Ontario Health
Cancer Care Ontario

Guideline MOTAC-6 Version 2

A Quality Initiative of the Program in Evidence-Based Care (PEBC), Ontario Health (Cancer Care Ontario)

Minimal Residual Disease Testing in Acute Leukemia

M. Sabloff, H. Feilotter, D. Sivajohanathan, C. Howlett, C. Ross, A. Schuh, A. Pollett and the Minimal Residual Disease Testing Guideline Development Group

An assessment conducted in November 2023 deferred the review of Guideline MOTAC-6. This means that the document remains current until it is assessed again next year. The PEBC has a formal and standardized process to ensure the currency of each document (PEBC Assessment & Review Protocol)

GL MOTAC-6 is comprised of 6 sections. You can access the summary and full report here:
https://www.cancercareontario.ca/en/guidelines-advice/types-of-cancer/63341

September 27, 2022

For information about this document, please contact Dr. Mitchell Sabloff or Dr. Harriet Feilotter, the lead authors, through the PEBC via:
Phone: 905-527-4322 ext. 42822 Fax: 905 526-6775 E-mail: ccopgi@mcmaster.ca

For information about the PEBC and the most current version of all reports, please visit the Ontario Health (CCO) website at http: https://www.cancercareontario.ca/en/guidelines-advice or contact the PEBC office at:
Phone: 905-527-4322 ext. 42822 Fax: 905 526-6775 E-mail: ccopgi@mcmaster.ca

PEBC Report Citation (Vancouver Style): Sabloff M, Feilotter H, Sivajohanathan D, Howlett C, Ross C, Schuh A, et al. Minimal residual disease testing in acute leukemia. Sabloff M, Feilotter H, Sivajohanathan D, reviewers. Toronto (ON): Ontario Health (Cancer Care Ontario); 2020 Mar 10; Endorsed 2022 Sep 27. Program in Evidence-Based Care Guideline No.: MOTAC-6 Version 2 ENDORSED.

Guideline Report History

GUIDELINE VERSION	SYSTEMATIC REVIEW		PUBLICATIONS	NOTES AND KEY CHANGES
	Search Dates	Data		
Original version March 10, 2020	2013 to 2019	Full Report	Web publication	NA
Current Version 2 September 27, 2022	2019 to 2022	New data found in Section 6: Document Assessment and Review	Updated Web publication	2020 recommendations are **ENDORSED**

Table of Contents

Minimal Residual Disease Testing in Acute Leukemia

Section 1: Recommendations

This section is a quick reference guide and provides the guideline recommendations only. For key evidence associated with each recommendation, see Section 2.

GUIDELINE OBJECTIVES

- To provide evidence surrounding the clinical utility[1] of minimal/measurable residual disease (MRD) testing using multiparameter flow cytometry (MFC), next-generation sequencing (NGS), or polymerase chain reaction (PCR)-based methods in patients with acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML).

TARGET POPULATION

Adult patients with a diagnosis of acute leukemia (i.e., AML or ALL).

INTENDED USERS

This guideline is targeted for:

- Clinicians, laboratory physicians, and scientists involved in the care and testing of patients with acute leukemia.
- Policy makers, health care administrators, and the Ontario Ministry of Health.

PREAMBLE

MRD testing refers to the evaluation of very small amounts of a biomarker(s) that signals the presence of residual disease beyond that detectable by conventional, less-sensitive testing methods. In acute leukemias, the use of MRD testing using bone marrow or blood is routine, as these assays provide prognostic information for clinicians and patients. Indeed, the most recent guidelines from international expert panels, the European LeukemiaNet [1] and the National Comprehensive Cancer Network [2], agree that MRD assessment in all ALL and AML is important to provide the most comprehensive prognostic information in order to be able to discuss the disease with patients and families.

What is not as well established is the ability of MRD testing to behave as a predictive factor to provide information that the course of treatment can be confidently escalated or de-escalated to provide the highest survival while minimizing the morbidity and mortality in appropriate cases. This ability has not yet been clearly demonstrated, particularly in the adult acute leukemias. Therefore, this review seeks to examine and collate the evidence that would allow expansion of the role of MRD testing in the adult population from a prognostic test to one that guides treatment.

This review does not evaluate optimal methods for MRD detection, nor does it recommend specific markers for testing. Both the testing modality and the set of biomarkers that could be assessed are rapidly evolving fields, and decisions about these practical aspects of MRD testing will require constant review.

[1] Clinical utility refers to the ability of the test to provide information that is useful to direct treatment and ultimately improve patient outcome. This is in contrast to prognostic utility (or prognosis), which gives information about likely survival time but does not address whether or not a treatment would be beneficial.

RECOMMENDATIONS

Recommendation 1
MRD testing may be considered as an aid to help select between various treatment options in adult patients with ALL (i.e., adjustment of treatment intensity/interventional treatment stratification), in addition to its prognostic capabilities.
Qualifying Statements
• The timing of MRD testing is variable and depends, in part, upon the treatment regimen and the type of test being used. • Although MRD can be measured in either bone marrow or peripheral blood, bone marrow is recommended as there may be discordance between the two measurements, potentially underestimating the disease burden if only blood is monitored. • While the prognostic information from MRD testing is accepted in clinical practice and used to inform patients of their prognosis, the ability of MRD testing results to be predictive of adult ALL patient response to different treatment options (escalation or de-escalation) is not yet established • While the adjustment of treatment intensity based on MRD testing results has been tested in the pediatric population with convincing results, similar analysis has not been conducted in the adult population. Confirmatory studies are needed in the adult literature. • Decisions to reduce or increase the intensity of treatment, based on the MRD results, should be restricted to clinical trials, if available, as a positive outcome has not yet been confirmed in adults with ALL. Ongoing trials with a focus on adjusting the intensity of therapy with novel agents that add efficacy with limited toxicity are needed to clarify whether this approach will improve survival. • MRD may be required for eligibility for specific therapy, in which case testing should be done when consideration is made for therapy as per treatment guidelines. • MRD testing should be conducted using clinically validated tests with suitable sensitivity and specificity metrics and clinically accepted thresholds to define MRD.

Recommendation 2
There is currently insufficient evidence for or against the use of MRD testing to guide the choice between various treatment options in adult patients with AML (i.e., adjustment of treatment intensity/interventional treatment stratification) outside of its pre-defined prognostic capabilities.
Qualifying Statement
The prognostic information from MRD testing in adult patients with AML is well understood and can be used to inform patients of their prognosis.

Minimal Residual Disease Testing in Acute Leukemia

Section 2: Guideline – Recommendations and Key Evidence

GUIDELINE OBJECTIVES

- To provide evidence surrounding the clinical utility[2] of minimal/measurable residual disease (MRD) testing using multiparameter flow cytometry (MFC), next-generation sequencing (NGS), or polymerase chain reaction (PCR)-based methods in patients with acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML).

TARGET POPULATION

Adult patients with a diagnosis of acute leukemia (i.e., AML or ALL).

INTENDED USERS

This guideline is targeted for:

- Clinicians, laboratory physicians, and scientists involved in the care and testing of patients with acute leukemia.
- Policy makers, health care administrators, and the Ontario Ministry of Health.

PREAMBLE

MRD testing refers to the evaluation of very small amounts of a biomarker(s) that signals the presence of residual disease beyond that detectable by conventional, less-sensitive testing methods. In acute leukemias, the use of MRD testing using bone marrow or blood is routine, as these assays provide prognostic information for clinicians and patients. Indeed, the most recent guidelines from international expert panels, the European LeukemiaNet (ELN) [1] and the National Comprehensive Cancer Network [2], agree that MRD assessment in all ALL and AML is important to provide the most comprehensive prognostic information in order to be able to discuss the disease with patients and families.

What is not as well established is the ability of MRD testing to behave as a predictive factor to provide information that the course of treatment can be confidently escalated or de-escalated to provide the highest survival while minimizing the morbidity and mortality in appropriate cases. This ability has not yet been clearly demonstrated, particularly in the adult acute leukemias. Therefore, this review seeks to examine and collate the evidence that would allow expansion of the role of MRD testing in the adult population from a prognostic test to one that guides treatment.

This review does not evaluate optimal methods for MRD detection, nor does it recommend specific markers for testing. Both the testing modality and the set of biomarkers that could be assessed are rapidly evolving fields, and decisions about these practical aspects of MRD testing will require constant review.

[2] Clinical utility refers to the ability of the test to provide information that is useful to direct treatment and ultimately improve patient outcome. This is in contrast to prognostic utility (or prognosis), which gives information about likely survival time but does not address whether or not a treatment would be beneficial.

RECOMMENDATIONS, KEY EVIDENCE, AND INTERPRETATION OF EVIDENCE

Recommendation 1

MRD testing may be considered as an aid to help select between various treatment options in adult patients with ALL (i.e., adjustment of treatment intensity/interventional treatment stratification), in addition to its prognostic capabilities.

Qualifying Statements for Recommendation 1

- The timing of MRD testing is variable and depends, in part, upon the treatment regimen and the type of test being used.
- Although MRD can be measured in either bone marrow or peripheral blood, bone marrow is recommended as there may be discordance between the two measurements, potentially underestimating the disease burden if only blood is monitored.
- While the prognostic information from MRD testing is accepted in clinical practice and used to inform patients of their prognosis, the ability of MRD testing results to be predictive of adult ALL patient response to different treatment options (escalation or de-escalation) is not yet established
- While the adjustment of treatment intensity based on MRD testing results has been tested in the pediatric population with convincing results, similar analysis has not been conducted in the adult population. Confirmatory studies are needed in the adult literature.
- Decisions to reduce or increase the intensity of treatment, based on the MRD results, should be restricted to clinical trials, if available, as a positive outcome has not yet been confirmed in adults with ALL. Ongoing trials with a focus on adjusting the intensity of therapy with novel agents that add efficacy with limited toxicity are needed to clarify whether this approach will improve survival.
- MRD may be required for eligibility for specific therapy, in which case testing should be done when consideration is made for therapy as per treatment guidelines.
- MRD testing should be conducted using clinically validated tests with suitable sensitivity and specificity metrics and clinically accepted thresholds to define MRD.

Key Evidence for Recommendation 1

Two randomized controlled trials (RCTs) (UKALL 2003 and AEIOPP-BFM ALL 2000) were found in the pediatric literature [3-7] that provided evidence around the clinical utility of MRD testing. Additionally, one retrospective study in the adult population [8] was found that conducted subgroup analyses to determine whether alternate therapies improved survival based on MRD levels. The overall certainty of this evidence is moderate, being rated down for indirectness.

The UKALL 2003 trial [3-5] tested whether adjustment of treatment intensity according to MRD risk stratification was feasible in pediatric patients using standardized real-time quantitative PCR (qPCR) for immunoglobulin and T-cell receptor gene rearrangements with a quantitative range of 10^{-4}.

- In patients who were stratified as MRD low risk (<0.01%) [3], no significant differences in five-year event-free survival (EFS) were reported between those who received reduced therapy (n=261) compared with those who received standard therapy (n=260; odds ratio [OR] 1.00; 95% confidence interval [CI] 0.43 to 2.31; p=0.99). Similarly, there was no difference in five-year overall survival [OS] (OR 0.67; 95% CI 0.19 to 2.30; p=0.53).
- No significant difference was found in the number of grade 3 to 4 toxic events (p=0.30), serious adverse events (p=0.26), or treatment-related deaths (p=0.08) between the two treatment arms.

- In patients who were stratified as MRD high risk (≥ 0.01%) [4], there was a statistically significant difference in five-year EFS between those who received augmented therapy (n=267; 89.6%; 95% CI 85.9 to 93.3) and those who received standard therapy (n=266; 82.8%; 95% CI 78.1 to 87.5) (OR 0.61; 95% CI 0.39 to 0.98; p=0.04). There was no difference in five-year OS (OR 0.67; 95% CI 0.38 to 1.17; p=0.16).
- No significant difference was reported for grade 3 to 4 adverse events (p=0.55). Patients who received augmented therapy experienced significantly more adverse events than those who received standard therapy (p=0.02).
- There were no statistically significant effects of randomization on health-related quality of life or parental care-giving burden within the MRD low-risk and MRD high-risk groups [5].

The AEIOPP-BFM ALL 2000 trial [6,7] also tested MRD-directed treatment augmentation and reduction to prove non-inferiority in pediatric patients with Philadelphia chromosome (Ph)-negative ALL with qPCR for immunoglobulin and T-cell receptor gene rearrangements with a quantitative range of at least 10^{-4} and a sensitivity of at least 10^{-4}. The results for the MRD intermediate-risk and MRD high-risk groups are not yet available.

- In patients determined by MRD to be at standard risk (i.e., MRD was negative on days 33 and 78 with at least two markers with a sensitivity of 1×10^{-4}), no significant differences were reported in eight-year OS between those who received the standard delayed intensification (n=583; 98.9% ± 0.6%) and those that received the reduced-intensity treatment (n=581; 96.1% ± 0.8%| hazard ratio [HR], 2.00; 95% CI 0.97 to 4.13; p=0.055.
- Non-life-threatening, life-threatening, and fatal adverse events were comparable in patients at standard risk who received either reduced intensity or standard delayed intensity treatments. P-values were not reported.
- The AEIOPP-BFM ALL 2000 has not provided any quality of life data to date.

In the adult population, one retrospective study [8] was identified that conducted a subgroup analysis of the GRAALL-2003 and GRAALL-2005 trials to determine whether alternate therapies improved survival based on MRD levels measured using real-time qPCR for immunoglobulin and T-cell receptor gene rearrangements with a quantitative range of 10^{-4} and a sensitivity of at least 10^{-4}.

- A subgroup analysis of 522 patients at high risk with Ph-negative ALL who were candidates for stem cell transplant at first complete remission was conducted. In a multivariable analysis of these patients, the significant interaction between MRD response and stem cell transplant effect was observed after adjustment of age, white blood cell count, and resistance to steroid prophase ($p_{interaction}$=0.002). For patients with MRD ≥10^{-3} who received a transplant, there was a significant survival benefit compared with those who did not receive a transplant, (HR, 0.41; 95% CI 0.22 to 0.76; p=0.005). For patients with MRD <10^{-3}, there was no difference in survival benefit whether or not they received a transplant (p=0.17). Other reviews and trials also support the use of MRD in patient selection for stem cell transplant [9-11].

Interpretation of Evidence for Recommendation 1

- The members of the Working Group considered OS to be a critical outcome and EFS, relapse, adverse events, and quality of life to be important outcomes. The Working Group was unanimous in their opinion that patients value survival benefit in addition to other outcomes, such as quality of life and adverse events, although patient input was not sought.
- For pediatric patients with ALL achieving an interim MRD below the target, a treatment reduction may be acceptable due to no reduction in OS. The Working Group

members determined the benefits outweighed the harms due to the value patients would place on fewer hospitals visits and a reduction in the number of chemotherapy rounds.
- For pediatric patients with ALL receiving treatment augmentation due to the failure to achieve an interim MRD below the target, the Working Group members determined the benefits outweighed the harms (i.e., increased EFS but not OS), due to its acceptability by patients and providers.
- The generalizability of this pediatric evidence to the adult population is not yet verified. However, the Working Group recognizes that the pediatric literature is more robust than the adult literature for ALL and pediatric protocols are often used to inform the treatment of adults.

Recommendation 2

There is currently insufficient evidence for or against the use of MRD testing to guide the choice between various treatment options in adult patients with AML (i.e., adjustment of treatment intensity/interventional treatment stratification) outside of its pre-defined prognostic capabilities.

Qualifying Statements for Recommendation 2

The prognostic information from MRD testing in adult patients with AML is well understood and can be used to inform patients of their prognosis.

Key Evidence for Recommendation 2

One study [12] was found that provided evidence around the prognostic clinical utility of MRD testing in adult patients with AML. Additionally, four studies [13-16] were found that conducted subgroup analyses to determine whether alternate therapies improved survival based on MRD levels. The overall certainty of this evidence is low, being rated down for inconsistency and imprecision.

The BMT CTN 0901 trial [12] randomized adult patients with AML in morphologic complete remission undergoing allogeneic hematopoietic cell transplantation (alloHCT) to either myeloablative conditioning (MAC) or reduced intensity conditioning (RIC). Retrospective non-protocol subgroup analysis by MRD status was conducted; MRD was determined using stored frozen blood samples collected during remission (prior to transplant) by a multiplex PCR-based panel including 13 commonly mutated genes with a limit of detection set at an allele frequency of 0.001. The authors reported significant differences in three-year OS according to conditioning intensity in patients with detectable mutations (MRD-positive; OS 61% MAC vs. 43% RIC, p=0.02) but not in patients who tested negative for MRD (OS 56% MAC vs 63% RIC, p=0.96).
Four studies were identified in which multivariable analyses were conducted to determine whether alternate therapies improved survival based on MRD levels.

Balsat et al [13] reported on a subgroup analysis of 64 patients with non-favourable AML from the ALFA-0702 trial that were eligible for alloHCT in first remission. Quantification of types A, B, and D nucleophosmin 1 mutations (*NPM*1m) transcript levels was performed using reverse transcriptase qPCR (RT-qPCR) with a detection limit of 0.01%.
- OS was significantly improved by alloHCT in those with a <4-log reduction in *NPM1* mutated (*NPM1*m) peripheral blood MRD (HR, 0.25; 95 % CI 0.06 to 0.98; p=0.047) while this was not observed in patients with a >4-log reduction (HR, 2.11; 95% CI 0.57 to 7.71; p=0.261). The interaction between alloHCT effect and MRD log reduction was significant (p=0.027).

