, more than 1-log reduction in BM-NPM1 is associated with a lower cumulative incidence of relapse (CIR) within 100 days, even after correcting for ELN risk classification (Hubmann et al., 2014). Another study could not show *absolute* values of NPM1 transcripts to be predictive of CIR, nor OS, at this time point (Krönke et al., 2011).

MRD after two cycles has been identified as the most important time point in NPM1-AML by different research groups.(Krönke et al., 2011; Hubmann et al., 2014; Ivey et al., 2016; Balsat et al., 2017) At this time point, *kinetics* are independent prognostic for relapse incidence: < 3-log reduction of BM-NPM1 was identified by Hubmann et al. as a significant threshold with a positive predictive value for relapse of 65% and negative predictive value of 74%, independent from ELN risk classification, WBC count and LDH on multivariate analysis (Hubmann et al., 2014). A French group found that the presence of PB-MRD after two cycles had predictive power, with less than 4-log reduction being associated with a higher relapse incidence (3Y CIR 65.8% versus 20.5%) and lower OS (40.8% versus +/- 92%). This was independent of pre-treatment variables such as abnormal karyotype, and FLT3-ITD status (Balsat et al., 2017). Hence, both research groups showed that MRD at this time point **adds valuable information to the already known pre-treatment risk factors** for predicting relapse and survival.

Even more striking are the findings of Ivey et al., who found the presence of *PB*-MRD for NPM1 after 2 cycles to be the *only* prognostic factor for relapse, independent of WBC count, DNMT3A status or FLT3 status, and the *sole* prognostic factor for mortality, independent of WBC count, thereby even **overruling pre-treatment risk factors** (Ivey et al., 2016). Similar findings are seen in the paper by Hubmann et al. who showed *BM*-MRD kinetics (3-log reduction) after induction to be the *sole* prognostic risk factor for relapse (not OS) in a multivariate analysis with ELN risk category, age and WBC count (Hubmann et al., 2014).

In addition, Ivey et al. showed that within the group of patients with high risk mutations such as DNMT3A and/or FLT3-ITD, the absence of NPM1 PB-MRD (with a sensitivity of at least 0.01%) after two cycles identified a subgroup of patients with favorable outcome (survival 76%). Also, in the absence of these mutations, the presence of PB-MRD after two cycles was associated with very poor outcome (Ivey et al., 2016). Although not yet tested in a prospective randomized trial, these finding clearly **challenge our current transplant indications.**

At the end of treatment, presence of BM-MRD is associated with an increased risk of relapse and decreased OS (Krönke et al., 2011; Lambert et al., 2014). However, during follow-up, low levels of BM-NPM1 can be present in long-term CR patients. In the study of Krönke et al., 51% of 136 patients had at least 1 positive follow-up sample (BM), with 1 in three patients remaining in CR with a median follow-up of 24.8 months (range 1.8–80.5). All had low levels and the majority turned negative in one of the following samples (Krönke et al., 2011).

Nevertheless, 1-log increase in BM-NPM1 transcripts is associated with clinical relapse, within a median time of 60–133 days (Schnittger et al., 2009; Krönke et al., 2011; Ivey et al., 2016). Therefore, it seems wise to analyse BM-MRD after treatment every 2–3 months for 18–24 months. However, it has not yet been established whether treating early molecular relapse is associated with a survival benefit, nor which kind of pre-emptive treatment would be preferred. Moreover, there is still some discussion on the stability of NPM1 as MRD marker. Data from the literature are ambiguous. In the study of Ivey et al., NPM1 mutation was lost in only 1 in 70 relapsed patients, whereas Hubmann et al. described 3 in 45 patients who lost NPM1 (Hubmann et al., 2014; Ivey et al., 2016). Relapse can occur from the initial AML clone, but also from an ancestral clone, with or without clonal evolution (Krönke et al., 2013), leading to NPM1-negative relapses. On the other hand, development of myelodysplastic syndromes or myeloproliferative disease without NPM1 mutation has been described as well. In the paper of Ivey et al., patients were young, and the phenomenon of preleukaemic clones and CHIP is more rare in this population. This could explain the stability of NPM1 as an MRD marker in their population.