Freeman et al [14] conducted a subgroup analysis of 204 *NPM*1-wild type (wt) patients with AML or high-risk myelodysplastic syndrome from the NCRI AML17 trial to determine the effect of alloHCT in first complete response according to MRD status. MFC, which screened for leukemia-associated immunophenotypes, was used at a sensitivity of 0.02% to 0.05% for pre-treatment blasts and a sensitivity of 0.05% to 0.1% on follow-up blast samples.

- Survival was not significantly improved in MRD-positive patients (HR, 0.72; 95% CI 0.31 to 1.69) or MRD-negative patients (HR, 1.68; 95% CI 0.75 to 3.85; $p_{interaction}$=0.16) of the *NPM1*-wt standard-risk group.

In the prospective cohort study by Jongen-Lavrencic et al. [15] 430 patients with newly diagnosed AML were evaluated to investigate whether molecular monitoring with NGS could predict recurrence.

- In a multivariable analysis, the interaction between the detection of residual disease and type of consolidation therapy (i.e., no therapy, chemotherapy, or autologous or allogeneic HSCT) was not significant for relapse or survival. The p-value was not reported.

The retrospective study by Chen et al [16] studied 245 patients with newly diagnosed, relapsed or refractory AML. MRD was measured using 10-colour flow cytometry. Further details regarding sensitivity and/or detection limit were not reported.

- A multivariable analysis determined that MRD is not predictive of a differential treatment effect by HCT for relapse, OS, or relapse-free survival (RFS) (interaction, p-value was not reported).

Interpretation of Evidence for Recommendation 2

- The Working Group members determined evidence from an abstract of an RCT and small multivariable analyses are insufficient to make definitive recommendations about the clinical utility of MRD testing in patients with AML at this time.

IMPLEMENTATION CONSIDERATIONS

The Complex Malignant Hematology program will address issues related to the implementation of MRD testing with their stakeholders.

Several barriers to implementation were identified in the external review of this guideline by target users in Ontario. They identified financial impact, staffing, access to validated testing, and awareness in the community as barriers. It was also noted that not all centres have the resources and expertise to provide these services and that, in practice, MRD testing is moving ahead of the published evidence. To address these issues, technology to support ongoing clinical evaluation and development of standardized structured synoptic reporting were suggested.

Minimal Residual Disease Testing in Acute Leukemia

Section 3: Guideline Methods Overview

This section summarizes the methods used to create the guideline. For the systematic review, see Section 4.

THE PROGRAM IN EVIDENCE-BASED CARE

The Program in Evidence-Based Care (PEBC) is an initiative of the Ontario provincial cancer system, Ontario Health (Cancer Care Ontario) (OH (CCO)). The PEBC mandate is to improve the lives of Ontarians affected by cancer through the development, dissemination, and evaluation of evidence-based products designed to facilitate clinical, planning, and policy decisions about cancer control.

The PEBC supports the work of Guideline Development Groups (GDGs) in the development of various PEBC products. The GDGs are composed of clinicians, other healthcare providers and decision makers, methodologists, and community representatives from across the province.

The PEBC is a provincial initiative of OH (CCO) supported by the Ontario Ministry of Health (MOH). All work produced by the PEBC is editorially independent from the MOH.

BACKGROUND FOR GUIDELINE

The Complex Malignant Hematology initiative of OH (CCO) determined that evidence-based recommendations were needed for the role of MRD testing in ALL and AML.

GUIDELINE DEVELOPERS

This guideline was developed by the Minimal Residual Disease Testing GDG (Appendix 1), which was convened at the request of the Molecular Oncology Testing Advisory Committee (MOTAC).

The project was led by a small Working Group of the MRD Testing GDG, which was responsible for reviewing the evidence base, drafting the guideline recommendations, and responding to comments received during the document review process. The Working Group had expertise in anatomical pathology, hematology, hematopathology, molecular genetics, and health research methodology. Other members of the MRD Testing GDG served as the Expert Panel and were responsible for the review and approval of the draft document produced by the Working Group. Conflict of interest declarations for all GDG members are summarized in Appendix 1, and were managed in accordance with the *PEBC Conflict of Interest Policy*.

GUIDELINE DEVELOPMENT METHODS

The PEBC produces evidence-based and evidence-informed guidance documents using the methods of the Practice Guidelines Development Cycle [17,18]. This process includes a systematic review, interpretation of the evidence by the Working Group and draft recommendations, internal review by content and methodology experts, and external review by Ontario clinicians and other stakeholders.

The PEBC uses the AGREE II framework [19] as a methodological strategy for guideline development. AGREE II is a 23-item validated tool that is designed to assess the methodological rigour and transparency of guideline development.

The currency of each document is ensured through periodic review and evaluation of the scientific literature and, where appropriate, the addition of newer literature to the original evidence base. This is described in the *PEBC Document Assessment and Review Protocol*. PEBC

guideline recommendations are based on clinical evidence, and not on feasibility of implementation; however, a list of implementation considerations such as costs, human resources, and unique requirements for special or disadvantaged populations is provided along with the recommendations for information purposes. PEBC guideline development methods are described in more detail in the *PEBC Handbook* and the *PEBC Methods Handbook*.

Search for Existing Guidelines

As a first step in developing this guideline, a search for existing guidelines was undertaken to determine whether an existing guideline could be adapted or endorsed. To this end, the following sources were searched for existing guidelines that addressed the research questions:

- Practice guideline databases: National Institute for Health and Care Excellence (NICE) Evidence Search; Canadian Partnership Against Cancer Database, and the Canadian Medical Association Infobase.
- Guideline developer websites: NICE, Scottish Intercollegiate Guidelines Network (SIGN), American Society of Clinical Oncology (ASCO), National Health and Medical Research Council Australia, and Cancer Council Australia.

The following criteria were used to search for and select potentially relevant guidelines:

- Guideline databases and websites were searched for guidelines on February 21, 2018 with the search term "minimal residual disease".
- Only guidelines published after 2015 (i.e., less than 3 years old) were considered to ensure currency.
- Guidelines based on consensus or expert opinion were excluded.

No guidelines met the inclusion criteria.

GUIDELINE REVIEW AND APPROVAL

Internal Review

For the guideline document to be approved, 75% of the content experts who comprise the GDG Expert Panel must cast a vote indicating whether or not they approve the document, or abstain from voting for a specified reason, and of those that vote, 75% must approve the document. In addition, the PEBC Report Approval Panel (RAP), a three-person panel with methodology expertise, must unanimously approve the document. The Expert Panel and RAP members may specify that approval is conditional, and that changes to the document are required. If substantial changes are subsequently made to the recommendations during external review, then the revised draft must be resubmitted for approval by RAP and the GDG Expert Panel.

External Review

Feedback on the approved draft guideline is obtained from content experts and the target users through two processes. Through the Targeted Peer Review, several individuals with content expertise are identified by the GDG and asked to review and provide feedback on the guideline document. Through Professional Consultation, relevant care providers and other potential users of the guideline are contacted and asked to provide feedback on the guideline recommendations through a brief online survey. This consultation is intended to facilitate the dissemination of the final guidance report to Ontario practitioners.

ACKNOWLEDGEMENTS

The MRD Testing GDG would like to thank the following individuals for their assistance in developing this report:

- Melissa Brouwers, Bill Evans, Jill Fulcher, Jennifer Hart, Sheila McNair, Emily Vella, Goran Klaric, Safia Mohamed, Jonathan Sussman, Graeme Quest for providing feedback on draft versions.
- Rebecca McClure for participating in the Working Group during project planning and early drafts of this work.
- Sitara Sharma for conducting a data audit.
- Glenn Fletcher for assisting with external review and document completion stages of the document.
- Sara Miller for copy editing.

Minimal Residual Disease Testing in Acute Leukemia

Section 4: Systematic Review

INTRODUCTION

Acute leukemia is a heterogenous disease, primarily of the bone marrow, resulting in an overproduction of immature blood cells. In AML, myeloblasts are overproduced, while in B-lymphoblastic leukemia/lymphoma (B-ALL) or T-lymphoblastic leukemia/lymphoma (T-ALL), lymphoblasts are overproduced. In all cases, this overproduction results in failure of normal hematopoiesis. In 2016, in Ontario, ALL was the most common type of leukemia diagnosed in young children. In adults, ALL was the least common type of leukemia diagnosed (incidence 1.4 per 100,000, all ages) and AML was the most common acute leukemia diagnosed (incidence 4.1 per 100,000, all ages) [20].

In the context of acute leukemia, MRD refers to a low level of disease that is only detectable using methods more sensitive than traditional morphologic assessment. Detection of MRD requires methods that can identify disease-specific markers, when these are present, at levels as low as 0.001% or 10^{-5} (i.e. considerably lower than <5% blasts in the bone marrow, the morphologic definition of complete remission) [21,22]. These methods, which currently include PCR, MFC, and NGS-based approaches, generally interrogate bone marrow samples to identify the presence of low levels of persistent or recurrent disease well before morphologic clues become evident. The quest for the detection of such minute levels of disease has been, in part, driven by the observation that many patients will achieve a morphologic remission after the initial chemotherapy, but only a fraction will remain in remission. The remainder will relapse, suggesting that smaller amounts of disease remain. However, this small amount of disease has been previously undetectable using traditional methods for assessing morphologic evidence of disease.

In ALL, two basic approaches to evaluate MRD have been used:

1. Multiparameter flow cytometry. This technique relies on characterization of tumour cell surface proteins that can be seen as more uniform on a clonal population of precursor B-cells than typically seen in normally developing B-cells. The sensitivity of this method can reach 0.01% (or 10^{-4}) [2].
2. Sequencing of immunoglobulin for B-ALL or T-cell receptor gene rearrangements for T-ALL. Because each B-cell or T-cell contains a unique rearrangement sequence from early in development, the unique clonal sequence of the neoplastic cells can be identified at the time of diagnosis and that sequence can subsequently be targeted during assessment for MRD. This approach can have a sensitivity of approximately 0.01% (or 10^{-4}).

In addition, some B-ALL or T-ALL may contain disease-defining translocations or other large variants that can be used to monitor MRD with good sensitivity, because these variants are thought to be present in every neoplastic cell of the clone and to remain stable throughout the course of disease (features required for reliable MRD evaluation). However, each of these variants (and any of their subvariations) typically require a variant-specific assay to be developed and maintained by the testing laboratories, which, aside from the p190 and p210 forms of breakpoint cluster region-abelson (BCR-ABL1), is rarely considered feasible.

In AML, the situation for MRD monitoring is more complex because myeloid cells do not have any universally available biomarkers.

1. Flow cytometry can be used in many cases to identify cells thought to be clonal based on evaluation of characteristic surface (or intracellular) proteins in a manner somewhat similar to ALL, again with a sensitivity of approximately 0.1% (or 10^{-3}) [22].

2. PCR or NGS techniques – these techniques rely on probes that are designed to identify certain abnormalities in the RNA or DNA of the abnormal cell. These techniques are quite sensitive, reaching detection levels of 0.01%. The use of these techniques for MRD monitoring in AML is challenging for two main reasons:
    a. There is a great diversity of DNA variants that have been identified in the clones at diagnosis with a variety of laboratory techniques currently being used to identify them.
    b. Very few AML clones have a variant in each neoplastic cell that remains stable throughout the course of disease, features that are generally considered necessary for reliable MRD monitoring. Variants that do meet this criterion include promyelocytic leukemia/retinoic acid receptor alpha (PML/RARA) and the other variants currently listed with the World Health Organization as "recurrent genetic abnormalities" [23]. As a group, these account for approximately 40% to 50% of AML cases [24]. In addition, as discussed above for the disease-defining variants in ALL, each of these variants (and any of their subvariations) typically require a variant-specific assay to be developed, and maintained, by the testing laboratories, which has led to MRD tests being routinely available for only a small number of the recurrent variants (mainly PML/RARA). Although a variety of small variants (single nucleotide variants and indels) in common "myeloid genes" can be identified in most AML clones at diagnosis, it is now known that most of these are not necessarily stable, such that they may not be present in all clonal cells and may appear or disappear over the course of disease as subclones emerge and regress, either naturally or as a consequence of therapy. As such, most of these variants are not suitable for MRD monitoring.

    A further challenge for "myeloid gene" variant monitoring for MRD in AML is that some of the variants identified in AML have also been identified in normal people with age-related clonal hematopoiesis [25,26]. Because of this, it is not always clear whether a variant identified while monitoring MRD is really indicative of the AML clone, or rather represents residual/new clonal hematopoiesis.

Despite these difficulties, the latest version of the ELN 2017 has divided complete remission for acute leukemia in adults into complete remission with and without MRD, based on the finding that MRD-negative patients have improved outcomes [1]. Based on prior data, however, it seems reasonable to suspect that monitoring MRD in all adult acute leukemias would be beneficial, and laboratories routinely receive requests for this type of testing from physicians treating these patients.

It needs to be emphasized that due to the technical challenges of MRD evaluation discussed above, laboratories will require substantial development/validation time for these assays, so it is important to have a clear picture of where the literature regarding clinical utility for adult leukemias is currently, as well as where it is likely headed in the near future.

Although the presence of MRD has prognostic value, there is a lack of evidence-based guidelines on their use in guiding treatment decisions in adult patients with acute leukemia. Coupled with a need to develop standards to ensure appropriate patients are deriving a benefit and that patients are being treated equitably across the province, the development of this guideline aims to evaluate whether MRD testing can further add to the clinical management of these patients, beyond its prognostic value. The current review is focused on this clinical utility of MRD testing and does not evaluate the best method of MRD testing to be used (i.e., PCR, MFC, droplet digital PCR, or NGS) nor which leukemia markers should be tested.

The Working Group of the MRD Testing Guideline Development Group developed this evidentiary base to inform recommendations as part of a clinical practice guideline. Based on the objectives of this guideline (Section 2), the Working Group members derived the research questions outlined below.

RESEARCH QUESTIONS

1. What benefit to clinical management does MRD testing contribute in the treatment of patients with ALL?
    a. If positive at a certain time point, does a change in treatment alter patient outcomes?
    b. If positive before transplant, does a delay in stem cell transplant and change in treatment, to achieve MRD negativity, alter patient outcomes?
    c. If positive before transplant and negativity is achieved, are patient outcomes different?

2. What benefit to clinical management does MRD testing contribute in the treatment of patients with AML?
    a. If positive at a certain time point, does change in treatment alter patient outcomes?
    b. If positive before transplant, does a delay in stem cell transplant and change in treatment, to achieve MRD negativity, alter patient outcomes?
    c. If positive before transplant and negativity is achieved, are patient outcomes different?

METHODS

This evidence review was conducted in two planned stages, including a search for systematic reviews followed by a search for primary literature. These stages are described in subsequent sections.

Search for Existing Systematic Reviews

A search was conducted for existing systematic reviews. This included original systematic reviews and systematic reviews published as a component of practice guidelines. The MEDLINE (2013 to October 9, 2019) and EMBASE (2013 to October 9, 2019) databases, as well as the Cochrane Database of Systematic Reviews (2013 to October 9, 2019) were searched. The full search strategy is available in Appendix 2. Systematic reviews were included if they met the following criteria:

1. The review addressed at least one research question with similar inclusion/exclusion criteria, and
2. The review comprehensively searched at least one database with the literature search date and search terms included, and
3. The review included an assessment of the quality of the evidence, and
4. The review extracted relevant information from each study, and
5. The review analyzed the data appropriately.

Search for Primary Literature

A search for primary literature was conducted to locate literature where no existing systematic reviews were found.

Acute Lymphoblastic Leukemia

In the absence of relevant, comprehensive systematic reviews in the adult population, the search for primary literature in adults was conducted as planned. A search of the pediatric literature was also conducted for primary literature limited to studies published after 2014 due to the presence of Health Quality Ontario's Health Technology Assessment of MRD testing in childhood ALL [27].

Acute Myeloid Leukemia

In the absence of relevant, comprehensive systematic reviews, a search for primary literature was conducted as planned for primary literature in the adult population.

Literature Search Strategy

The MEDLINE (2000 to October 9, 2019) and EMBASE (2000 to October 9, 2019) databases were searched for RCTs. If no RCTs were found then the databases were searched for non-randomized comparative studies. If no non-randomized comparative studies were found then the databases were searched for single-arm studies with multivariable analyses. The full search strategy is available in Appendix 2. Reference lists of included primary literature were scanned for additional citations. The following conference proceedings were also searched from 2015 to 2018: Summit of American Society of Hematology, Congress of European Hematology Association, and Society of Hematopathology.

Study Selection Criteria and Process

Inclusion Criteria

- RCTs (if no RCTs then non-randomized comparative studies) with ≥30 participants and if no RCTs or non-randomized comparative studies then single-arm studies with ≥100 participants where confounders are controlled for; and
- Studies assessing adult patients with a diagnosis of acute leukemia (i.e., AML or ALL), if no relevant studies in the adult population are found then the pediatric literature is to be searched; and
- Studies using MRD testing and reporting the following clinical outcomes: OS, EFS, relapse rate, adverse events, and quality of life.

Exclusion Criteria

- Abstracts of non-randomized studies (single-arm clinical trials, case series, etc.); or
- Abstracts of interim analyses; or
- Papers or abstracts not available in English; or
- Letters and editorials that reported clinical trial outcomes; or
- Papers and abstracts published before 2000.

A review of the titles and abstracts that resulted from the search was conducted by one reviewer (DS). For items that warranted full-text review, one reviewer (DS) reviewed each item.

Data Extraction and Assessment of Study Quality and Potential for Bias

All included primary studies underwent data extraction by one reviewer (DS), with all extracted data and information audited subsequently by an independent auditor (SS). Ratios, including HRs, were expressed with a ratio of <1.0 indicating benefit for the experimental group for a given outcome. Important quality features, such as generation of allocation sequence, allocation concealment, blinding, intention-to-treat analysis, withdrawals, loss to follow-up, funding source, statistical power calculations, length of follow-up, differences in baseline patient characteristics, and early termination, were extracted for each study. Risk of bias was

assessed for each included trial using Cochrane's Risk of Bias tool, http://handbook.cochrane.org/ (Part 2, Section 8.5). Criteria from the Cochrane Risk of Bias for Non-randomized Studies of Interventions (ROBINS-I) tool were used to assess the risk of bias for all non-randomized studies. The overall certainty of the evidence for ALL and AML was assessed using criteria from the GRADE method: risk of bias, inconsistency, indirectness, and imprecision.

Synthesizing the Evidence

A meta-analysis was not planned due to the heterogeneity of the trials.

RESULTS

Search for Existing Systematic Reviews

A search for systematic reviews yielded 164 documents examining the use of MRD testing in adult and pediatric patients with ALL and AML. A total of five underwent full-text review. All systematic reviews were excluded for not answering the research questions of interest. An online search yielded Health Quality Ontario's Health Technology Assessment of MRD in childhood ALL [27]; this was used to identify primary literature published in the pediatric ALL population prior to 2014.

Search for Primary Literature

Literature Search Results

A PRISMA flow diagram of the complete search is available in Appendix 2. Table 4-1 [3-8] and Table 4-2 [12-16] summarize the characteristics of the included studies. Where multiple reports and abstracts were published for a single trial, only the most recent full publication was included, unless other reports contained data that were not available in the most recent publication.

Acute Lymphoblastic Leukemia

Two RCTs [3-7] in the pediatric ALL population and one study from the adult ALL population [8] were included addressing whether MRD testing contributes to a change in clinical management of these patients (i.e., a differential benefit from therapy exists depending on the status of the MRD test result). Please refer to Table 4-1 for details.

Acute Myeloid Leukemia

One RCT comparing intensity of conditioning reported an unplanned subgroup analysis of survival according to MRD status [12] [Note: this study was initially included as an abstract; however, the full report was published prior to document completion and is instead referenced.] Four studies [13-16] in the adult population were found addressing whether MRD testing contributes to a change in clinical management of these patients (i.e., a differential benefit from therapy exists depending on the status of the MRD test result). Please refer to Table 4-2 for details.

Table 4-1. Studies selected for inclusion for ALL

Study	Inclusion	Classification	Treatment	Method of MRD testing, sensitivity and other details
Paediatric population				
Randomized controlled trials				
UKALL 2003 Vora et al (2013, 2014) [3,4]; Eiser, 2017 [5]	Consecutive patients aged 1 to 24 with ALL diagnosed at 46 centres in the UK and Ireland between Oct 1, 2003, and June 30, 2011. Excluded: Patients <1 yr of age or with mature B-cell or Ph-positive ALL.	At diagnosis, patients were stratified according to their risk of relapse on the basis of three metrics: NCI risk criteria, high-risk cytogenetics and early response to induction therapy as assessed by bone-marrow morphology on days 8 and 15 of treatment. Clinical standard and intermediate risk groups were stratified by MRD; clinical high risk patients were not eligible. MRD low risk: Patients with undetectable MRD after induction (day 29) and before interim maintenance. MRD high risk: Patients with at least 0.01% MRD at day 29 of induction. MRD indeterminate: Patients in whom MRD could not be measured because no or poor-quality samples were available and those with persistent disease that was less than 0.01% MRD before the start of interim maintenance.	**MRD low risk** Control arm (n=261): Two delayed intensification courses[a] separated by interim maintenance[b] course followed by continuing therapy[c] Experimental (reduced) arm (n=260): One delayed intensification[a] course followed by continuing therapy[c] **MRD high risk** Control arm (n=266): *Regimen A (clinical standard risk):* Consolidation then two delayed intensification courses[a] separated by interim maintenance[d] course followed by continuing therapy[c]. *Regimen B (clinical high risk):* Consolidation plus 4 weeks of BFM consolidation then two delayed intensification courses[a] separated by interim maintenance[d] course followed by continuing therapy[c] Experimental (augmented) arm (n=267): Consolidation plus 4 weeks of BFM consolidation[e] then two delayed intensification courses[a, f] separated by interim maintenance[b, g] course followed by continuing therapy[c]	MRD was measured with a standardized real-time qPCR method for immunoglobulin and T-cell receptor antigen gene rearrangements with a quantitative range of 10^{-4}. All patients aged 16 years or older were treated as clinical intermediate risk irrespective of day 8 or 15 bone marrow response and were eligible for MRD stratification and randomization. Patients who were not in complete remission at day 29 of induction were not eligible for MRD stratification and randomization.
AEIOP-BFM ALL 2000 Schrappe et al (2018) [6] Conter et al (2010) [7]	Patients 1 to 17 years of age with ALL in one of the participating centres in Italy, Germany, Austria, and Switzerland	MRD standard risk: If MRD was negative on days 33 and 78 with at least two markers with a sensitivity of 1×10^{-4} MRD intermediate risk: MRD was positive at one or both days 33 and 78, but at a level $<10^{-3}$ at day 78 with at least 2 markers. If MRD levels differed between the 2 markers, the highest MRD level was chosen for the	**MRD standard risk** **Control arm (n=583):** Standard delayed intensification **Experimental arm (n=581):** Reduced-intensity regimen **MRD intermediate risk** **Control treatment:** Standard delayed intensification	Response assessment was performed by early cytologic assessment as well as by PCR-MRD on the basis of immunoglobulin and T-cell receptor gene rearrangements. PCR-MRD targets were tested for specificity and sensitivity with the aim for each patient to select 2 targets with a sensitivity of at least 10^{-4} and a quantitative range of at least 10^{-4} for one target and at least 5×10^{-4} for the second target.

Study	Inclusion	Classification	Treatment	Method of MRD testing, sensitivity and other details
Paediatric population				
Randomized controlled trials				
		final MRD classification, provided that the selected markers had a sensitivity of at least 10^{-3} MRD high risk: Genetic characterization of ALL (presence of BCR-ABL1, KMT2A-AFF1) and slow cytologic and molecular response to treatment (prednisone poor response, no CR on day 33, or MRD $\geq 5 \times 10^{-4}$ on day 78).	**Experimental treatment:** Reduced-intensity regimen given twice **MRD high risk** **Control treatment:** 3 blocks of non–cross-resistant drugs followed by reduced intensity regimen given 3 times **Experimental treatment:** 3 blocks followed by standard delayed intensification given twice in the AIEOP group, Standard delayed intensification in the BFM group.	
Adult population				
Retrospective study of trials				
GRAAL-2003 and GRAAL-2005 Dhedin et al (2015) [8]	Patients aged 15 to 60 years with newly diagnosed Ph-negative high-risk ALL	MRD was measured after the first induction course (6 weeks after induction initiation) and after the first 3 blocks of consolidation (12 weeks after induction initiation).	Patients received allogeneic stem cell transplant (n=282) or no stem cell transplant (n=240).	MRD-level quantification was based on patient-specific Ig/TCR gene rearrangement monitoring from BM samples using real-time qPCR. For each patient, preferably 2 independent Ig/TCR targets with a sensitivity of at least 10^{-4} and a quantitative range of 10^{-4} for at least 1 of 2 targets were selected for MRD monitoring.

[a] Intensification course: one dose of pegylated asparaginase on day 4; vincristine, dexamethasone (alternate weeks), and doxorubicin for 3 weeks; and then 4 weeks of cyclophosphamide and cytarabine as during the BFM consolidation course

[b] Interim maintenance: daily oral mercaptopurine and weekly methotrexate with monthly vincristine and steroid pulses for 8 weeks

[c] Continuing therapy: oral mercaptopurine and methotrexate, monthly vincristine and steroid pulses, and intrathecal methotrexate every 3 months.

[d] Interim maintenance: escalating doses of intravenous methotrexate without folinic acid rescue, and vincristine and pegylated asparaginase for 8 weeks

[e] Augmented consolidation contains an additional 4 doses of vincristine and two doses of pegylated asparaginase

[f] Augmented delayed intensification contains an additional 2 doses of vincristine and 1 dose of pegylated asparaginase

[g] Augmented interim maintenance contains increasing doses of intravenous methotrexate without folinic acid rescue, and vincristine and pegylated asparaginase

Abbreviations: AIEOP: Associazione Italiana di Ematologia e Oncologia Pediatrica; ALL: acute lymphoblastic lymphoma; BFM: Berlin–Frankfurt–Munster; CR: complete remission; Ig/TCR: immunoglobulin/T-cell receptor; MRD: minimal residual disease; NCI: National Cancer Institute; PCR: polymerase chain reaction; Ph: Philadelphia chromosome; qPCR: quantitative polymerase chain reaction; RCT: randomized controlled trial

Table 4-2. Studies selected for inclusion for AML

Study	Inclusion	Classification	Treatment	Method of MRD testing, sensitivity and other details
Adult population				
Randomized controlled trial (subgroup analysis)				
BMT CTN 0901 trial NCT01339910 Hourigan et al, 2020 [12]	Patients with AML undergoing alloHCT while in morphologic CR (<5% marrow myeloblasts at assessment) were randomized to myeloablative conditioning (MAC) vs. reduced intensity conditioning (RIC). Frozen blood samples taken during remission prior to alloHCT and conditioning was available for 190/218 patients	NR No mutations: 32% of MAC and 37% of RIC recipients	Retrospective, non-protocol analysis of survival according to conditioning treatment and MRD status	A custom 51 kb anchored multiplex PCR panel (ArcherDx, CO) with coverage of 13 commonly mutated genes in AML was used. A limit of detection was set at an allele frequency of 0.001.
Retrospective studies of trials				
Balsat et al (2017) [13]	Patients (aged 18-60 years) with previously untreated de novo AML with *NPM1*m who were treated in the ALFA-0702 trial and achieved CR/CR_p	MRD assessed after one or two induction courses. MRD negative: MRD levels below 0.01% MRD positive: MRD levels above 0.01%	All patients received 1 or 2 induction courses[a] to achieve a CR or CRp. Eligible patients[b] were then treated with an alloHCT (if they had a matched sibling or 9 or 10/10 HLA-matched unrelated donor) or randomized to one of two types of consolidation[c]. MRD evaluation available in 152 patients.	Quantification of types A, B, and D *NPM*1m transcript levels was performed using a mutation-specific RT-qPCR. The quantitative detection limit of the assays was 0.01%. At AML diagnosis, *NPM*1m transcript level evaluation was assessed in PB and/or in BM. The maximum of these values was considered as the normalized *NPM*1m baseline level.
Freeman et al (2018) [14]	Patients (aged <60 years) enrolled in the NCRI AML 17 trial – high risk myelodysplastic syndrome and secondary AML. Excluded: Patients with APL.	MRD negative: Not defined MRD positive: Samples with any level of MRD detected above a diagnostic LAIP or different-from-normal follow-up LAIP threshold	Clinically standard risk patients received two cycles of the same induction and were then randomized to receive either 1 or 2 courses of high-dose cytosine arabinoside. Clinically high-risk patients were randomized to either FLAG-Ida or daunorubicin/clofarabine for the second induction followed by transplant if eligible. *FLT3*-ITD	Samples for MFC-MRD were requested at baseline (bone marrow and/or blood) and following each course (bone marrow). MFC-MRD analysis was performed centrally, using standardized gating strategy that screened for different from-normal LAIPs on blasts pre-treatment and tracked these (approximately 0.02%

Study	Inclusion	Classification	Treatment	Method of MRD testing, sensitivity and other details
			mutant patients were directed to the lestaurtinib randomization until 2012. Post course 1, data available in 1443 patients. Post course 2, data available in 806 patients.	to 0.05% sensitivity thresholds) but also applied the different-to-normal approach in follow-up samples to detect changes in blast LAIPs (approximately 0.05% to 0.1% sensitivity threshold).
Prospective study				
Jongen-Lavrencic et al (2018) [15]	Patients (aged 18-65 years) who had a confirmed diagnosis of previously untreated AML or had refractory anemia with excess of blasts indicating high or very high risk of relapse on the Revised International Prognostic Scoring System. Patients had to be in either CR or CRi after receiving 2 cycles of induction chemotherapy.	MRD negative: MRD levels below 0.1% on MFC MRD positive: MRD levels above 0.1% on MFC	Patients (n=430) were treated according to the clinical protocol of either the Dutch-Belgian Cooperative Trial Group for Hematology-Oncology (HOVON) or the Swiss Group for Clinical Cancer Research (SAKK)	Targeted next-generation sequencing and MFC[d]
Retrospective study				
Chen et al (2015) [16]	Patients with newly diagnosed or relapsed or refractory AML who achieved CR, CRp, or CRi after induction therapy. Excluded: Patients with APL	MRD positive: Any level of abnormal blast population detected	Patients (n=245) were treated according to institutional review board-approved protocols.	Ten-color flow cytometry analysis was performed.

[a] A timed-sequential induction chemotherapy consisted of with daunorubicin, cytarabine, and granulocyte-colony stimulating factor priming. A salvage course with idarubicin and high-dose cytarabine was permitted in patients not achieving CR or CRp after this first induction course.

[b] Patients in CR/CRp with nonfavorable AML according to the ELN classification or those who needed the salvage (late CR/CRp)

[c] Three HDAC cycles or three CLARA cycles if no identified donor

[d] The Illumina TruSight Myeloid Sequencing Panel was used

Abbreviations: ALFA: Acute Leukemia French Association; AML: acute myeloblastic leukemia; APL: acute promyelocytic leukemia; alloHCT: allogeneic hematopoietic cell transplantation; BM: bone marrow; CLARA: clofarabine plus cytarabine; CR: complete remission; CRi: incomplete blood count recovery; CRp: complete remission with incomplete platelet recovery; ELN: European LeukemiaNet; FLAG-Ida: Fludarabine, Cytarabine, Idarubicin, Granulocyte-colony stimulating factor; FLT3: FMS-like tyrosine kinase 3; HDAC: high-dose cytarabine; ITD: internal tandem duplication; LAIP: leukemia-associated-immunophenotypes; MFC: multiparameter flow cytometry; MRD: minimal residual disease; *NPM*1*m*: nucleophosmin-2 gene mutation; PB: peripheral blood; PCR: polymerase chain reaction; RT-qPCR: reverse-transcriptase quantitative PCR

Study Design and Quality

Risk of bias assessments for both RCTs and non-RCTs are reported in Appendix 4, Tables A4-1 and A4-2, respectively. The quality characteristics of the RCTs are reported in Table A4-3. All published reports of the trials were searched for the necessary details.

Acute Lymphoblastic Leukemia

Risk of Bias and Quality Characteristics

Randomized Controlled Trials

Two RCTs [3,4,6,7] were included and assessed. Both RCTs scored 'low' on most domains of the risk of bias tool, while scoring 'high' for performance bias and detection bias; however, it is not feasible to blind participants and personnel to treatment reductions or treatment augmentations.

Both RCTs provided details about randomization and sample size calculations. No patients were lost to follow-up. Baseline characteristics were balanced and intention-to-treat analysis was conducted in both. The UKALL2003 trial identified EFS as the primary outcome while the AEIOPP-BFM ALL 2000 trial noted disease-free survival (DFS) as its primary outcome.

Non-randomized Controlled Studies

One study [8] was included and was assessed as having a moderate risk of bias.

Certainty of the Evidence for ALL

The certainty of the evidence for all outcomes is moderate due to risk of bias and indirectness (i.e., variation in patient population).

Acute Myeloid Leukemia

Risk of Bias and Quality Characteristics

Four non-randomized studies [13-16] were included and assessed. They were all assessed as having a moderate risk of bias. While the BMT CTN 0901 trial [12] randomized patients to conditioning intensity, there was no randomization regarding MRD, and analysis according to MRD status was an unplanned retrospective subgroup analysis. It was determined that assessment as a randomized trial was inappropriate.

Certainty of the Evidence for AML

According to GRADE, observational studies without special strengths or important limitations provide evidence with a low level of certainty.

Outcomes

Acute Lymphoblastic Leukemia

1. What benefit to clinical management does MRD testing contribute in the treatment of patients with ALL?

Two RCTs [3,4,6] in the pediatric population and one retrospective analysis of an RCT [8] in the adult population that address the clinical utility of MRD testing in patients with ALL were found. Refer to Table 4-1 for details on clinical and MRD risk classification details as well as for details on the regimens used.

The UKALL 2003 trial [3,4] tested whether adjustment of treatment intensity according to MRD risk stratification was feasible. Patients in clinically standard- and intermediate-risk groups were stratified by bone marrow MRD qPCR for immunoglobulin and T-cell receptor

antigen gene rearrangements) at the end of induction (day 29) and recovery from consolidation. Clinically high-risk patients were not eligible for MRD stratification. Patients identified as MRD low risk were randomly assigned to receive one (reduced treatment) or two (standard treatment) delayed intensifications, while MRD high-risk patients were randomly assigned to receive augmented treatment or standard treatment based on clinical risk.