Finally, there is the question whether peripheral blood should be preferred over bone marrow for MRD-analysis. After reviewing the literature, the answer depends on the time point of the MRD-analysis. Krönke et al. show a higher sensitivity of BM-MRD **during the early time points**, because 46% of the MRD-negative PB-samples had a paired positive BM-sample (Krönke et al., 2011). However, they did not correlate these findings with outcome. Ivey et al. did and could show that - although less sensitive - PB-MRD at early time points was more discriminatory for relapse and OS (Ivey et al., 2016). In contrast, **during follow-up**, Ivey et al. showed that BM-MRD was 1 log more sensitive than PB (Ivey et al., 2016). Therefore, one could state that MRD testing in BM àn PB is preferred *during* treatment, in contrast to the *follow-up phase*, where BM is still more sensitive.

3.2. Core-binding factor-AML (CBF-AML)

CBF consists of a group of transcription factors that are involved in normal haematopoiesis. Inv(16) and t(8;21) lead to loss of function of this complex, resulting in differentiation blockade. These leukaemias are generally chemo-sensitive, but nonetheless, 30% will relapse. MRD can potentially identify those CBF-AML patients with high risk of relapse, who might benefit from a therapeutic upgrade, such as an allogeneic stem cell transplantation in first complete remission.

3.2.1. CBFB-MYH11 or inv(16)

Pre-treatment copy numbers (CN) are not informative on CR, CIR or OS (Yin et al., 2012; Corbacioglu et al., 2010; Perea et al., 2006). At the early time points during therapy, there is some evidence that molecular MRD informs on outcome and can be used to adjust treatment strategies. Based on solid data in more than 100 patients, Yin et al showed in a prospective study that PB-*CBFB-MYH11* CN > 10/10E5 ABL1 copies after the 1st cycle was predictive of relapse and survival (e.g. with a CIR 21% if < 10 CN and 100% if CN > 500) in multivariate analysis. (Yin et al., 2012). In contrast, other authors could not confirm these findings in smaller series (Corbacioglu et al., 2010; Perea et al., 2006; Pigazzi et al., 2015; Qin et al., 2015).

Later during consolidation, different studies identified thresholds informative for relapse incidence and/or mortality (e.g. PB CN > 10 (Yin et al., 2012)or 3-log reduction (Qin et al., 2015) after the 3rd cycle). In a study by Corbacioglu and colleagues, evaluating MRD in prospective German-Austrian AML Study Group treatment trials, patients with at least two PCR-negative BM-MRD during consolidation and early (3 months) follow up, have the highest probability of remaining in CR, with a 90% relapse-free survival rate (Corbacioglu et al., 2010).

During follow-up, persistence of inv(16) has been described in patients in long-term remission: in the series of Corbacioglu et al., 11 out of 35 patients in continuing CR had positive PCR, albeit with low copy numbers (Corbacioglu et al., 2010). This could be explained by the 'two-hit theory' of AML ontogenesis: in this model, the CBF translocation is the '1st hit', leading to impaired differentiation and giving rise to a preleukaemic phase, that might remain clinically silent until true leukaemia develops after a '2nd hit', being a mutation/genetic aberration leading to excessive proliferation (e.g. KIT or RAS mutations) (Mosna and Gottardi, 2016). Thus, persistent PCR-positivity is possible

during follow-up, and does not need therapeutic intervention per se. Therefore, Ivey et al. and Corbacioglu et al. identified critical MRD thresholds in both BM and PB above which relapse occurs (BM *CBFB-MYH11* > 50/10E5 ABL1 copies or PB *CBFB-MYH11* > 10/10E5 ABL1 copies). (Corbacioglu et al., 2010; Ivey et al., 2016). However, due to lack of standardization (as mentioned above), implementation of those cut-offs into your own clinical practice remains challenging. But as a general rule, it can be stated that conversion from negative to positive PCR and ≥ 1 log rise in PCR levels are predictive of imminent relapse (Corbacioglu et al., 2010; Lane et al., 2009).

Finally, on the question of what source of samples to use, Corbacioglu et al. showed higher sensitivity of bone marrow samples during therapy, whereas peripheral blood samples were equally informative in follow up (Corbacioglu et al., 2010).

3.2.2. *RUNX1-RUNX1T1 or t(8;21)*

Different study groups showed in pediatric patients that absolute values of *RUNX1-RUNX1T1*-transcripts during treatment do not correlate with outcome, but log-reduction does (Yin et al., 2012; Jourdan et al., 2013; Zhu et al., 2013). These findings however, could not be confirmed in a multivariate analysis, probably because of lack of power. The prospective CBF-2006 trial (ALFA group) identified the time point before 2nd consolidation as the only point of prognostic value, with < 3-log reduction being associated with worse prognosis, which was more powerful than the identification of cKIT mutations (Jourdan et al., 2013). In this trial, a transplant was offered to poor MRD-responders, but the study was not designed to demonstrate whether allogeneic SCT could overcome this poor prognosis and improve survival. Moreover, failing to achieve a 3-log reduction in *RUNX1-RUNX1T1* after the second consolidation could also discriminate high-risk patients in the prospective studies of Zhu et al (AML05 trial) (Zhu et al., 2013) and Yin et al (MRC AML-15 trial) (Yin et al., 2012).