The AEIOPP-BFM ALL 2000 trial [6] also tested MRD-directed treatment augmentation and reduction to prove non-inferiority. Patients with Ph-negative ALL were stratified by MRD (standard risk, intermediate risk and high risk, Table 4-1). The results for the MRD intermediate-risk and MRD high-risk groups are not yet available and will not be discussed in this review. In the MRD standard-risk cohort, patients were randomized to receive standard delayed intensification or a reduced-intensity regimen, while stratified by allocation to a preceding random assignment and treatment centre. The duration of the experimental arm was shorter than the control arm (28 vs. 49 days).

One retrospective study [8] of the GRAALL-2003 and GRAAL-2005 trials was found that studied whether MRD levels are predictive of survival outcomes in adults. This study conducted a subgroup analysis of 522 patients at high-risk with Ph-negative ALL who were candidates for stem cell transplant at first complete remission. MRD was measured using real-time qPCR for immunoglobulin/T-cell receptor gene rearrangements. Of the 522 patients, 54.0% received a stem cell transplant. There were no differences in characteristics between those who received a stem cell transplant and those that did not with the exception of more patients with t(4;11)/*MLL* abnormalities receiving a transplant (p=0.015).

Survival

Table 4-3 [3,4,6] presents survival outcomes for the RCTs; survival data for any retrospective analyses are presented only in text and not within the table.

The UKALL 2003 trial [3,4] reported no significant differences in five-year EFS in patients identified as MRD low risk (<0.01% at day 29 induction and before interim maintenance) who received reduced therapy (49.9%) compared with those who received standard therapy (50.1%; OR 1.00; 95% CI 0.43 to 2.31; p=0.99). Similarly, there was no difference in five-year OS (OR 0.67; 95% CI 0.19 to 2.30; p=0.53). However, in patients identified as MRD high risk (≥0.01%), there was a statistically significant difference in five-year EFS between those who received augmented therapy (89.6%; 95% CI 85.9 to 93.3) and those who received standard therapy (82.8%; 95% CI 78.1 to 87.5, and OR 0.61; 95% CI 0.39 to 0.98; p=0.04). There was no difference in five-year OS (OR 0.67; 95% CI 0.38 to 1.17; p=0.16).

The AEIOPP-BFM ALL 2000 trial [6] used DFS as its primary outcome and reported no significant differences in eight-year DFS between those who received standard delayed intensification (91.7% ± 1.2%) and those who received the reduced-intensity treatment (89.6% ± 1.3%; HR 1.36; 95% CI 0.92 to 2.00; p=0.12) in the intention-to-treat analysis. However, it is important to note that the per protocol analysis reported a statistically significant difference between the control (92.3% ± 1.2%) and experimental arms (89.2% ± 1.3%%; HR, 1.50; 95% CI 1.01 to 2.22; p=0.041). The difference in significance between the per protocol and intention-to-treat analyses is the net result of one event and should be interpreted with caution. No significant differences in eight-year OS between patients who received the standard delayed intensification (98.9% ± 0.6%) and those that received the reduced-intensity treatment (96.1% ± 0.8%; HR, 2.00; 95% CI 0.97 to 4.13; p=0.055) were reported in the intention-to-treat analysis.

The retrospective study of the GRAALL-2003 and GRAAL-2005 trials [8] reported statistically significant interactions between poor post-induction MRD level (MRD $\geq 10^{-3}$) and stem cell transplant for OS (p=0.002) in adult patients with Ph-negative ALL. In further subgroup analyses, similar results were obtained for patients with B-cell precursor ALL (OS, $p_{interaction}$=0.050) and patients with T-ALL (OS, $p_{interaction}$=0.010). In a multivariable analysis of

these patients, the significant interaction between MRD response and stem cell transplant effect was observed after adjustment of age, white blood cell count and resistance to steroid prophase ($p_{interaction}$=0.002). For patients with MRD $\geq 10^{-3}$, there was a significant survival benefit, (HR, 0.41; 95% CI 0.22 to 0.76; p=0.005). For patients with MRD $<10^{-3}$, there was no difference in survival benefit (p=0.17). The interaction term for OS was p=0.002.

Adverse Events

Table 4-4 [3,4,6] presents adverse event data for the RCTs; if any adverse event data were reported in any retrospective analyses, they are only presented within the text.

The UKALL 2003 trial [3,4] reported no significant difference between MRD low-risk patients who received standard or reduced therapy in grade 3 to 4 toxic events (p=0.30), serious adverse events (p=0.26), or treatment-related deaths (p=0.08). In MRD high-risk patients who received standard or augmented therapy, no significant difference was reported for grade 3 to 4 toxic events (p=0.55). However, there was a significant difference in patients who experienced serious adverse events between those who received augmented therapy (45%) and those who received standard therapy (34%, p=0.02).

In the AEIOP-BFM ALL 2000 trial [6], non-life-threatening, life –threatening, and fatal adverse events were comparable in standard-risk patients who received either reduced intensity or standard delayed intensity treatments. P-values were not reported.

Quality of Life

A total of 874 (61.2%) patients aged four to 18 years in the UKALL 2003 trial participated in the health-related quality of life study. The following questionnaires were completed by parents at five time points: a) PedsQL4.0 generic core, a 23-item scale assessing total, physical and psychosocial health-related quality of life for each time point; b) PedsQL 3.0 Cancer Module, a modified 19-item questionnaire assessing the impact of disease and treatment on pain and hurt, nausea, procedural anxiety, worry about side effects, concern for appearance and communication; and c) a modified measure of parents' perceived care giving burden in families with a child with asthma, consisting of 11 items asking parents how often they were bothered about specific tasks associated with their child's illness.

There were no statistically significant effects of randomization on health-related quality of life or parental care-giving burden within the MRD low-risk and MRD high-risk groups.

Table 4-3. Outcomes of RCTs with pediatric patients with ALL

Author	Method of MRD testing	N	Median age (yr)	Median follow-up	Time point of MRD testing	MRD risk group	Treatment arm (%)[b]	Outcomes		
								5yr EFS	5yr relapse rate	5yr OS
UKALL 2003 Vora et al (2013) (2014) [3,4]	Real-time qPCR with a quantitative range of 10^{-4}	2721[a]	NR	57 mo (42-72)	Induction (day 29) and before interim maintenance	$MRD_{lowrisk}$: <0.01% (38.9%)	Reduced therapy: (49.9%)	94.4% (95% CI 91.1-97.7)	5.6% (95% CI 2.3-8.9)	97.9% (95% CI 95.7-100.1)
							Standard therapy: (50.1%)	95.5% (95% CI 92.8-98.2) OR 1.00; 95% CI 0.43-2.31; p=0.99	2.4% (95% CI 0.2-4.6) p=0.23	98.5% (95% CI 96.9-100.0) OR 0.67; 95% CI 0.19-2.30; p=0.53
				70 mo (52-91)	Induction (day 29)	$MRD_{highrisk}$: ≥0.01% (29.7%)	Augmented therapy: (50.1%)	89.6% (95% CI 85.9-93.3)	NR	92.9% (95% CI 89.8-96.0)
							Standard therapy: (49.9%)	82.8% (95% CI 78.1-87.5) OR 0.61; 95% CI (0.39-0.98); p=0.04	NR	88.9% (95% CI 85.0-92.8) OR 0.67; 95% CI (0.38-1.17); p=0.16[c]
AEIOP-BFM ALL 2000 Schrappee t al (2018) [6]	qPCR with a quantitative range of at least 10^{-4} for one target and at least 5 × 10^{-4} for the second target	1164[d] randomly assigned	NR	103.2 mo	NR	$MRD_{standardrisk}$	Reduced intensity: (49.9%)	NR	NR	96.1% ± 0.8%
							Standard DI: (50.1%)	NR	NR	98.9% ± 0.6% HR, 2.00; 95% CI 0.97-4.13; p=0.055

[a] Eligible for MRD stratification

[b] Percentage of those randomly assigned

[c] Unadjusted

[d] Number randomly assigned

Abbreviations: ALL: acute lymphoblastic leukemia; CI: confidence interval; DI: delayed intensification; EFS: event-free survival; HR: hazard ratio; IQR: interquartile range; mo: months; MRD: minimal residual disease; N: number; NR: not reported; OR: odds ratio; OS: overall survival; qPCR: quantitative polymerase chain reaction; RCT: randomized controlled trial; yr: year

Table 4-4. Adverse event outcomes from RCTs of pediatric patients with ALL

Author	Method of MRD testing	MRD risk group	Treatment	Grade 3-4 toxic events	Serious adverse events	Treatment-related deaths
UKALL 2003 Vora et al (2013) (2014) [3,4]	Real-time qPCR with a quantitative range of 10^{-4}	$MRD_{low\ risk}$: <0.01%	Reduced therapy	73%	27%	0%
			Standard therapy	77% (p=0.30) *17% of patients had toxic events during second course	31% (p=0.26)	1.2% (95% CI 0-2.6) p=0.08
		$MRD_{high\ risk}$: ≥0.01%	Augmented therapy	86%	45%	Death during remission, 2.6%
			Standard therapy	84% (p=0.55)	34% (p=0.02)	3.4%
AEIOP-BFM ALL 2000 Schrappe et al (2018) [6]	qPCR with a quantitative range of at least 10^{-4} for one target and at least 5 × 10^{-4} for the second target	$MRD_{standard\ risk}$	Reduced intensity	Non-life threatening, 4.2%	Life threatening, 1.2%	Fatal, 0.7%
			Standard DI	5.0%	1.7%	0.4%

Abbreviations: ALL: acute lymphoblastic leukemia; CI: confidence interval; DI: delayed intensification; MRD: minimal residual disease; RCT: randomized controlled trial; qPCR: quantitative polymerase chain reaction

Table 4-5. Outcomes of trials of patients with AML

Author	Method of MRD testing	N	Median age (yr)	Median follow-up	Time point of MRD testing	MRD risk group	Outcomes	
							OS	other
Hourigan et al (2020) [12]	PCR panel with detection limit of 0.001.	190	55	NR >49 months in survivors	Prior to alloHCT	$MRD_{negative}$	3-yr OS 56% MAC vs. 63% RIC, p=0.96	In multivariate analysis differences (RIC vs MAC) not significant: relapse HR=1.78 (95% CI 0.72-4.38 p=0.210); and OS HR=1.05 (95% CI 0.50-2.18, p=0.905)
						$MRD_{positive}$	3-yr OS 61% MAC vs. 43% RIC, p=0.02	In multivariate analysis, RIC associated with increased relapse (HR=6.38, 95% CI 3.37-12.10, p<0.001), decreased relapse-free survival (HR=2.94, 95% CI 1.84-4.69, p<0.001, decreased OS (HR=1.97, 95% CI 1.17-3.30, p=0.01)

Abbreviations: alloHCT: allogeneic hematopoietic cell transplantation; AML: acute myeloid leukemia; CI: confidence interval; DFS: disease-free survival; HR: hazard ratio; MAC: myeloablative conditioning; MRD: minimal residual disease; NR: not reported; OS: overall survival; PCR: polymerase chain reaction; RIC: reduced intensity conditioning; yr: years

Acute Myeloid Leukemia

1. What benefit to clinical management does MRD testing contribute in the treatment of patients with AML?

One trial was found that provided evidence around the clinical utility of MRD testing adult patients with AML. Additionally, five studies were found that conducted multivariable analyses to determine whether alternate therapies improved survival based on MRD status. Please refer to Table 4-3 for details on clinical and MRD risk classification details as well as for details on the regimens used.

The BMT CTN 0901 trial [12] compared conditioning intensity in adult patients with AML in morphological complete remission undergoing an alloHCT. MRD status was determined by a multiplex PCR panel. A retrospective subgroup analysis compared patients who were MRD-positive or MRD-negative and according to whether they were administered MAC or RIC. Patients in the two consolidation intensity groups were matched for baseline characteristics including age, sex, comorbidity, disease risk, disease duration, cytogenetics, donor type and match, graft type, and anti-thymocyte globulin use.

Of the remaining four studies that conducted multivariable analyses to determine whether alternate therapies improved survival based on MRD levels, the first by Balsat et al. [13] reported on patients with non-favourable AML from the ALFA-0702 trial that were eligible for alloHCT in first remission. MRD was measured using real-time RT-qPCR.

The second by Freeman et al. [14] reported on patients with AML or high-risk myelodysplastic syndrome from the NCRI AML17 trial. MRD was measured using MFC, which screened for leukemia-associated immunophenotypes. This study conducted a subgroup analysis of 204 *NPM*1-wt patients to determine the effect of alloHCT in first complete response according to MRD status at the second course.

In the prospective cohort study by Jongen-Lavrencic et al. [15], 430 patients with newly diagnosed AML were evaluated to investigate whether molecular monitoring with NGS could predict recurrence.

Finally, a retrospective study by Chen et al. [16] studied 245 patients with newly diagnosed, relapsed, or refractory AML. MRD was measured using 10-colour flow cytometry.

Survival

Table 4-5 [12] presents survival outcomes for the RCTs; survival data for any studies presenting multivariable analyses are presented within the text only.

The BMT CTN 0901 trial reported significant differences in three-year OS in MRD-positive patients who were randomized to receive either MAC (61%) or RIC (43%, p=0.02) before alloHCT, while no difference was found in three-year OS in MRD-negative patients who were randomized to receive either MAC (56%) or RIC (63%, p=0.96).

In a multivariable analysis of 64 patients in the study by Balsat et al. [13], survival was significantly improved by alloHCT in those with a <4-log reduction in *NPM1*m peripheral blood MRD (HR, 0.25; 95% CI 0.06 to 0.98; p=0.047) while this was not observed in patients with a >4-log reduction (HR 2.11; 95% CI 0.57 to 7.71; p=0.261). The interaction between alloHCT effect and MRD log reduction was significant (p=0.027).

In the study by Freeman et al. [14], survival was not significantly improved in MRD-positive patients (HR 0.72; 95% CI 0.31 to 1.69) or MRD-negative patients (HR 1.68; 95% CI 0.75 to 3.85; $p_{interaction}$=0.16) of the *NPM1*-wt standard-risk group.

In the study by Jongen-Lavrencic et al. [15], a multivariable analysis confirmed that the persistence of non-DTA (DNMT3A, TET2, or ASXL1) mutations maintained significant independent prognostic value for relapse and survival; however, the interaction between the

detection of residual disease and type of consolidation therapy (i.e., no therapy, chemotherapy, or autologous or allogeneic HSCT) was not significant. The p-value was not reported.

In the final included study by Chen et al [16], a multivariable analysis determined that MRD is not predictive of a differential treatment effect by HCT for relapse, OS, or RFS (interaction, p-value was not reported).

Adverse Events

No adverse event data were reported.

Quality of Life

No quality of life data were presented.

Ongoing, Unpublished, or Incomplete Studies

There were no ongoing, unpublished, or incomplete studies found that met the inclusion criteria of this guideline. This search was conducted October 9, 2019 at clinicaltrials.gov.

DISCUSSION

In acute leukemia, methods to confirm responsiveness to therapy early in the therapeutic course are actively sought as it is clear that some patients, using the current models of risk assessment and treatment assignment, have long remissions after less-intensive therapy while others are insufficiently treated. Basing post-remission therapy exclusively on risk of relapse determined by pretreatment information is problematic since in many cases those metrics have limited ability to forecast prognosis [28]. Post-therapy assessment, which has been in routine practice at a morphological level, has helped stratify patients into those more or less responsive to the administered treatment. However, MRD testing, which represents analyses that are several-fold more sensitive than the traditional morphological assessment, have been reported to further enhance the achievable level of prognostication. It is of prognostic value in both pediatric and adult patients with ALL and AML and its routine use in treatment assessment of these patients is important and beneficial in informing patients and health care providers of their prognoses.

As is true of any emerging technology, approaches to measure MRD currently require appropriate laboratories that meet stringent levels of standardization and meet the appropriate technical metrics [2,22]. There are several methods for measuring MRD (i.e., PCR, droplet digital PCR, flow cytometry, and NGS) and each has been applied in various studies at different times during therapy with different thresholds that define positivity. However, in order to both signify that MRD testing has a role and in an attempt to standardize the approaches to measurement, the latest ELN guidelines have modified the complete remission designation into complete remission with and without MRD [1]. Technical details on minimum standards have been recently updated as well [2,22].

The field of utilizing MRD testing as a predictive factor in guiding treatment decisions in acute leukemia is being led by those in the area of pediatric ALL where there has been regulated testing within many clinical trials. This has driven knowledge and enthusiasm for using peri- and post- therapy MRD to stratify patients into higher- and lower-risk disease groups, even beyond the traditional prognostic markers [29] in the ELN guidelines. These approaches have provided some evidence that interventions such as minimizing therapy in those deemed at lower risk of relapse or increasing treatment intensity in those with higher risk of relapse, may alter outcomes [30]. Small, non-randomized, studies have explored this concept further in higher-risk disease by limiting the more intensive therapy to those who continued to be MRD

positive at a given time point during therapy with positive results [30-32]. This has enabled clinical trials to prospectively randomize patients to different therapies based on the outcome of these tests. It has also been noted that the interpretation of these results may vary depending on the known genetic profile, pretreatment, and/or the timing of the testing.