At the end of treatment, only 30% of patients reach complete molecular remission in bone marrow (Willekens et al., 2016) and a significant portion of t(8;21)-patients remains positive during a long time without a haematological relapse (e.g. 9% of all patients after up to 2 years of follow up according to Yin et al. (Yin et al. (2012))). This probably reflects persistence of *RUNX1-RUNX1T1* positive non-leukaemic cells such as haematopoietic stem cells, B cells, monocytes and/or mast cells. However, values of BM- and PB-MRD are consistently lower in patients in continued remission (Jourdan et al., 2013), and conversion of MRD negativity to positivity during follow-up is associated with a high risk of relapse, after a median of 4–5 months(Yin et al. (2012); Jourdan et al., 2013). During follow-up bone marrow is more sensitive than peripheral blood, but only PB-MRD is predictive for relapse (Willekens et al., 2016): i.e. the latter is more discriminatory.

To date, it is not clear at what time point and at which level of MRD, treatment should be upgraded. We keep in mind that a high percentage of CBF-AML patients can be salvaged with rescue chemotherapy, which makes it difficult to correlate MRD to survival. It is not yet clear whether allogeneic SCT in CR1 is improving outcome in good risk AML with poor MRD response.

4. Multiparameter flow cytometry

Residual leukaemic cells can be identified by MFC using 'leukaemia-associated immunophenotypes' (LAIP) or 'different-from-normal approach' in the majority of the patients (~80%). The first is based on patient-specific antibody panels that are defined at diagnosis and used during follow up, the latter is based on the identification of an aberrant differentiation and maturation pattern in follow-up. It is recommended to combine both approaches to define MRD, allowing detection of immunophenotype shifts and monitoring patients for which information from diagnosis is not available. By using sufficiently large panels of antibodies (minimum 8 colors) and same tubes at diagnosis and at follow-up, both approaches can be combined (Schuurhuis et al., 2018).

The recently published consensus document on MRD from the ELN provides an excellent overview of the most important studies on the prognostic value of MRD by MFC (Schuurhuis et al., 2018). With significant cut-off levels between 0.035% to 0.2%, a general cut-off of 0.1% was found to be most relevant to distinguish MRD-positive from MRD-negative patients (Schuurhuis et al., 2018), although there is still some discussion on this subject. The threshold is possibly dependent on the time point of collection.

At present, the prognostic significance of MFC-MRD has been established by many different study groups, with a clear advantage for risk analysis at early time points. As it might be difficult to discriminate true MRD based on LAIP from regenerating haematopoietic cells, the phase of aplasia after the first induction might be the ideal moment for MFC-MRD. Könkhe et al. showed that MRD positivity at day 15 post-induction is an independent risk factor for RFS and OS, independent of age or cytogenetic risk. They could not demonstrate additional value from MRD at the end of induction (Köhnke et al., 2015). This was in contrast to Freeman et al, who prospectively identified the time point after the 1st induction as the most informative one in patients > 60 years, independent of other pre-treatment risk factors. (Freeman et al., 2013). In addition, Terwijn et al. prospectively showed that MRD after the 1st cycle, the 2nd cycle and even during consolidation, is prognostic for relapse incidence, RFS and OS, mainly during the first year, and much less for relapse beyond one year (Terwijn et al., 2013).

The load of leukaemic stem cells (LSC) at diagnosis has been proven to be of prognostic value in both adults and children (Hanekamp et al., 2017). Adding information on the presence or absence of LSC to MFC-MRD after the 2nd cycle improves prognostication with MRD-pos/LSC-pos patients having the worst prognosis with 100% 3Y CIR and 100% 3Y mortality (in contrast to resp. 35% and 66% for MRD-neg/LSC-neg patients). This added value was confirmed in a multivariate analysis, correcting for all known risk factors, including HOVON risk group and post-remission treatment (Zeijlemaker et al., 2017).

Although retrospective, Buccisano et al. clearly showed that MFC-MRD-positive good- and intermediate-risk AML behave as bad as poor-risk and FLT3-positive AML with regard to RFS and OS. On the other hand, MRD-negative good risk and intermediate risk have a similar outcome, concluding that the combination of pre-treatment factors and post-consolidation quality of response improves the definition of prognostic categories (Buccisano et al., 2010).

BM-MFC is preferred above PB-MFC because of a higher sensitivity, although pre-analytical issues are of great concern. As larger amounts of bone marrow are aspired to allow analysis of more events (up to 10E7), there is a higher risk of dilution with PB and subsequent loss of sensitivity, unless the first portion of the aspirate is used.

5. Next generation sequencing (NGS)

MRD detection by qPCR analysis and flow cytometry have proven their use in AML. However, there is still an portion of patients that lack a traceable marker, and clonal evolution or phenotypic switching limits the use of current MRD techniques. NGS could fill this gap as a broad variety of mutations are detected (Cruz et al., 2017).

After the first sequencing of an AML genome in 2008 (Ley et al., 2008), several other NGS projects demonstrated that AML is characterized by a low number of mutations, with ~25 frequently mutated genes. The number of mutations per patient depends on the gene set analyzed and ranges from 3 to 12 (Jongen-Lavrencic et al., 2018; Hirsch et al., 2017; Klco et al., 2015) and in most studies, ≥90% of patients have at least one mutation, demonstrating the theoretical applicability of NGS to detect MRD in a large group of patients. Moreover, a combined analysis of several markers could circumvent the false-negative results that are sporadically seen when only a single target is analyzed. The introduction of MRD detection by NGS was hampered by the limited sensitivity (1–5%) when large gene panels are used. A recent report solves this problem by performing MRD with targeted deep sequencing

in a patient-tailored analysis (Malmberg et al., 2017). The analysis of a patient-specific set of genes, where mutations were detected at diagnosis, enables a very high coverage (and sensitivity) of these regions at an acceptable cost. Other groups use error-corrected NGS with unique molecular identifiers (UMI) (Hirsch et al., 2017; Young et al., 2016). UMI's are random DNA-barcodes that are attached to individual DNA fragments before the first amplification reaction. When variants are only detected in fragments containing the same barcode, this variant is most likely due to a technical error, introduced during PCR amplification. Only when the same variant is found in fragments with different barcodes, it can reliably be called as a true positive variant. The use of UMI's thus increases the discriminatory power between true positive variants and sequencing artefacts, increasing the sensitivity of the assay. A third approach uses statistical methods to reliably discriminate true positives from sequencing background noise (Jongen-Lavrencic et al., 2018).

The existence of age-related CHIP in elderly individuals complicates the MRD analysis, as it has been shown that patients in remission can show even high levels of these mutations, which are not necessarily associated with an increased chance of relapse. Additionally, the clinical context in which the NGS analysis is performed is critical for a correct interpretation. In theory, CHIP mutations are not good targets to detect MRD, although in the context of alloSCT, such mutations could have potential interest (Hirsch et al., 2017), as also these mutations will be cleared by a successful treatment.

While early reports showed the applicability of NGS to detect MRD in single genes, more recent publications use an unbiased approach and detect multiple mutations in one analysis. For instance, in a prospective, single center study of 50 patients where genome analysis was done at diagnosis and 30 days after successful induction therapy, almost half had one or more persistent leukaemia-associated mutations (> 2.5% allelic frequency) at remission (Klco et al., 2015). These patients showed reduced EFS and OS when compared to the group that did not show persistent mutations. Another prospective study of 69 patients could only detect a significant impact on LFS and OS when more than one mutation was present above the threshold (0.4%) when the patient was in CR (Hirsch et al., 2017).

A recent, larger study of 482 patients with a 54-gene panel detected an average of 2.9 mutations per patient (Jongen-Lavrencic et al., 2018). Similar to the smaller studies described above, half of the patients showed persistent mutations during CR (51.4%) and these were most common in *DNMT3A*, *TET2* and *ASXL1* (abbreviated as *DTA*), genes known to be frequently mutated in patients with CHIP (Hirsch et al., 2017; Acuna-Hidalgo et al., 2017). The frequency of the persistent mutations ranged from 0.02% to 47%, with higher allele frequencies for the *DTA* mutations. It has been suggested before that this can be explained by the expansion of non-leukaemic clones that repopulate the bone marrow after induction therapy (Wong et al., 2016). The detection of persistent DTA mutations alone during CR was not correlated with an increased risk of relapse. However, persistent non-*DTA* mutations were seen in 28.4% of patients and this was associated with a significantly higher relapse rate (4-year relapse rate of 55.7% with detection vs 34.6% without detection), and with a lower 4-year RFS and OS. This significance was retained in a multivariate analysis that accounted for the major relevant prognostic factors (Jongen-Lavrencic et al., 2018).