The studies identified in this systematic review to determine the clinical utility of MRD testing are prospective, randomized trials, while the ability of MRD testing to predict patient outcomes are based on studies that have conducted multivariable analyses. Studies involving adult patients with ALL and AML have been following the lead in pediatrics but appear to have more variability in methodology and timing of MRD testing, making it very difficult to conclude definitively that these assays can impact patient outcome when there is a need to decide between different therapies [33]. However, those studying adult ALL have taken lessons from their pediatric colleagues through the adoption of pediatric-style chemotherapy regimens and applying a similar level of MRD monitoring. The UKALL 2003 randomized trial in pediatrics [3,4] demonstrated that lowering the intensity of therapy after confirming an MRD-negative state was safe and maintained efficacy [3], while intensifying therapy based on MRD positivity did not result in an OS advantage, although an increased EFS was reported [4]. Although there was increased toxicity due to the chemotherapy, the authors stated that, "it was not associated with a higher treatment-related mortality or a significant effect on quality of life and, with the exception of pancreatitis, the various toxicities resolved fully without late complications". Therefore, we can conclude that reducing intensity for MRD-negative disease, at a defined time point that is dependent on the regimen being studied, may be safe. Similarly, the AEIOPP-BFM ALL 2000 trial [6], also in pediatric patients, found no significant differences in eight-year OS between those who received standard delayed intensification and those that received the reduced-intensity treatment. Increased intensity, however, is not as straightforward and should continue to be restricted to clinical trials, if available. Ongoing trials intensifying the therapy with novel agents that add efficacy with limited toxicity are likely to provide a solution to this problem.

Unlike in pediatrics, hematologists treating adult disease have tended to transplant many of their patients with ALL until it was demonstrated that a pediatric-style regimen without transplant could produce superior results in some subgroups of patients [34], with treatment defaulting to transplant if information arises to suggest that patients will have an unsatisfactory outcome with the pediatric-style chemotherapy [9]. With this background, the study by Dhedin et al. [8] followed a pediatric-style regimen for all those without high risk features, including post-therapy MRD positivity. Those that were identified as higher risk and had a suitable donor were directed to the stem cell transplant and were compared with those who did not have a donor or had a donor but could not proceed to transplant for other reasons. That analysis demonstrated that MRD positivity and the presence of an *IKZF1* gene deletion correlated with particularly poor outcome, which could be mitigated by a stem cell transplant. These results, however, have to be carefully considered as 61 of the 240 patients treated with just chemotherapy had a suitable donor but did not proceed with transplant for a variety of reasons and were included in the chemotherapy arm, which may have influenced the results.

Other than the study by Dhedin et al. [8] with limitations as noted, the literature review was not able to address the subquestions of whether positive MRD before transplant in either ALL or AML and subsequent treatment to achieve MRD negativity delays stem cell transplant and alters patient outcomes. The literature does not allow an in-depth review of these questions and this topic will be revisited as the literature matures.

MRD testing in AML is much more heterogeneous than in adult ALL, as there is not a single reliable assay for biological, historical, and technical reasons [22]. Patients with acute promyelocytic leukemia (APL) (bearing the *PML-RARA* or variant *RARA* translocation), and with other AMLs bearing core binding factor (CBF) translocations, or alterations in nucleophosmin

(*NPM1*) or FMS-like tyrosine kinase 3 (*FLT3*), have leukemia-specific markers that can be reliably measured at low levels by PCR [13,35-41]. In the case of APL and CBF-leukemia, this testing has been available and used for monitoring for many years. In some cases, it has been shown that re-treating patients who have become MRD positive after being negative (so-called molecular relapse) post treatment and prior to a morphological relapse, has resulted in improved outcomes, particularly in cases of APL, and CBF-translocation-positive AML. However, such early intervention in patients who fail to reach the defined milestones is less well understood, and transplant is often the preferred back-up plan. Only one study attempted to examine a risk adapted approach in AML, but although it appeared to demonstrate a usefulness to altering therapy based on MRD status, it was not randomized and has not been repeated or replicated [42].

The intermediate AML risk group, defined by a normal karyotype, is very heterogeneous. While this group is often directed toward transplant, there are many patients who have been cured without a transplant [43,44]. Additional prognostication in these patients relies on measurement of other gene mutations to either define them as higher risk (i.e., *FLT3* mutation) or lower risk (i.e., *NPM1* mutation in the absence of a *FLT3* mutation). The interpretation of molecular prognostic markers is in constant evolution. This additional information is generally used to assist in decisions regarding the pursuit of an allogeneic transplant in first complete remission.

Limited evidence was found in the adult or pediatric AML literature addressing the clinical utility of MRD testing in this population. Additional questions can also be asked further down the treatment path such as whether MRD-positive patients should receive additional therapy prior to transplant to potentially improve their post-transplant outcome and/or whether post-transplant interventions such as preemptive donor lymphocyte infusion, rapid immunosuppression taper or maintenance therapy, should be implemented. None of this is definitively known, although many of these questions are being studied (i.e., NCT02272478) [45].

The many influences on outcome are not well understood and may be additive; a study in patients with CGF with additional mutations found only residual MRD was predictive of a poor outcome [39]. On the other hand, there are some selected genetic profiles for which simply intensifying the therapy does not appear to change outcome and so the clinical utility of MRD testing in those groups is unclear [46-49]. Using MRD to distinguish patients who could be spared the intensive pathway, generally associated with long-term remission, or to identify those who are incurable by traditional treatment paradigms would help deliver enough therapy to those who need it while avoiding unnecessary toxicity in those who are not likely to benefit or do not need the higher intensity.

Despite interest in this area, many questions remain. What is the optimal time to test? Should testing be assessed after one or two cycles of chemotherapy or later? Should the decision be based on a single time point or a trend? Which is the most appropriate and reliable test? What is the ideal level of detection? These and many other issues are being studied and discussed, with an aim of reaching consensus [22,50]. Nonetheless, despite these remaining questions, in some patient groups, MRD detection has unambiguously been shown to have a high positive predictive value for subsequent morphologic relapse, which typically occurs within 12 months of detection of MRD regardless of the traditional risk group [40,51,52].

Given the building evidence for MRD being not only a prognostic marker but one which may enable treatment stratification, it is incumbent on our community to provide MRD testing more universally, as an initial step, and rigorously evaluate the literature in order to implement the information gained from MRD as fully as possible to optimize therapy decisions for patients.

CONCLUSIONS

Although the presence of MRD is of prognostic value in patients with ALL and AML, interventional studies demonstrating clinical utility in the adult population using that information are few. In the development of this guideline to evaluate the evidence for the ability of MRD testing to help with decisions of escalation or de-escalation of the intensity of the therapeutic pathway, it was noted that there remains a paucity of level 1 evidence. Pediatric patients with ALL have the most established data, which has resulted in some clinicians extrapolating the approach to the adult ALL population, while noting that confirmatory studies are needed in this population. Treatment augmentation or reduction based on MRD is not straightforward and should continue to be restricted to clinical trials, if available. For patients with AML, while also providing prognostic information, there is currently insufficient evidence for or against the use of MRD in guiding decisions between various treatment options in adult patients with AML (i.e., adjustment of treatment intensity/treatment stratification). This is in contrast to the established prognostic role of MRD testing in both adult and pediatric ALL and AML. There is a need, however, to evaluate ongoing clinical trials in the near future evaluating the clinical utility of MRD testing in adult patients with acute leukemia.

Minimal Residual Disease Testing in Acute Leukemia

Section 5: Internal and External Review

INTERNAL REVIEW

The guideline was evaluated by the GDG Expert Panel and the PEBC RAP (Appendix 1). The results of these evaluations and the Working Group's responses are described below.

Expert Panel Review and Approval

Of the 21 members of the GDG Expert Panel, 18 members voted and 0 abstained, for a total of 85.7% response in August 2019. Of those who voted, 15 approved the document (83.3%). The main comments from the Expert Panel and the Working Group's responses are summarized in Table 5-1.

Table 5-1. Summary of the Working Group's responses to comments from the Expert Panel

Comments	Responses
1. Some reviewers noted that both NGS and MPS (massively parallel sequencing) are used intermittently in the document – it would be better to decide on one term and replace the other.	We have corrected this.
2. Defining clinical utility would be helpful in the document.	We have added a definition in the introduction.
3. Suggestion to not use the term predictive/predict when referring to correlation with prognosis, otherwise there may be confusion with predictive of benefit of treatment (i.e., predictive marker) as markers that detect recurrence do not necessarily predict it.	We have now clarified the use of the term predictive within the guideline.
4. This paper provides the basis for considering the question of why MRD status is meaningful since this point is not made strongly in this manuscript and heads straight into the question of whether interventions in MRD-stratified groups have made an impact. (Berry et al. Association of minimal residual disease with clinical outcome in pediatric and adult acute lymphoblastic leukemia: a meta-analysis. JAMA Oncol. 2017;3(7):e170580.)	We have now clarified the basis for considering the question of why MRD status is meaningful by adding a preamble to the Recommendations.
5. Reviewers found it unclear that MRD testing is actually very useful for Ph-positive ALL and for some subtypes of AML (i.e., APL and CBF leukemias) and do not agree that there is not a single reliable methodology out there for testing.	The purpose of this guideline was not to provide the evidence for prognostic capabilities of MRD testing but rather its predictive clinical utility. The prognostic capabilities are accepted and we have now clarified this by adding a preamble to the Recommendations. We have also added within the preamble the following, "This review does not evaluate optimal methods for MRD detection, nor does it recommend specific markers for testing. Both the testing modality and the set of biomarkers that

	could be assessed are rapidly evolving fields, and decisions about these practical aspects of MRD testing will require constant review."
6. There is no question that MRD in adult (as well as pediatric) ALL provides critical prognostic information about both EFS and OS and is presumably the basis for the qualifying statement in the document: "While the prognostic information from MRD testing is accepted in clinical practice and used to inform patients of their prognosis, the ability of MRD testing results to predict differential treatment effect of stem cell transplant on OS in adult patients with ALL is not yet established" So if it accepted in clinical practice, why is it not available in clinical practice in Ontario? The question is not IF we need to do proper testing, but rather HOW (by what recognized assay with appropriate quality control) and WHERE (which laboratories should do it–one centralized or several distributed?). As there is inarguably abundant evidence to support the prognostic impact of MRD testing, especially when predicting those patients at high risk of relapse, MRD testing should be recommended and accessible.	The Working Group members understand the concerns made by this reviewer and have now added a preamble to Section 2: Recommendations to clarify the purpose and role of this guideline. Within this preamble it states, "This review seeks to examine and collate the evidence that would allow expansion of the role of MRD testing in the adult population from a prognostic test to one that guides treatment. This review does not evaluate optimal methods for MRD detection, nor does it recommend specific markers for testing. Both the testing modality and the set of biomarkers that could be assessed are rapidly evolving fields, and decisions about these practical aspects of MRD testing will require constant review." Availability of MRD testing or current practice in Ontario was not within the scope of the document.
7. The recommendation regarding MRD testing in AML is misleading if not inaccurate, and should be replaced by the qualifying statement, as once again the utility of the test in determining prognosis should not be minimized. It also ignores the body of evidence for prognostic information by MRD testing in APL, and CBF leukemia that is well established.	The purpose of this guideline was not to provide the evidence for prognostic capabilities of MRD testing but rather its predictive clinical utility. The prognostic capabilities are accepted and we have now clarified this by adding a preamble to the Recommendations.
8. Suggestion to include the abstract of the 0901 trial presented at EHA 2019 by Dr. Christopher Hourigan.	The abstract was included with the search update; this was substituted with the full publication after external review.
9. In the section "Study Selection Criteria and Process" (page 11), it states that "*..if no RCTs or non-randomized comparative studies then single-arm studies with ≥100 participants where confounders are controlled for*". How was the number "100" chosen?	The Working Group members had decided that single-arm studies with greater than 100 participants would be sufficient in identifying any patterns.
10. Why is the role of MRD in specific subgroups not addressed (e.g. AYA ALL, Ph-positive ALL, CBF AML, APL)? Ph-positive ALL, CBF AML and APL are mentioned in the guideline (paragraph 3, page 26) and it states that "in the case of APL, CBF leukemia and BCR-ABL1-positive AMLs, this testing has been available and used for monitoring for many years"	The Working Group members agree the that prognostic capabilities of MRD testing for APL, CBF leukemia and BCR-ABL1-positive AMLs have been shown and have been used for many years. However, the purpose of this guideline was not to provide the evidence for prognostic capabilities of MRD testing but rather its predictive clinical utility.
11. There are no specific references/recommendations concerning	

MRD assessment for patients with CBF AML, APL, or BCR-ABL1-positive (Philadelphia-positive) B-ALL in this guideline.	

RAP Review and Approval

Three RAP members reviewed and conditionally approved this document in August 2019. The main comments from the RAP and the Working Group's responses are summarized in Table 5-2.

Table 5-2. Summary of the Working Group's responses to comments from RAP

Comments	Responses
1. The two recommendations are clear but do not align with the guideline objectives. All the necessary information is in the Recommendations, Key Evidence and Interpretation of the Evidence section but should be reorganized.	The Working Group has now modified the guideline to make the recommendations and systematic review more clear.
2. The qualifying statement for Recommendation 1 could be improved. What is meant by differential effect of stem cell transplant on OS could be stated more clearly. The bullet on the timing of MRD testing is weak and the evidence for doing MRD testing on bone marrow is lacking	Thank you for your comment. The Qualifying Statements of Recommendation 1 are based on expert opinion.
3. The analysis of the quality characteristics and risk of bias of the limited number of studies is well presented and clearly explained. The discussion section is very helpful in summarizing the potential role of MRD testing in ALL and AML, the existing evidence and where additional research needs to be done.	Thank you for your comment.

EXTERNAL REVIEW

External Review by Ontario Clinicians and Other Experts

Targeted Peer Review

Three targeted peer reviewers from Canada who are considered to be clinical and/or methodological experts on the topic were identified by the Working Group. Two agreed to be the reviewers (Appendix 1). Two responses were received. Results of the feedback survey are summarized in Table 5-3. The main comments from targeted peer reviewers and the Working Group's responses are summarized in Table 5-4.

Table 5-3. Responses to nine items on the Targeted Peer Reviewer questionnaire

	Reviewer Ratings (N=2)				
Question	Lowest Quality (1)	(2)	(3)	(4)	Highest Quality (5)
1. Rate the guideline development methods.	0	0	0	1	1

2. Rate the guideline presentation.	0	0	0	1	1
3. Rate the guideline recommendations.	0	0	0	2	0
4. Rate the completeness of reporting.	0	0	0	2	
5. Does this document provide sufficient information to inform your decisions? If not, what areas are missing? [3]	0	0	1	0	0
6. Rate the overall quality of the guideline report.	0	0	1	1	0
	Strongly Disagree (1)	(2)	Neutral (3)	(4)	Strongly Agree (5)
7. I would make use of this guideline in my professional decisions.[1]	0	0	0	1	0
8. I would recommend this guideline for use in practice.[1]	0	0	0	1	0
9. What are the barriers or enablers to the implementation of this guideline report?	Barriers Laboratory: • A variety of testing methods used across the province (RT-PCR, NGS, MFC) with inherent variability in access favouring larger academic centres. • Inter-laboratory quality assurance is not routinely performed (particularly at low MRD levels) but will be essential to ensure the accuracy of testing results and qualification of testing sites • Need for an element of harmonization/standardization required, given discrepancies between NGS and MFC testing between laboratories • Given the technical and professional difficulty of MRD testing, consideration of qualification of laboratories and funding models to support adequate validation, external quality assessment, and clinical testing will be required Clinical: • Lack of availability (and funding) for targeted agents limits the application of MRD-directed therapy to consolidative chemotherapy and alloHCT • Role of clinical trials for novel agents/ regimens to evaluate for survival/quality of				

[3] One reviewer gave comments but no numeric rating for these questions.

	life benefit to MRD-adapted therapy should be strongly emphasized Routine MRD analysis for prognostic or any purpose has not yet been standardized, implemented or routinely funded in Ontario which negates any practical application of the current guideline. This applies especially to ALL where the value of MRD analysis as a prognostic tool has been clearly demonstrated in the literature.

Table 5-4. Summary of the Working Group's responses to comments from Targeted Peer Reviewers

Comments	Responses
1. Well organized with clear delineation of recommendations, and pertinent qualifications and rationale for the statements.	Thank you for your comment.
2. Recommendation 1: It may be of value to indicate that targeted agents have received United States Food and Drug Administration approval for use based on MRD testing (i.e., blinatumomab for B-ALL with MRD >10^{-3}). Additional studies could be noted: Bassan R et al 2009 113:4153-4162. Ribera JM et al. J Clin Onc 2014. 32:1595-1604. Gokbuget N et al. Blood 2012. 120: 1868-1876. Gokbuget N et al 2019. Hematology 24(1): 337-348.	These studies were prognostic and therefore did not meet the inclusion criteria for the review. We have added the qualifying statement that MRD may be required for eligibility for specific therapy, in which case testing should be done when consideration is made for therapy and as per treatment guidelines
3. Recommendation may warrant further discussion of evidence for transplantation in CR1 for CBF and NPM1-mut 'favourable' AML with positive MRD. Consideration of studies by Rubnitz JE at el. (Lancet Oncol 2010. 11(6): 543-552) and Zhu HH et al. (Blood 2013. 121(20): 4056-4062) may provide further support for role of addition of gemtuzumab ozogamicin or transplantation	The Working Group has included the following in the discussion, "In the case of APL and CBF-leukemia, this testing has been available and used for monitoring for many years. In some cases, it has been shown that re-treating patients who have become MRD positive after being negative (so-called molecular relapse) post treatment and prior to a morphological relapse, has resulted in improved outcomes, particularly in cases of APL, and CBF-translocation-positive AML." If MRD remains positive at the end of therapy or becomes positive during monitoring with rising, serially tested samples, consideration of an allogeneic transplantation should be made. If an effective and relatively non-toxic therapy is available that can result in a return to MRD negativity, it should be considered prior to allogenic transplant [22,53]. This is included in the Discussion. As there were adult studies, pediatric studies were excluded per the inclusion/exclusion criteria. There were inconsistencies within the Zhu et al article and therefore a decision to exclude it was made.