These papers show promising data for the use of NGS to detect MRD, but further delineation of the threshold and which genes to include in the algorithm seems necessary.

6. Combination of different MRD techniques

Whatever the technique used, an important percentage of MRD-negative patients still relapse. Combining different techniques for MRD monitoring will probably improve prognostication. Perea at al. combined MFC (≤0.1%) with CBF-transcript levels (< 10 CN) at the end of treatment to identify a patient group with excellent outcome (Perea et al., 2006). Similar results were described by other authors, especially in case PCR results were not very low or high (and therefore less informative). Recently, Jongen-Lavrencic et al. combined the results from the NGS-MRD analyses described above, with MFC for a group of 340 patients (Jongen-Lavrencic et al., 2018). Concordant results were found in 69.1% of patients. However, 64 NGS positive patients did not show MFC-MRD, and 41 MFC positive patients did not show NGS-MRD. Patients that showed MRD in both NGS and MFC showed the highest 4-year relapse rate (73.3%). When only one method was positive, similar relapse rates were seen for NGS and MFC-positive patients (52.3% and 49.8%, respectively). Patients for whom neither of the assays was positive had the lowest relapse rate (26.7%) (Jongen-Lavrencic et al., 2018). These data show that combining of MRD techniques results in refinement of prognostication.

7. Implications for treatment

7.1. Effect of treatment on MRD

Several studies show that intensifying chemotherapy (e.g. reinforced induction in CBF-2006 trial, addition of Clofarabine to standard chemotherapy in HOVON102 study or addition of gemtuzumab ozogamycin in the ALFA0701-study) leads to a better quality of remission in patients in CR (i.e. more rapid decline in MRD and/or more MRD-negative patients), sometimes (but not every time) resulting in better outcome (Jourdan et al., 2013; Lambert et al., 2014; Löwenberg et al., 2017). These data suggest that MRD might be a reasonable candidate to be a surrogate marker for survival. Using MRD as surrogate endpoint could speed up approval of new drugs, but for now, the evidence is not yet strong enough.

7.2. Impact of MRD on treatment

Currently, there are no prospective randomized controlled trials proving that MRD-directed treatment improves outcome. However, in daily practice and in certain study protocols, MRD is already used, for example to decide on alloSCT in CBF-AML (Jourdan et al., 2013). However, and especially for CBF-AML, we need to be cautious about this, because survival is not always compromised by relapse in this chemo-sensitive group, thanks to highly effective salvage regimens: e.g. < 3-log reduction in t(8;21)-AML is predictive for relapse risk (with a CIR of 30% vs 4%), but not predictive for 5Y OS (Yin et al., 2012). AlloSCT on the other hand will not necessarily improve outcome in poor MRD-responders due to the important treatment-related mortality. Prospective randomized trials could answer this question, but are not available yet. However, there is some retrospective data supporting the strategy of alloSCT for poor MRD-responders. Buccisano et al. clearly showed a benefit of alloSCT over autologous SCT or chemotherapy in MRD-positive good- and intermediate-risk AML patients, but not for MRD-negative patients (Buccisano et al., 2010). Similar results were seen by Qin et al., who showed a benefit from alloSCT for inv(16) patients with poor MRD-response: 3Y-CIR 7.1% vs 87.7%, DFS 86.7% vs 12.3% and OS 93.3% vs 40.0%, all highly significant. Again, this could not be shown for good MRD-responders (Qin et al., 2015). For NPM1-positive patients, the ALFA group retrospectively demonstrated an added value of transplantation in patients with less than 4-log reduction in PB-MRD after induction (1 or 2 cycles), but not in those with > 4-log reduction. (Balsat et al., 2017). The only prospective study on MRD-directed treatment is the one by Zhu et al. in t(8;21)-AML, in which allogeneic SCT was offered to patients not reaching 3-log reduction in transcript levels after the second consolidation, or losing MRD-negativity within 6 months after the 2nd consolidation (defined as 'high risk') (Zhu et al., 2013). However, this study is also at high risk of bias because of a lack of randomization: MRD-directed treatment was proposed for all patients, but many patients did not follow the study proposal. Patients undergoing the proposed MRD-directed treatment

were compared with those who did not undergo the proposed treatment because of patient's choice. AlloSCT was shown to improve relapse risk, DFS and OS in high-risk patients (CIR 22.1% vs 78.9%; DFS 61.7% vs 19.6% and OS 71.6% vs 26.7%, all significant), but not for low-risk patients (Zhu et al., 2013). However, as all of the above studies were retrospective and/or non-randomized studies, with a high risk of bias, we need to conclude that there is still not enough evidence to support MRD-directed treatment in AML.

Another question that has not been addressed properly so far is the value of MRD-monitoring in follow-up post-therapy. In most AML subtypes, there is a delay of a few months between the molecular and haematological relapse. Theoretically, this offers a window to administer pre-emptive therapy. To date, there is little data as to which kind of treatment is preferred in case of molecular relapse (intensive rescue chemotherapy, hypomethylating agents, transplant, donor-lymphocyte infusion), and whether it could improve outcome. Sockel et al. reported on a small series of MRD-positive NPM1 + AML patients (n = 10) in 1st and 2nd CR after intensive chemotherapy with or without autologous or alloSCT, receiving pre-emptive therapy with azacytidine (Sockel et al., 2011). After a median follow-up time of 10 months, 70% of them were still in morphological CR (more than is expected from historical series) and had MRD-response (at least 1-log decrease), however only temporary for 4 of them. Moreover, Ragon et al. reported on 17 CBF-patients, who were MRD-positive at the start of maintenance with a hypomethylating agent (HMA) after intensive chemotherapy: 12/17 remained in complete morphological remission, with molecular response in 11/12 after 1 or 2 cycles (Ragon et al., 2017). In conclusion, there seems to be some kind of sensitivity of low-level MRD-positive good-risk AML to HMA therapy. However, there is no data on the effect of these kind of strategies on outcome, and randomized prospective trials are lacking.

8. Conclusion

In conclusion, there is currently sufficient evidence that MRD evaluation has an added value in prognostication in non-APL AML, and that it is more than just a surrogate marker for known risk factors. Different retrospective studies suggest that MRD might even overrule known pre-treatment outcome predictors such as mutational status: as such, the behavior of the disease during treatment might become equally or even more important to decide whether or not an upgrade of treatment (such as an alloSCT) is necessary to improve outcome. However, there are still many open issues as to what the ideal time point is to measure MRD, which threshold is clinically significant, what sample (peripheral blood or bone marrow) should be used and how we can standardize tests so that results from different labs become comparable. Moreover, survival benefit of MRD-directed post-induction treatment strategies needs to be confirmed in prospective randomized controlled trials, before MRD can be incorporated in clinical guidelines. The same is true for pre-emptive therapy in molecular relapse. And of course, all taking into account cost-effectiveness of these strategies. We expect MRD to be a major contributor in the process of clinical decision-making for individual AML patients with regard to treatment and follow-up in the (near) future.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Referee suggestions

Gert Ossenkoppele, VUMC, De Boelelaan 1117, 1081 Amsterdam, G.Ossenkoppele@vumc.nl.

Peter Valk, Erasmus MC, 's-Gravensdijkwal 230, 3015 CE Rotterdam, P.Valk@erasmusmc.nl, Adam Ivey, Alfred Hospital, Melbourne, Australia

Conflicts of interest

None.

Acknowlegdements

On behalf of the Acute Leukemias Committee of the Belgian Hematological Society (BHS): S. Anguille, V. Beckers, Z. Berneman, S. Buntinx, A. De Becker, C. Graux, V. Havelange, J. Kerger, F. Lambert, W. Lee, P. Lewalle, J. Maertens, T. Mercier, V. Robin, K. Theunissen, A. Van De Velde, F. Van Obbergh, M. Vekemans, S. Wittnebel.

References

Acuna-Hidalgo, R., Sengul, H., Steehouwer, M., et al., 2017. Ultra-sensitive sequencing identifies high prevalence of clonal hematopoiesis-associated mutations throughout adult life. Am. J. Hum. Genet. 101 (1), 50–64.

Balsat, M., Renneville, A., Thomas, X., et al., 2017. Postinduction minimal residual disease predicts outcome and benefit from allogeneic log in acute myeloid leukemia with NPM1 mutation: a study by the Acute Leukemia French Association Group. J. Clin. Oncol. 35 (2), 185–193.

Benton, C.B., Ravandi, F., 2017. A mind map for managing minimal residual disease in acute myeloid leukemia. Clin. Adv. Hematol. Oncol. 15 (11), 859–867.