4. Recommendation 2: Given the recognition of the significant adverse prognostic effect of MRD+ complete remission, a qualifying statement indicating value of clinical trials for targeted therapies (if any available) may present benefit. It may be of value to include qualifying statements (similar to Recommendation 1) indicating that MRD assessment is optimally performed on marrow, and that timing/frequency of MRD assessment is variable and underlying biology must be considered (e.g., high-risk APL vs. AML with inv(16)).”	Qualifying statements have been added to Recommendation 2.
5. Given the emerging nature of targeted therapies and studies evaluating MRD-directed application of targeted agents, it is expected that clear and high-quality evidence remains lacking. With regards to Recommendation2, specific reference to some evidence in CBF and NPM1-mutated AML may be considered based on the presence of some evidence.	This is included in the Introduction and Discussion, and some additional references have been added in the Discussion of this topic.
6. An excellent review of current evidence, the authors rightly note that clear and high-quality evidence for actionability of MRD testing remains lacking.	Thank you for the comment
7. While the Guideline Objective 2 highlights the assessment of prognostic value, the topic is not reviewed in depth (rather discusses as standard-of-care with reference to ELN and NCCN guidelines) – the document appears highly focused on the predictive value, which may be well warranted, though consideration of expanding review of the prognostic value or clearly restricting the guideline to predictive applications may be warranted.	This objective has been revised to reflect that prognosis was outside the scope. The prognostic value of MRD testing is accepted as a standard of care in this guideline. The preamble states, “In acute leukemias, the use of MRD testing using bone marrow or blood is routine, as these assays provide prognostic information for clinicians and patients.”
8. Although the recommendations acknowledge that current standard of care for optimal prognostication of ALL should include MRD analysis, routine MRD analysis for prognostic purposes has not yet been developed, standardized, implemented or funded in Ontario. For this reason, I find the current guideline recommendations misrepresentative of what is current reality in Ontario. This point is also highlighted in comment number 6 from the Expert Panel in section 5 of the draft document Routine MRD analysis for prognostic- or any purpose has not yet been standardized, implemented, or routinely funded in Ontario which negates any practical application of the current guideline – this applies	There is some funding in place for MRD, but it is not well fleshed out. University Health Network offers this testing and OH (CCO) recognizes that it is being done and is being paid for acute leukemia patients. However, this document is not intended to justify the current funding situation, nor to recommend which or how many laboratories should do testing or what methods they should use.

especially to ALL where the value of MRD analysis as a prognostic tool has been clearly demonstrated in the literature. For answers to general questions 8 and 9, it is not possible to comment whether I would make use of this guideline in my professional decisions or recommend the guideline for use in practice, when MRD analysis is not yet standard practice in Ontario even as a prognostic tool.	
9. While I agree with the recommendations and the interpretation of available evidence that is presented in this guideline, it must be acknowledged that in ALL, it is not realistic/practical to wait for adult ALL RCTs in order to decide on the value of MRD in guiding therapy – the disease is rare in adults and therefore confirmatory randomized controlled studies of similar power to those conducted in pediatric ALL is not possible with adult ALL.	This is certainly a valid point. Further reviews or guidelines on this topic may also include real world (retrospective) data; this was not included in the scope and protocol as determined at the onset of this work.

Professional Consultation

Feedback was obtained through a brief online survey of healthcare professionals and other stakeholders who are the intended users of the guideline. Clinicians with an interest in hematology, leukemia and pathology and lab medicine in the PEBC database were contacted by email to inform them of the survey. Two hundred twenty-four professionals were contacted, all of which practice in Ontario. Twenty-four (10.4%) responses were received. Ten stated that they did not have interest in this area and one stated they were now retired. The results of the feedback survey from 13 people are summarized in Table 5-5. The main comments from the consultation and the Working Group's responses are summarized in Table 5-6.

Table 5-5. Responses to four items on the professional consultation survey

	N=13 (5.8%)				
General Questions: Overall Guideline Assessment	Lowest Quality (1)	(2)	(3)	(4)	Highest Quality (5)
1. Rate the overall quality of the guideline report.	0	0	0	8	5
	Strongly Disagree (1)	(2)	(3)	(4)	Strongly Agree (5)
2. I would make use of this guideline in my professional decisions.	0	0	2	5	6
3. I would recommend this guideline for use in practice.	0	0	0	5	8
4. What are the barriers or enablers to the implementation of this guideline report?	Barriers • Financial impact				

	• Staffing • Access to validated testing • Awareness in the community of MRD testing • Not all centres have the comprehensive panels and expertise to determine MRD • The use of MRD testing in practice are 'ahead' of published evidence - this will be an ongoing issue as laboratories adopt diagnostics ahead of published evidence showing their clinical utility. We need to find a way to rapidly introduce technology with a system to support ongoing clinical evaluation and feedback. • Need to develop standardized structured synoptic reporting • Strength of pediatric literature but lack of ability to confidently translate to adult population

Table 5-6. Summary of the Working Group's responses to comments from professional consultants

Comments	Responses
1. Within the Key Evidence for Recommendation 1, rather than saying MRD low and standard risk, state what test was used and the level of detection and the level considered to show low risk. Page 4-need to state what MRD standard risk is.	The Working Group has added this information to the Key Evidence for Recommendation 1.
2. Page 27. Cannot extrapolate one study to another if induction or other factors vary.	The methods section (see Synthesizing the Evidence) indicates a meta-analysis was not planned due to heterogeneity of the trials. We agree, in principal, that one must be cautious in extrapolating results from one type of regimen to another. There are no two studies that have the exact same induction and consolidation regimens making it difficult to compare results across studies.
3. Page 27. Double induction is a problem as it is not standard of care in Ontario; the protocol in Jongen-Lavrencic et al, 2018 does not apply to Ontario	The scope of the review was not restricted to protocols used in Ontario.
4. Need to define prognostic and predictive within the guideline	Predictive (markers; clinical utility) refers to the whether a patient is likely to benefit from a treatment. Prognostic (prognosis) is used to refer to a patient's probable long-term outcome (untreated or with a standard treatment). A

	footnote has been added to the start of Section 2.
5. Reviewers noted this guideline was well written, comprehensive and through.	Thank you for your comment.
6. Comment regarding the introduction to the systematic review. Can likely get below 0.01%	Maybe, but this is the currently common and most frequent value.
7. 6.Comment regarding the introduction to the systematic review. Need to distinguish clonal hematopoiesis of indeterminate potential abnormalities and consider their importance	This was addressed in the introduction. "A further challenge for 'myeloid gene' variant monitoring for MRD in AML is that some of the variants identified in AML have also been identified in normal people with age-related clonal hematopoiesis [25,26]. Because of this, it is not always clear whether a variant identified while monitoring MRD is really indicative of the AML clone, or rather represents residual/new clonal hematopoiesis."
8. 6. Comment regarding the introduction to the systematic review. Need to discuss/consider CBF in addition to APL 12	This was addressed in the introduction. "Variants that do meet this criterion include promyelocytic leukemia/retinoic acid receptor alpha (PML/RARA) and the other variants currently listed with the World Health Organization as "recurrent genetic abnormalities" [23]. As a group, these account for approximately 40% to 50% of AML cases [24]."

CONCLUSION

The final guideline recommendations contained in Section 2 and summarized in Section 1 reflect the integration of feedback obtained through the external review processes with the document as drafted by the GDG Working Group and approved by the GDG Expert Panel and the PEBC RAP.

References

1. Dohner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Buchner T, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood. 2017;129(4):424-47.

2. Alvarnas JC, Brown PA, Aoun P, Ballen KK, Barta SK, Borate U, et al. Acute lymphoblastic leukemia, version 2.2015. J Natl Compr Canc Netw. 2015;13(10):1240-79.

3. Vora A, Goulden N, Wade R, Mitchell C, Hancock J, Hough R, et al. Treatment reduction for children and young adults with low-risk acute lymphoblastic leukaemia defined by minimal residual disease (UKALL 2003): a randomised controlled trial. Lancet Oncol. 2013;14(3):199-209.

4. Vora A, Goulden N, Mitchell C, Hancock J, Hough R, Rowntree C, et al. Augmented post-remission therapy for a minimal residual disease-defined high-risk subgroup of children and young people with clinical standard-risk and intermediate-risk acute lymphoblastic leukaemia (UKALL 2003): a randomised controlled trial. Lancet Oncol. 2014;15(8):809-18.

5. Eiser C, Stride CB, Vora A, Goulden N, Mitchell C, Buck G, et al. Prospective evaluation of quality of life in children treated in UKALL 2003 for acute lymphoblastic leukaemia: a cohort study. Pediatr Blood Cancer. 2017;64(11).

6. Schrappe M, Bleckmann K, Zimmermann M, Biondi A, Moricke A, Locatelli F, et al. Reduced-intensity delayed intensification in standard-risk pediatric acute lymphoblastic leukemia defined by undetectable minimal residual disease: results of an international randomized trial (AIEOP-BFM ALL 2000). J Clin Oncol. 2018;36(3):244-53.

7. Conter V, Bartram CR, Valsecchi MG, Schrauder A, Panzer-Grumayer R, Moricke A, et al. Molecular response to treatment redefines all prognostic factors in children and adolescents with B-cell precursor acute lymphoblastic leukemia: results in 3184 patients of the AIEOP-BFM ALL 2000 study. Blood. 2010;115(16):3206-14.

8. Dhedin N, Huynh A, Maury S, Tabrizi R, Beldjord K, Asnafi V, et al. Role of allogeneic stem cell transplantation in adult patients with Ph-negative acute lymphoblastic leukemia. Blood. 2015;125(16):2486-96.

9. Bassan R, Spinelli O. Minimal residual disease monitoring in adult ALL to determine therapy. Curr Hematol Malig Rep. 2015;10(2):86-95.

10. Gokbuget N, Dombret H, Giebel S, Bruggemann M, Doubek M, Foa R, et al. Minimal residual disease level predicts outcome in adults with Ph-negative B-precursor acute lymphoblastic leukemia. Hematology. 2019;24(1):337-48.

11. Gokbuget N, Kneba M, Raff T, Trautmann H, Bartram CR, Arnold R, et al. Adult patients with acute lymphoblastic leukemia and molecular failure display a poor prognosis and are candidates for stem cell transplantation and targeted therapies. Blood. 2012;120(9):1868-76.

12. Hourigan CS, Dillon LW, Gui G, Logan BR, Fei M, Ghannam J, et al. Impact of conditioning intensity of allogeneic transplantation for acute myeloid leukemia with genomic evidence

of residual disease. DOI https://doi.org/10.1200/JCO.19.03011, Epub: 2019 Dec 20. J Clin Oncol. 2020;38.

13. Balsat M, Renneville A, Thomas X, de Botton S, Caillot D, Marceau A, et al. Postinduction minimal residual disease predicts outcome and benefit from allogeneic stem cell transplantation in acute myeloid leukemia with NPM1 mutation: a study by the Acute Leukemia French Association Group. J Clin Oncol. 2017;35(2):185-93.

14. Freeman SD, Hills RK, Virgo P, Khan N, Couzens S, Dillon R, et al. Measurable residual disease at induction redefines partial response in acute myeloid leukemia and stratifies outcomes in patients at standard risk without NPM1 mutations. J Clin Oncol. 2018;36(15):1486-97.

15. Jongen-Lavrencic M, Grob T, Hanekamp D, Kavelaars FG, Al Hinai A, Zeilemaker A, et al. Molecular minimal residual disease in acute myeloid leukemia. N Engl J Med. 2018;378(13):1189-99.

16. Chen X, Xie H, Wood BL, Walter RB, Pagel JM, Becker PS, et al. Relation of clinical response and minimal residual disease and their prognostic impact on outcome in acute myeloid leukemia. J Clin Oncol. 2015;33(11):1258-64.

17. Browman GP, Levine MN, Mohide EA, Hayward RS, Pritchard KI, Gafni A, et al. The practice guidelines development cycle: a conceptual tool for practice guidelines development and implementation. J Clin Oncol. 1995;13(2):502-12.

18. Browman GP, Newman TE, Mohide EA, Graham ID, Levine MN, Pritchard KI, et al. Progress of clinical oncology guidelines development using the practice guidelines development cycle: the role of practitioner feedback. J Clin Oncol. 1998;16(3):1226-31.

19. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. CMAJ. 2010;182(18):E839-42.

20. Statistics Canada. Table 13-10-0111-01. Number and rates of new cases of primary cancer, by cancer type, age group and sex [webpage]. Ottawa: Statistics Canada. Cited 2019 Jun 12. DOI: https://doi.org/10.25318/1310011101-eng.

21. Estey EH. Acute myeloid leukemia: 2019 update on risk-stratification and management. Am J Hematol. 2018;93(10):1267-91.

22. Schuurhuis GJ, Heuser M, Freeman S, Bene MC, Buccisano F, Cloos J, et al. Minimal/measurable residual disease in AML: a consensus document from the European LeukemiaNet MRD Working Party. Blood. 2018;131(12):1275-91.

23. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. Blood. 2016;127(20):2391-405.

24. Bullinger L, Dohner K, Dohner H. Genomics of acute myeloid leukemia diagnosis and pathways. J Clin Oncol. 2017;35(9):934-46.

25. Short NJ, Ravandi F. How close are we to incorporating measurable residual disease into clinical practice for acute myeloid leukemia? Haematologica. 2019;104(8):1532-41.

26. Steensma DP, Bejar R, Jaiswal S, Lindsley RC, Sekeres MA, Hasserjian RP, et al. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. Blood. 2015;126(1):9-16.

27. Health Quality Ontario. Minimal residual disease evaluation in childhood acute lymphoblastic leukemia: a clinical evidence review. Ont Health Technol Assess Ser. 2016;16(7):1-52.

28. Estey E, Gale RP. How good are we at predicting the fate of someone with acute myeloid leukaemia? Leukemia. 2017;31(6):1255-8.

29. Papaemmanuil E, Gerstung M, Bullinger L, Gaidzik VI, Paschka P, Roberts ND, et al. Genomic classification and prognosis in acute myeloid leukemia. N Engl J Med. 2016;374(23):2209-21.

30. Campana D, Pui CH. Minimal residual disease-guided therapy in childhood acute lymphoblastic leukemia. Blood. 2017;129(14):1913-8.

31. Biondi A, Schrappe M, De Lorenzo P, Castor A, Lucchini G, Gandemer V, et al. Imatinib after induction for treatment of children and adolescents with Philadelphia-chromosome-positive acute lymphoblastic leukaemia (EsPhALL): a randomised, open-label, intergroup study. Lancet Oncol. 2012;13(9):936-45.

32. Roberts KG, Pei D, Campana D, Payne-Turner D, Li Y, Cheng C, et al. Outcomes of children with BCR-ABL1-like acute lymphoblastic leukemia treated with risk-directed therapy based on the levels of minimal residual disease. J Clin Oncol. 2014;32(27):3012-20.

33. Morita K, Kantarjian HM, Wang F, Yan Y, Bueso-Ramos C, Sasaki K, et al. Clearance of somatic mutations at remission and the risk of relapse in acute myeloid leukemia. J Clin Oncol. 2018;36(18):1788-97.

34. Seftel MD, Neuberg D, Zhang MJ, Wang HL, Ballen KK, Bergeron J, et al. Pediatric-inspired therapy compared to allografting for Philadelphia chromosome-negative adult ALL in first complete remission. Am J Hematol. 2016;91(3):322-9.

35. Yanada M, Matsuo K, Suzuki T, Kiyoi H, Naoe T. Prognostic significance of FLT3 internal tandem duplication and tyrosine kinase domain mutations for acute myeloid leukemia: a meta-analysis. Leukemia. 2005;19(8):1345-9.

36. Grimwade D, Jovanovic JV, Hills RK, Nugent EA, Patel Y, Flora R, et al. Prospective minimal residual disease monitoring to predict relapse of acute promyelocytic leukemia and to direct pre-emptive arsenic trioxide therapy. J Clin Oncol. 2009;27(22):3650-8.

37. Chendamarai E, Balasubramanian P, George B, Viswabandya A, Abraham A, Ahmed R, et al. Role of minimal residual disease monitoring in acute promyelocytic leukemia treated with arsenic trioxide in frontline therapy. Blood. 2012;119(15):3413-9.

38. Yin JA, O'Brien MA, Hills RK, Daly SB, Wheatley K, Burnett AK. 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. 2012;120(14):2826-35.

39. Jourdan E, Boissel N, Chevret S, Delabesse E, Renneville A, Cornillet P, et al. Prospective evaluation of gene mutations and minimal residual disease in patients with core binding factor acute myeloid leukemia. Blood. 2013;121(12):2213-23.

40. Ivey A, Hills RK, Simpson MA, Jovanovic JV, Gilkes A, Grech A, et al. Assessment of minimal residual disease in standard-risk AML. N Engl J Med. 2016;374(5):422-33.

41. Willekens C, Blanchet O, Renneville A, Cornillet-Lefebvre P, Pautas C, Guieze R, et al. 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. 2016;101(3):328-35.