Buccisano, F., Maurillo, L., Spagnoli, A., et al., 2010. Cytogenetic and molecular diagnostic characterization combined to postconsolidation minimal residual disease assessment by flow cytometry improves risk stratification in adult acute myeloid leukemia. Blood 116 (13), 2295–2303.

Corbacioglu, A., Scholl, C., Schlenk, R.F., et al., 2010. Prognostic impact of minimal residual diseas in CBFB-MYH11-positive acute myeloid leukemia. J. Clin. Oncol. 28 (23), 3724–3729.

Cruz, N.M., Mencia-Trinchant, N., Hassane, D.C., et al., 2017. Minimal residual disease in acute myelogenous leukemia. Int. J. Lab. Hematol. 39 (S1), 53–60.

Freeman, S.D., Virgo, P., Couzens, S., et al., 2013. Prognostic relevance of treatment response measured by flow cytometric residual disease detection in older patients with acute myeloid leukemia. J. Clin. Oncol. 31 (32), 4123–4132.

Hanekamp, D., Denys, B., Kaspers, G.J.L., et al., 2017. Leukaemic stem cell load at diagnosis predicts the development of relapse in young acute myeloid leukaemia patients. Br J Haematol. Wiley-Blackwell.

Hirsch, P., Tang, R., Abermil, N., et al., 2017. Precision and prognostic value of clone-specific minimal residual disease in acute myeloid leukemia. Haematologica 102 (7), 1227–1237.

Hubmann, M., Köhnke, T., Hoster, E., et al., 2014. Molecular response assessment by quantitative real-time polymerase chain reaction after induction therapy in NPM1-mutated patients identifies those at high risk of relapse. Haematologica 99 (8), 1317–1325.

Ivey, A., Hills, R.K., Simpson, M.A., et al., 2016. Assessment of minimal residual disease in standard-risk AML. N. Engl. J. Med. 374 (5), 422–433.

Jongen-Lavrencic, M., Grob, T., Hanekamp, D., et al., 2018. Molecular minimal residual disease in acute myeloid leukemia. N. Engl. J. Med. 378 (13), 1189–1245.

Jourdan, E., Boissel, N., Chevret, S., et al., 2013. Prospective evaluation of gene mutations and minimal residual diseas in patients with core binding factor acute myeloid leukemia. Blood 121 (12), 2213–2223.

Klco, J.M., Miller, C.A., Griffith, M., et al., 2015. Association between mutation clearance after induction therapy and outcomes in acute myeloid leukemia. JAMA 314 (8), 811–822.

Köhnke, T., Sauter, D., Ringel, K., et al., 2015. Early assessment of minimal residual disease in AML by flow cytometry during aplasia identifies patients at increase risk of relapse. Leukemia 29, 377–386.

Krönke, J., Schlenk, Rf., Jensen, K., et al., 2011. Monitoring of minimal residual disease in NPM1-mutated acute myeloid leukemia: a study from the German-Austrian Acute Myeloid Leukemia Study Group. J. Clin. Oncol. 29 (19), 2709–2716.

Krönke, J., Bullinger, L., Teleanu, V., et al., 2013. Clonal evolution in relapsed NPM1-mutated acute myeloid leukemia. Blood 122 (1), 100–108.

Lambert, J., Lambert, J., Nibourel, O., et al., 2014. MRD assessed wy WT1 and NPM1 transcript levels identifies distinct outcomes in AML patients and is influenced by gemtuzumab ozogamicin. Oncotarget 5 (15), 6280–6288.

Lane, S., Saal, R., Mollee, P., et al., 2009. A ≥ 1 log rise in RQ-PCR transcript levels defines molecular relapse in core binding factor acute myeloid leukemia and predicts subsequent morphologic relapse. Leuk. Lymphoma 49 (3), 517–523.

Ley, T.J., Mardis, E.R., Ding, L., et al., 2008. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. Nature 456 (7218), 66–72.

Löwenberg, B., Pabst, T., Maertens, J., 2017. Therapeutic value of clofarabine in younger and middle-aged (18-65 years) adults with newly diagnosed AML. Blood 129, 1636–1645.

Malmberg, E.B., Stahlman, S., Rehammar, A., et al., 2017. Patient-tailored analysis of minimal residual disease in acute myeloid leukemia using next-generation sequencing. Eur. J. Haematol. 98 (1), 26–37.