42. Miyamoto T, Weissman IL, Akashi K. AML1/ETO-expressing nonleukemic stem cells in acute myelogenous leukemia with 8;21 chromosomal translocation. Proc Natl Acad Sci U S A. 2000;97(13):7521-6.

43. Koreth J, Schlenk R, Kopecky KJ, Honda S, Sierra J, Djulbegovic BJ, et al. Allogeneic stem cell transplantation for acute myeloid leukemia in first complete remission: systematic review and meta-analysis of prospective clinical trials. JAMA. 2009;301(22):2349-61.

44. Dombret H, Gardin C. An update of current treatments for adult acute myeloid leukemia. Blood. 2016;127(1):53-61.

45. Sung PJ, Luger SM. Minimal residual disease in acute myeloid leukemia. Curr Treat Options Oncol. 2017;18(1):1.

46. Brands-Nijenhuis AV, Labopin M, Schouten HC, Volin L, Socie G, Cornelissen JJ, et al. Monosomal karyotype as an adverse prognostic factor in patients with acute myeloid leukemia treated with allogeneic hematopoietic stem-cell transplantation in first complete remission: a retrospective survey on behalf of the ALWP of the EBMT. Haematologica. 2016;101(2):248-55.

47. Cassier PA, Castets M, Belhabri A, Vey N. Targeting apoptosis in acute myeloid leukaemia. Br J Cancer. 2017;117(8):1089-98.

48. Eisfeld AK, Kohlschmidt J, Mrozek K, Blachly JS, Walker CJ, Nicolet D, et al. Mutation patterns identify adult patients with de novo acute myeloid leukemia aged 60 years or older who respond favorably to standard chemotherapy: an analysis of Alliance studies. Leukemia. 2018;32(6):1338-48.

49. Maganti HB, Jrade H, Cafariello C, Manias Rothberg JL, Porter CJ, Yockell-Lelievre J, et al. Targeting the MTF2-MDM2 axis sensitizes refractory acute myeloid leukemia to chemotherapy. Cancer Discov. 2018;8(11):1376-89.

50. Paietta E. Consensus on MRD in AML? Blood. 2018;131(12):1265-6.

51. Terwijn M, van Putten WL, Kelder A, van der Velden VH, Brooimans RA, Pabst T, et al. High prognostic impact of flow cytometric minimal residual disease detection in acute myeloid leukemia: data from the HOVON/SAKK AML 42A study. J Clin Oncol. 2013;31(31):3889-97.

52. Short NJ, Jabbour E, Albitar M, de Lima M, Gore L, Jorgensen J, et al. Recommendations for the assessment and management of measurable residual disease in adults with acute lymphoblastic leukemia: a consensus of North American experts. Am J Hematol. 2019;94(2):257-65.

53. Freeman SD, Hourigan CS. MRD evaluation of AML in clinical practice: are we there yet? Hematology Am Soc Hematol Educ Program. 2019;2019(1):557-69.

Appendix 1: Affiliations and Conflict of Interest Declarations

Table A1-1: Working Group members

Name	Affiliation	Conflict of Interest
Harriet Feilotter Molecular Geneticist	Kingston General Hospital Kingston, ON	Received ≥$5000 in a single year to act in a consulting capacity from Roche Diagnostics, Janssen and Bristol-Myers Squibb; received research support from Astra Zeneca for an unrelated topic
Christopher Howlett Pathologist	London Health Science Centre London, ON	No conflict of interest declared
Rebecca McClure* Pathologist	Health Sciences North Sudbury, ON	No conflict of interest declared
Aaron Pollett Pathologist	Mount Sinai Hospital Toronto, ON	No conflict of interest declared
Catherine Ross Hematopathologist	Juravinski Hospital Hamilton, ON	No conflict of interest declared
Mitchell Sabloff Hematologist	Ottawa Hospital Ottawa, ON	Received an unrestricted research grant from Sanofi that is used to fund the tissue bank
Andre Schuh Hematologist	Princess Margaret Cancer Centre Toronto, ON	No conflict of interest declared
Duvaraga Sivajohanathan Health Research Methodologist	Program in Evidence-Based Care, OH (CCO), McMaster University Hamilton, ON	No conflict of interest declared

*withdrew from Working Group for personal reasons subsequent to Internal Review

Table A1-2: Report Approval Panel

Name	Affiliation	Conflict of Interest
Melissa Brouwers Professor & Director	School of Epidemiology and Public Health University of Ottawa Ottawa, ON	No conflict of interest declared
Willian (Bill) Evans Medical Oncologist	Oncosynthesis Consulting Inc.	No conflict of interest declared
Jonathan Sussman Scientific Director Radiation Oncologist	Department of Oncology, Juravinski Cancer Centre; Program in Evidence-Based Care, OH (CCO), McMaster University Hamilton, ON	No conflict of interest declared

Table A1-3: Expert Panel members

Name	Affiliation	Conflict of Interest
Philip Berardi Hematopathologist	Ottawa Hospital Ottawa, ON	Has received $500 or more as Advisory Board member for Astellas Pharma, Novartis and Jannsen - all not relevant to MRD testing
Chris Bredeson Hematologist	Ottawa Hospital Ottawa, ON	No conflict of interest declared
Rena Buckstein Hematologist	Odette Cancer Centre Toronto, ON	Has received $500 or more as an Advisory Board member; has received research support as a principal or co-investigator for an MDS registry from Celgene and Otsuka; has been a principal investigator of trials testing FLT3 inhibitors and azacidine
Janet Dancey Medical Oncologist	Cancer Centre of Southeastern Ontario Queen's University Kingston, ON	Has received research funding as a principal or co-investigator to conduct trials run by CCTG, Roche, AstraZeneca, Pfizer, and BMS; is the Director of CCTG which has received $5000 or more from AstraZeneca, BMS, Roche, Pfizer, Celgene, Janssen, Senhwa and others
Andrea Eisen Medical Oncologist	Odette Cancer Centre Toronto, ON	Has received a grant from Genomic Health for a research study as a principal or co-investigator; has been a principal investigator in a clinical trial of Oncotype DX in node positive breast cancer; has provided guidance regarding the objects of study to OHTAC
Sathish Kumar Gopalakrishnan Hematologist	Northeast Cancer Centre Sudbury, ON	No conflict of interest declared
Kang Howson-Jan Hematologist	London Regional Cancer Centre London, ON	No conflict of interest declared
Suzanne Kamel-Reid Geneticist	University Health Network Toronto, ON	No conflict of interest declared
Rouslan Kotchetkov Hematologist	Royal Victoria Regional Heath Centre Barrie, ON	No conflict of interest declared
Tom Kouroukis Hematologist	Juravinski Cancer Centre Hamilton, ON	No conflict of interest declared
Nicole Laferriere Hematologist	Regional Cancer Care Northwest Thunder Bay, ON	No conflict of interest declared
Brian Leber Hematologist	Juravinski Cancer Centre	No conflict of interest declared

	Hamilton, ON	
Bryan Lo Molecular Geneticist	Ottawa Hospital Ottawa, ON	Has received $500 or more in participating in advisory boards for Novartis, AstraZeneca, Roche, Bayer and Pfizer; has stocks valued greater than $1000 at Merck and Johnson & Johnson; has received start-up research funds in the form of an unrestricted grant from Amgen to the Ottawa Hospital Foundation; has been the principal or co-investigator for the OCTANE and EXACTIS trials
Elizabeth McCready Geneticist	McMaster University Hamilton, ON	Has received $500 or more for the Amgen Canada National Leukemia Consultancy Meeting in Toronto on November 2, 2018
Jean McGowan-Jordan Geneticist	Ottawa Hospital Ottawa, ON	No conflict of interest declared
Trevor Pugh Molecular Geneticist	Princess Margaret Cancer Centre University of Toronto Toronto, ON	Has received $500 or more to act in a consulting capacity for Merck, Prosigna, Chrysalis Biomedical Advisors and DynaCare; has received a training grant for a postdoctoral fellowship from Boehringer Ingelheim from 2015 to 2016
Michael Rutherford Molecular Geneticist	Hôpital Régional de Sudbury Regional Hospital Sudbury, ON	No conflict of interest declared
Marsha Speevak Geneticist	Credit Valley Hospital University of Toronto Toronto, ON	No conflict of interest declared
Anne Tierens Hematopathologist	University Health Network Toronto, ON	No conflict of interest declared
Hong Wang Geneticist	North York General Hospital Toronto, ON	No conflict of interest declared
Karen Yee Hematologist	Princess Margaret Cancer Centre Toronto, ON	Has received $500 or more as an Advisory Board member for Celgene, Novartis, Otsuka, Pfizer, Roche and Shire; has received $500 or more in travel support by Astex; has had managerial responsibility for an organization receiving $5000 or more in a year by Norvartis for organizing a Latin America AML Preceptorship in 2018

Table A1-4: Targeted Peer Reviewers

Name	Affiliation	Conflict of Interest
Graeme Quest Hematopathologist	Kingston General Hospital Kingston, ON	No conflict of interest declared
Jill Fulcher Hematologist	Ottawa Hospital Ottawa, ON	Has received $500 or more in a year from Amgen for work as a consultant and from Novartis and Jazz Pharmaceuticals for participation on Advisory Board; has received the Jazz Pharmaceutical Educational Grant as a principal or co-investigator which has been paid into the OHRI cost centre

Appendix 2: Literature Search Strategy

MEDLINE

1 (systematic adj (review: or overview:)).mp. (171430)
2 (meta-analy: or metaanaly:).mp. (186563)
3 (pooled analy: or statistical pooling or mathematical pooling or statistical summar: or mathematical summar: or
quantitative synthes?s or quantitative overview:).mp. (11180)
4 (exp review literature as topic/ or review.pt. or exp review/) and systematic.tw. (135673)
5 (cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinhal or cinahl or science citation index
or scisearch or bids or sigle or cancerlit or pubmed or pub-med or medline or med-line).ab. (209320)
6 (reference list: or bibliograph: or hand-search: or handsearch: or relevant journal: or manual search:).ab.
(42176)
7 or/1-6 (394567)
8 (selection criteria or data extract: or quality assess: or jadad score or jadad scale or methodologic:
quality).ab. (72949)
9 (stud: adj1 select:).ab. (24354)
10 (8 or 9) and review.pt. (45751)
11 7 or 10 (399314)
12 (guideline or practice guideline).pt. (32676)
13 exp consensus development conference/ (11607)
14 consensus/ (11364)
15 (guideline: or recommend: or consensus or standards).ti. (154985)
16 12 or 13 or 14 or 15 (176366)
17 11 or 16 (564867)
18 exp Randomized Controlled Trial/ or Clinical Trial, Phase III/ or Clinical Trial, Phase IV/ or Phase 3 Clinical
Trial/ or Phase 4 Clinical Trial/ or ((exp Clinical Trial/ or Prospective Study/ or Prospective Studies/) and
Random$.tw.) or exp Randomized Controlled Trials as topic/ or Clinical Trials, Phase III as Topic/ or Clinical Trials, Phase IV as Topic/ or exp "Randomized Controlled Trial (Topic)"/ or "Phase 3 Clinical Trial (Topic)"/ or "Phase 4 Clinical Trial (Topic)"/ or ((exp Clinical Trials as Topic/ or exp "Clinical Trial (Topic)"/) and random$.tw.) or Random Allocation/ or Randomization/ or Single-Blind Method/ or Double-Blind Method/ or Single Blind Procedure/ or Double Blind Procedure/ or Triple Blind Procedure/ or Placebos/ or Placebo/ or ((singl$ or doubl$ or tripl$) adj3 (blind$3 or mask$3 or dummy)).tw. or (random$ control$ trial? or rct or phase III or phase IV or phase 3 or phase 4).tw. or (((phase II or phase 2 or clinic$) adj3 trial$) and random$).tw. or (placebo? or (allocat$ adj2 random$)).tw. or (random$ adj3 trial$).mp. or "clinicaltrials.gov".mp. (1013151)
19 (minimal adj2 residual adj2 disease).mp. (6731)
20 (measurable adj2 residual adj2 disease).mp. (140)
21 19 or 20 (6815)
22 exp LEUKEMIA/ (227040)
23 leukemia.mp. (303133)
24 22 or 23 (304854)

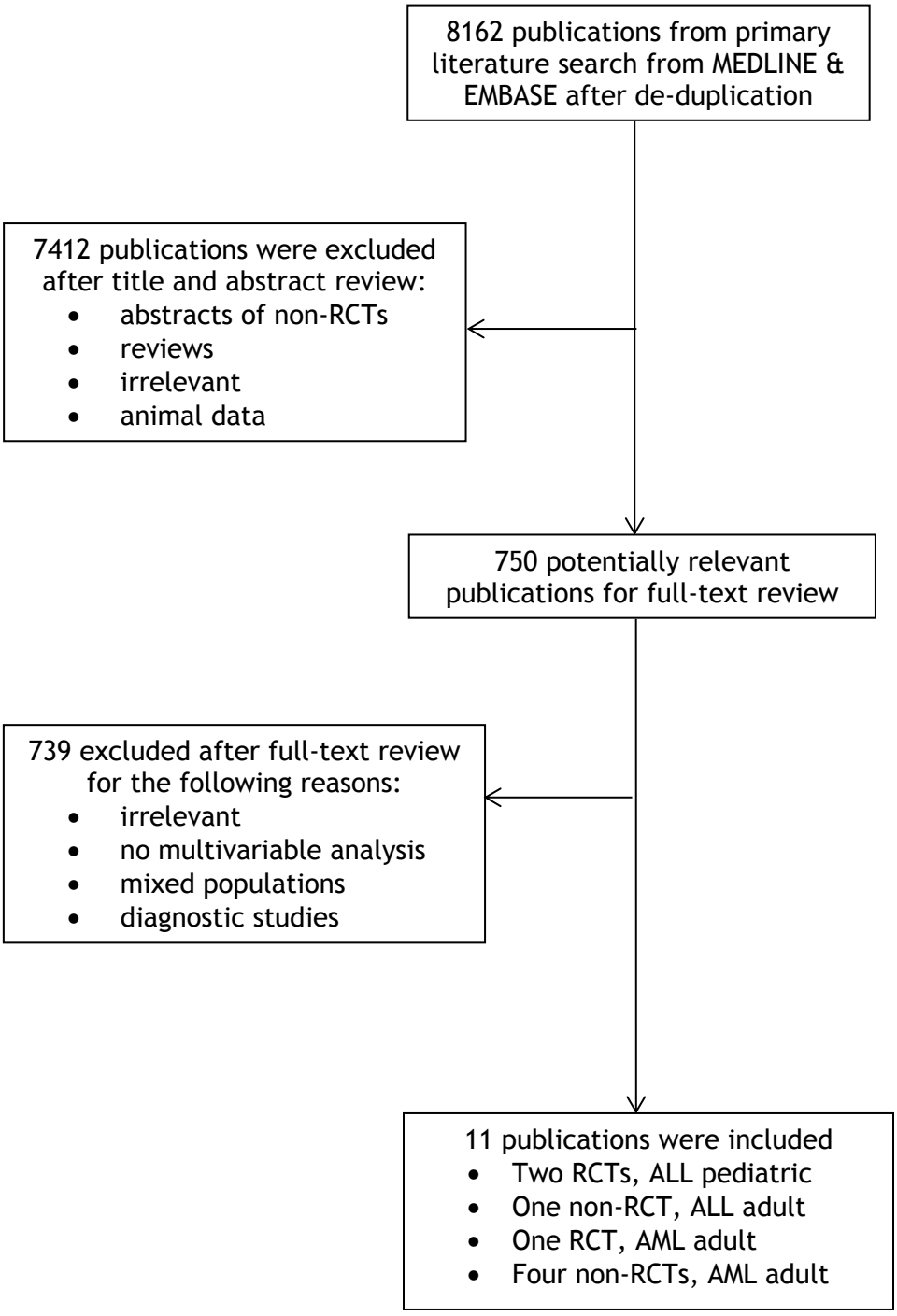
25 17 and 21 and 24 (89)
26 limit 25 to (english language and yr="2013 -Current") (43)
27 18 and 21 and 24 (236)
28 (comment or letter or editorial or news or newspaper article or patient education handout or case reports or historical article).pt. (4140039)
29 27 not 28 (231)
30 exp animals/ not humans/ (4626398)
31 29 not 30 (231)
32 limit 31 to (english language and yr="2000 -Current") (209)
33 21 and 24 (4274)
34 33 not 17 (4185)
35 34 not 28 (3790)
36 35 not 30 (3745)

EMBASE

(systematic adj (review: or overview:)).mp. (210929)
(meta-analy: or metaanaly:).mp. (213677)
(pooled analy: or statistical pooling or mathematical pooling or statistical summar: or mathematical summar: or quantitative synthes?s or quantitative overview:).mp. (13549)
(exp review literature as topic/ or review.pt. or exp review/) and systematic.tw. (147250)
(cochrane or embase or psychlit or psyclit or psychinfo or psycinfo or cinhal or cinahl or science citation index or scisearch or bids or sigle or cancerlit or pubmed or pub-med or medline or med-line).ab. (194587)
(reference list: or bibliograph: or hand-search: or handsearch: or relevant journal: or manual search:).ab. (43555)
(selection criteria or data extract: or quality assess: or jadad score or jadad scale or methodologic:quality).ab. (74879)
(stud: adj1 select:).ab. (24560)
(7 or 8) and review.pt. (34907)
or/1-6 (438901)
9 or 10 (443213)
consensus development conference/ (21287)
practice guideline/ (322257)
*consensus development/ or *consensus/ (7707)
*standard/ (3650)
(guideline: or recommend: or consensus or standards).kw. (41308)
(guideline: or recommend: or consensus or standards).ti. (169764)
or/12-17 (451307)
11 or 18 (863868)
exp Randomized Controlled Trial/ or Clinical Trial, Phase III/ or Clinical Trial, Phase IV/ or Phase 3 Clinical Trial/ or Phase 4 Clinical Trial/ or ((exp Clinical Trial/ or Prospective Study/ or Prospective Studies/) and Random$.tw.) or exp Randomized Controlled Trials as topic/ or Clinical Trials, Phase III as Topic/ or Clinical Trials, Phase IV as Topic/ or exp "Randomized Controlled Trial (Topic)"/ or "Phase 3 Clinical Trial (Topic)"/ or "Phase 4 Clinical Trial (Topic)"/ or ((exp Clinical Trials as Topic/ or exp "Clinical Trial (Topic)"/) and random$.tw.) or Random Allocation/ or Randomization/ or Single-Blind Method/ or Double-Blind Method/ or Single Blind Procedure/ or Double Blind Procedure/ or Triple Blind Procedure/ or Placebos/ or Placebo/ or ((singl$ or doubl$ or tripl$) adj3 (blind$3 or mask$3 or dummy)).tw. or (random$ control$ trial? or rct or phase III or phase IV or phase 3 or phase 4).tw. or (((phase II or phase 2 or clinic$) adj3

trial$) and random$).tw. or (placebo? or (allocat$ adj2 random$)).tw. or (random$ adj3 trial$).mp. or "clinicaltrials.gov".mp. (1204508)
(minimal adj2 residual adj2 disease).mp. (23482)
(measurable adj2 residual adj2 disease).mp. (103)
21 or 22 (23506)
exp leukemia/ (256053)
leukemia.mp. (392135)
24 or 25 (393915)
19 and 23 and 26 (267)
limit 27 to (english language and yr="2013 -Current") (151)
23 and 26 (8962)
29 not 19 (8695)
(editorial or note or letter or short survey).pt. or letter/ or case study/ (2575305)
30 not 31 (8054)
animal/ not human/ (1003160)
32 not 33 (8046)
limit 34 to (english language and yr="2000 -Current") (6639)

Appendix 3: PRISMA Flow Diagram



Appendix 4. Quality Assessment and Risk of Bias

Table A4-1: Quality assessment of included RCTs

Study	Primary outcome	Randomization details	Statistical power and required sample size	ITT analysis	Baseline characteristics balanced	Loss to follow-up (# of pts)	Withdrawals	Industry funding	Terminated early
Acute lymphoblastic leukemia									
Vora et al (2013) (2014) [3,4] UKALL 2003	EFS	Randomized centrally by a computer and stratified by MRD result, sex, age, and white blood cell count at diagnosis by method of minimization	80% power to detect a reduction in 5-year EFS in the group given one DI course from 95% to 88%; 400 MRD low-risk patients 80% power to detect a 10% improvement in EFS (80% to 90%) in the intensification group; 450 MRD high-risk patients	Yes	Yes	None	None	No	No
Schrappe et al (2018) [6] AEIOP-BFM ALL 2000	DFS	Randomized centrally by each country's data centre in accordance with random permuted blocks and stratified by allocation to a preceding random assignment (dexamethasone vs. prednisone) and treatment centre	90% power to assess non-inferiority ($\Delta<4\%$) under the assumption of a 96% 4-year DFS in the reference arm; 1024 patients	Yes	Yes	None	None	No	No

Abbreviations: DFS: disease-free survival; DI: delayed intensification; EFS: event-free survival; ITT: intention to treat; MRD: minimal residual disease; NR: not reported; OS: overall survival; pts: patients; RCT: randomized controlled trial

Table A4-2: Risk of Bias for Included Randomized Controlled Trials

Trial	SELECTION BIAS		PERFORMANCE BIAS	DETECTION BIAS	ATTRITION BIAS	REPORTING BIAS	OTHER BIAS
	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective outcome reporting	Other sources of bias
Acute lymphoblastic leukemia							
Vora et al (2013)(2014) [3,4] UKALL 2003	+	+	-	-	+	+	+
Schrappe et al (2018) [6] AEIOP-BFM ALL 2000	+	?	-	-	+	+	+

Table A4-3: Risk of Bias for Included Non-Randomized Studies Assessed Using Cochrane's ROBIN-I

Study	Bias due to confounding	Bias in selection of participants into the study	Bias in classification of interventions	Bias due to departures from intended interventions	Bias due to missing data	Bias in measurement of outcomes	Bias in selection of the reported result	Overall
Acute lymphoblastic leukemia								
Dhedin et al (2015) [8]	Moderate	Low	Low	Moderate	Low	Low	Low	Moderate
Acute myeloid leukemia								
Balsat et al (2017) [13]	Moderate	Low	Low	Moderate	Low	Low	Low	Moderate
Freeman et al (2018) [14]	Moderate	Low	Low	Moderate	Low	Low	Low	Moderate
Jongen-Lavrencic et al (2018) [15]	Moderate	Low	Low	Moderate	Low	Low	Low	Moderate
Chen et al (2015) [16]	Moderate	Low	Low	Moderate	Low	Low	Low	Moderate

Ontario Health
Cancer Care Ontario

Section 6: Document Assessment and Review Results

Minimal Residual Disease Testing in Acute Leukemia

Document Assessment and Review

M. Sabloff, H. Feilotter, D. Sivajohanathan and Members of the Minimal Residual Disease Testing Guideline Development Group

Septemer 27, 2022

The 2020 guideline recommendations are

ENDORSED

This means that the recommendations are still current and relevant for decision making

OVERVIEW

The original version of this guidance document was released by Ontario Health (Cancer Care Ontario)'s Program in Evidence-based Care in March 2020.

In February 2022, this document was assessed in accordance with the PEBC Document Assessment and Review Protocol and was determined to require a review. As part of the review, a PEBC methodologist (DS) conducted an updated search of the literature. Two clinical experts (MS and HF) reviewed and interpreted the new eligible evidence and proposed the existing recommendations could be endorsed. The Members of the Expert Panel of the Minimal Residual Disease Testing Guideline Development Group (GDG) (Appendix 1) endorsed the recommendations found in Section 1 (Clinical Practice Guideline) in September 2022.

Questions considered

1. What benefit to clinical management does minimal residual disease (MRD) testing contribute to in the treatment of patients with acute lymphoblastic leukemia (ALL)?
2. What benefit to clinical management does MRD testing contribute to in the treatment of patients with acute myeloid leukemia (AML)?

Literature Search and New Evidence

The new search (January 2019 to March 2022) yielded three predictive studies for AML. No new studies meeting the inclusion criteria were found for ALL. An additional search for ongoing studies on clinicaltrials.gov yielded no potentially relevant ongoing randomized contolled trials (RCTs). Brief results of these searches are shown in the Document Review Tool.

The preamble of the original guideline refers to the ELN guidelines published in 2017 for the diagnosis and management of AML. A new version of this guideline was published in 2022 and this reference has been updated to reflect this.

Impact on the Guideline and Its Recommendations
The new data support existing recommendations. Hence, the members of the Expert Panel of the Minimal Residual Disease Testing GDG ENDORSED the 2020 recommendations on minimal residual disease testing for acute leukemia.

Ontario Health
Cancer Care Ontario

Document Review Tool

Number and Title of Document under Review	MOTAC-6: Minimal Residual Disease Testing in Acute Leukemia
Original Report Date	March 10, 2020
Date Assessed (by DSG or Clinical Program Chairs)	February 17, 2022
Health Research Methodologist	Duvaraga Sivajohanathan
Clinical Expert	Dr. Mitchell Sabloff & Dr. Harriet Feilotter
Approval Date and Review Outcome (once completed)	September 27, 2022 ENDORSE

Original Question(s):
1. What benefit to clinical management does MRD testing contribute to in the treatment of patients with ALL?

2. What benefit to clinical management does MRD testing contribute to in the treatment of patients with AML?

Target Population: Adult patients with a diagnosis of acute leukemia (i.e., AML or ALL).

Study Selection Criteria:

Inclusion Criteria
- RCTs (if no RCTs then non-randomized comparative studies) with ≥30 participants and if no RCTs or non-randomized comparative studies then single-arm studies with ≥100 participants where confounders are controlled for; and
- Studies assessing adult patients with a diagnosis of acute leukemia (i.e., AML or ALL); and
- Studies using MRD testing and reporting the following clinical outcomes: OS, EFS, relapse rate, adverse events, and quality of life.

Exclusion Criteria
- Abstracts of non-randomized studies (single-arm clinical trials, case series, etc.); or
- Abstracts of interim analyses; or
- Papers or abstracts not available in English; or
- Letters and editorials that reported clinical trial outcomes; or
- Papers and abstracts published before 2000.

Search Details:
- Full Search is in Appednix 2 and dates were modified to January 2019 to March 17, 2022 (MEDLINE, EMBASE)

- January 2019 to March 2022 (clinicaltrials.gov, and abstracts from the Summit of American Society of Hematology, Congress of European Hematology Association, and Society of Hematopathology)

Summary of new evidence:

There was a total of 248 hits (after deduplication) of guidelines and systemic reviews from MEDLINE and EMBASE; none met the inclusion criteria for this guideline.

There was a total of 491 hits for primary literature after deduplication from MEDLINE and EMBASE; three studies assessing AML were included.

No ongoing trials or relevant abstracts were found.

Details from the included trials are summarized in the tables below.

1. Does any of the newly identified evidence contradict the current recommendations? (i.e., the current recommendations may cause harm or lead to unnecessary or improper treatment if followed)	No
2. Does the newly identified evidence support the existing recommendations?	Yes
3. Do the current recommendations cover all relevant subjects addressed by the evidence? (i.e., no new recommendations are necessary)	Yes. There is a need for more trial data to establish the role of a MRD risk-adapted approach to guiding treatment options for adult patients with ALL and AML. It is important to note that there are recent publications of studies using MRD for risk-adapted management. However, these studies do not meet the current inclusion criteria of this guideline as published. Nonetheless, these studies do lay the foundation for future studies which are expected to meet this guideline's inclusion criteria and be able to support or refute a risk-adapted approach.
Review Outcome as recommended by the Clinical Expert	ENDORSE
If outcome is UPDATE, are you aware of trials now underway (not yet published) that could	

affect the recommendations?	
DSG/Expert Panel Commentary	

Table 1. Outcomes of predictive studies for AML

Author, year, reference	Procedures and Population	N	Methods	Intervention/ Comparison	Brief results
Craddock et al (2020) [1] FIGARO Phase II	Patients aged 22 to 75 with AML or high-risk MDS, who were undergoing their first allo-SCT from a matched sibling or unrelated donor and had been deemed ineligible for a MAC regimen.	244	BMs for MFC detection were obtained pretransplant (within four weeks of transplant and day +42 post-transplant.	Patients were randomized to a fludarabine-based RIC regimen or FLASMA-Bu.	• No significant difference in two-year OS between the control and FLAMSA-Bu arms (p=0.81). • No interaction was found between MRD status and conditioning intensity in the preplanned subgroup analysis for OS (p=0.56) or relapse risk (p=0.92).
Lambert et al (2021) [2] ALFA-0702	Patients aged 18 to 59 with newly diagnosed de novo AML.	314	BM and PB samples were used to quantify WT1 transcripts using qRT-PCR and were evaluated on days 28 and 45 after chemotherapy.	All patients received induction chemotherapy with an optional second induction course in those who did not achieve CR after the first course. Patients in CR with intermediate- or unfavourable-risk AML without a donor for allo-SCT were randomly assigned to HDAC or CLARA.	• The interaction between MRD status and the effect of allo-SCT was not statistically significant for RFS (p=0.48), OS (p=0.12) or risk of relapse (p=0.32). This shows that the effect of allo-SCT was similar in postinduction patients with WT1 MRD_{high} and MRD_{low}.
Ahn et al (2021) [3]	Patients diagnosed with NK-AML who had achieved CR1 between 2002 and 2014 at two institutions.	124	NGS was performed in 278 samples collected from BM (n=255) or PB (n=23) at initial diagnosis and CR1.	Patients who achieved morphologic CR received consolidation with or without allo-HCT depending on the availability of a matched related or unrelated donor.	• There was no difference in OS according to MRD status in the subgroup receiving allo-HCT; however, a statistically significant interaction was found between allo-HCT and MRD status for OS (p=0.036).

Abbreviations: allo-HCT: allogeneic hematopoietic stem cell transplantation; AML: acute myeloid leukemia; BM: bone marrow; CLARA: clofarabine and cytarabine; CR: complete remission; CR1: first remission; FLASMA-Bu: fludarabine/amsacrine/cytarabine-busulphan; HDAC: high-dose cytarabine; MAC: myeloblative conditioning; MDS: myelodysplasia; MFC: multiparameter flow cytometric; MRD: measurable residual disease; NGS: next-generation sequencing; NK-AML: normal karyotype acute myeloid leukemia; OS: overall survival; PB: peripheral blood; qRT-PCR: quantitative real time-quantitative PCR; RFS: relapse-free survival; RIC: reduced-intensity conditioning.

REFERENCES

1. Craddock C, Jackson A, Loke J, Siddique S, Hodgkinson A, Mason J, et al. Augmented Reduced-Intensity Regimen Does Not Improve Postallogeneic Transplant Outcomes in Acute Myeloid Leukemia. J Clin Oncol. 2021;39(7):768-78.
2. Lambert J, Thomas X, Marceau-Renaut A, Micol JB, Renneville A, Clappier E, et al. Early detection of WT1 measurable residual disease identifies high-risk patients, independent of transplantation in AML. Blood Advances. 2021;5(23):5258-68.
3. Ahn JS, Kim T, Jung SH, Ahn SY, Jung SY, Song GY, et al. Allogeneic transplant can abrogate the risk of relapse in the patients of first remission acute myeloid leukemia with detectable measurable residual disease by next-generation sequencing. Bone Marrow Transplant. 2021;56(5):1159-70.

Appendix 1. Affiliations and Conflict of Interest Declarations

Name	Affiliation	Conflict of Interest Declaration
Authors		
Harriet Feilotter Molecular Geneticist	Kingston General Hospital Kingston, ON	No conflict of interest
Duvaraga Sivajohanathan Health Research Methodologist	Program in Evidence-Based Care, OH (CCO), McMaster University Hamilton, ON	No conflict of interest
Expert Panel		
Mitchell Sabloff Hematologist	Ottawa Hospital Ottawa, ON	Has served on the Advisory Board for Bristol Myers Squibb, Astellas, AbbVie, Pfizer, Taiho, Jazz, Celgene, Novartis and Roche; has received a grant from Taiho and Bristol Myers Squibb as a principal or co-investigator for Biobank
Philip Berardi Hematopathologist	Ottawa Hospital Ottawa, ON	Has served on the advisory board for discussing methods for FLT3 mutation testing in AML supported by Astellas Pharma Canada (February 2020); has been a panelist discussing the clinical utility of NGS in myeloid malignancies supported by Novartis Pharma Canada (April 2021); has served on the biomarker advisory board for diffuse large B cell lymphoma supported by Janssen Pharma (May 2019)
Chris Bredeson Hematologist	Ottawa Hospital Ottawa, ON	No conflict of interest
Rena Buckstein Hematologist	Odette Cancer Centre Toronto, ON	Has received $500 or more in consulting for Bristol Myers Squibb and Taiho; has received grants as a principal or co-investigator from Bristol Myers Squibb, Taiho and Takeda for MDS registry.
Rouslan Kotchetkov Hematologist	Royal Victoria Regional Heath Centre Barrie, ON	No conflict of interest
Tom Kouroukis Hematologist	Juravinski Cancer Centre Hamilton, ON	No conflict of interest
Brian Leber Hematologist	Juravinski Cancer Centre Hamilton, ON	Has served on the Medical Advisory Board Membership for and received honoraria from Pfizer, AbbVie, Novartis, BMS/Celgene, AMGEN, Jazz, Astellas, Astex, Paladin, Alexion, Roche, Otsuka, Janssen and Treadwell
Andre Schuh Hematologist	Princess Margaret Cancer Centre Toronto, ON	No conflict of interest

Anne Tierens Hematopathologist	University Health Network Toronto, ON	No conflict of interest

DEFINITIONS OF REVIEW OUTCOMES

1. **ARCHIVE** – ARCHIVE means that a Clinical Expert and/or Expert Panel has reviewed new evidence pertaining to the guideline topic and determined that the guideline is out of date or has become less relevant. The document will no longer be tracked or updated but may still be useful for academic or other informational purposes. The document is moved to a separate section of our website and each page is watermarked with the words "ARCHIVE."

2. **ENDORSE** – ENDORSE means that a Clinical Expert and/or Expert Panel has reviewed new evidence pertaining to the guideline topic and determined that the guideline is still useful as guidance for clinical decision making. A document may be endorsed because the Expert Panel feels the current recommendations and evidence are sufficient, or it may be endorsed after a literature search uncovers no evidence that would alter the recommendations in any important way.

3. **UPDATE** – UPDATE means the Clinical Expert and/or Expert Panel recognizes that the new evidence pertaining to the guideline topic makes changes to the existing recommendations in the guideline necessary but these changes are more involved and significant than can be accomplished through the Document Assessment and Review process. The Expert Panel advises that an update of the document be initiated. Until that time, the document will still be available as its existing recommendations are still of some use in clinical decision making, unless the recommendations are considered harmful.