Maurillo, L., Buccisano, F., Spagnoli, A., et al., 2007. Monitoring of minimal residual diseas in adult acute myeloid leukemia using peripheral blood as an alternative

source to bone marrow. Haematologica 92 (5), 605–611.
Mosna, F., Gottardi, M., 2016. Stem cell modeling of core binding factor acute myeloid leukemia. Stem Cells Int. 2016, 7625827. https://doi.org/10.1155/2016/7625827. Epub 2016 Jan 13.
Perea, G., Lasa, A., Aventin, A., et al., 2006. Prognostic value of minimal residual disease (MRD) in acute myeloid leukemia (AML) with favorable cytogenetics [t(8;21) and inv (16)]. Leukemia 20, 87–94.
Pigazzi, M., Manara, E., Buldini, B., et al., 2015. Minimal residual disease monitored after induction therapy by RQ-PCR can contribute to tailor treatment of patients with t (8;21) RUNX1-RUNX1T1 rearrangement. Haematologica 100, e99.
Qin, Y., Xu, L., Chen, H., et al., 2015. Allogeneic stem cell transplant may improve the outcome of adult patients with inv(16) acute myeloid leukemia in first complete remission with poor molecular responses to chemotherapy. Leuk. Lymphoma 56 (11), 3116–3123.
Ragon, B.K., Daver, N., Garcia-Manero, G., et al., 2017. Minimal residual disease eradication with epigenetic therapy in core binding factor acute myeloid leukemia. Am. J. Hematol. 92, 845–850.
Schnittger, S., Kern, W., Tschulik, C., et al., 2009. Minimal residual diseas levels assessed by NPM1 mutation-specific RQ-PCR provide important prognostic information in AML. Blood 114 (11), 2220–2231.
Schuurhuis, Gj., Heuser, M., Freeman, S., et al., 2018. Minimal/measurable disease in AML: consensus document from ELN MRD working party. Blood 131 (12), 1275–1291.
Sockel, K., Wermke, M., Radke, J., et al., 2011. Minimal residual disease-directed preemptive treatment with azacitidine in patients with NPM1-mutant acute myeloid leukemia and molecular relapse. Haematologica 96 (10), 1568–1570.
Terwijn, M., van Putten, W.L.J., Kelder, A., et al., 2013. High prognostic impact of flow cytometric minimal residual disease detection in acute myeloid leukemia: data from the HOVON/SAKK AML42A study. J. Clin. Oncol. 31 (31), 3889–3897.
Willekens, C., Blanchet, O., Renneville, A., et al., 2016. Prospective long-term minimal residual disease monitoring using RQ-PCR in RUNX1-RUNX1T1-positive acute myeloid leukemia: results of the French CBF-2006 trial. Haematologica 101 (3), 328–335.
Wong, T.N., Miller, C.A., Klco, J.M., et al., 2016. Rapid expansion of preexisting nonleukemic hematopoietic clones frequently follows induction therapy for de novo AML. Blood 127 (7), 893–897.
Yin, J.A.L., O'Brien, M.A., Hills, R.K., et al., 2012. Minimal residual disease monitoring by quantitative RT-PCR in core binding factor AML allows risk stratification and predicts relapse: results of the United Kingdom MRC AML-15 trial. Blood 120 (14), 2826–2835.
Young, A.L., Challen, G.A., Birmann, B.M., et al., 2016. Clonal haematopoiesis harbouring AML-associated mutations is ubiquitous in healthy adults. Nat. Commun. 7 (12484).
Zeijlemaker, W., Kelder, A., Oussoren-Brockhoff, Y.J., et al., 2016. Peripheral blood minimal residual disease may replace bone marrow minimal residual disease as an immunophenotypic biomarker for impending relapse in acute myeloid leukemia. Leukemia 30 (3), 708–715.
Zeijlemaker, W., Meijer, R., Kelder, A., et al., 2017. Leukemic stem cell frequency combined with MRD is an important biomarker to predict relapse in acute myeloid leukemia. Results from a prospective H102 study. Abstract at EHA, Presented 23rd June 2017.
Zhang, L., Li, Q., Li, W., et al., 2013. Monitoring of minimal residual disease in acute myeloid leukemia with t(8;21)(q22;q22). Int. J. Hematol. 97, 786–792.
Zhu, H., Zhang, X., Qin, Y., et al., 2013. MRD-directed risk stratification treatment may improve outcomes of t(8;21) AML in first complete remission: results from the AML05 multicenter trial. Blood 121 (20), 4056–4062.