A systematic literature review and meta-analysis of minimal residual disease as a prognostic indicator in adult B-cell acute lymphoblastic leukemia

Renato Bassan,1 Monika Brüggemann,2 Hoi-Shen Radcliffe,2 Elizabeth Hartfield,4 Georg Kreuzbauer5 and Sally Wetten2

1Complex Operative Unit of Haematology, dell’Angelo Hospital and Santissimi Giovanni and Paolo Hospital, Mestre and Venice, Italy; 2Department of Medicine II, Schleswig-Holstein University Hospital, Kiel, Germany; 3Amgen Ltd, Uxbridge, UK; 4Oxford PharmaGenesis, Oxford, UK and 5Amgen (Europe) GmbH, Zug, Switzerland

ABSTRACT

Minimal (or ‘measurable’) residual disease in acute lymphoblastic leukemia appears to be a prognostic indicator, with potential value in informing individualized treatment decisions. Complete understanding of the strength of the association between minimal residual disease and long-term outcomes is, however, lacking. A systematic literature review and meta-analysis were performed to elucidate the clinical significance of minimal residual disease with respect to relapse-free survival and overall survival in precursor B-cell acute lymphoblastic leukemia. A total of 23 articles and abstracts, most published between 2012 and 2016, were identified for inclusion in the primary meta-analysis. Typically, patients were in their first complete remission at the time of minimal residual disease assessment; in two studies, all patients were in their second, or later, complete remission. The primary analysis revealed improved relapse-free survival across all studies for patients who achieved minimal residual disease negativity (random effects hazard ratio, 2.34; 95% confidence interval, 1.91–2.86). Improved overall survival for patients who achieved minimal residual disease negativity was also observed (hazard ratio, 2.19; 95% confidence interval, 1.63–2.94). There was no observed difference in the impact of minimal residual disease status in subgroups based on disease stage, minimal residual disease sensitivity threshold level, Philadelphia chromosome status, histological phenotype, risk group, minimal residual disease testing location, minimal residual disease timing after induction, or minimal residual disease detection method. Despite heterogeneity in study design and patient populations between the contributing studies, these data provide a compelling argument for minimal residual disease as a clinical tool for assessing prognosis and guiding treatment decisions in precursor B-cell acute lymphoblastic leukemia.

Introduction

Acute lymphoblastic leukemia (ALL) is a heterogeneous disease that derives from lymphoid cell populations.1 Precursor B-cell ALL (B-ALL) is the most common immunological subtype,2 and the most common genetic abnormality in B-ALL is the Philadelphia chromosome (Ph-positive ALL), found in approximately one-quarter of adult patients.3 Genetic or molecular profiling can be used to characterize disease prognosis, predict response to therapy, and inform treatment decisions.4,5 The mainstay of treatment for patients with newly diagnosed B-ALL has historically involved induction chemotherapy, followed by consolidation and maintenance chemotherapies; allogeneic hematopoietic stem cell transplantation (HSCT) is recommended following consolidation in selected high-risk groups.6 However,
for patients with Ph-positive disease in particular, outcomes with standard combination chemotherapy are poor. The tyrosine kinase inhibitors imatinib, dasatinib, nilotinib, and ponatinib have shown efficacy in patients with Ph-positive disease, and there are some preliminary suggestions that they may also be efficacious in selected subsets of Ph-negative ALL. Moreover, several targeted therapies have been developed for patients with B-ALL, such as the antibody-based therapies rituximab, inotuzumab ozogamicin and blinatumomab, these latter demonstrating greater effectiveness than salvage chemotherapy in the very high-risk setting of relapsed/refractory ALL, including both Ph-positive and Ph-negative patients.

The introduction of targeted therapies, alongside advances in diagnostic procedures, have improved outcomes for patients with B-ALL. However, despite a substantial proportion (74% to 91%) of patients achieving complete remission (CR), one-third or more will eventually relapse because of the presence of submicroscopic levels of leukemic cells in the bone marrow. The presence of these remaining cancer cells is known as minimal residual disease (MRD; alternatively termed ‘measurable residual disease’).

MRD is increasingly being used in clinical practice as an independent prognostic marker for the duration of CR and long-term outcomes in patients with ALL, and for informing treatment decisions. The European Society for Medical Oncology (ESMO) and National Comprehensive Cancer Network clinical practice guidelines for adult patients with ALL recommend the quantification of MRD whenever possible, but it is not yet clear whether MRD status alone is sufficient to predict prognosis, or whether it should be combined with other parameters, such as patient-related (e.g., age) or disease-related (e.g., cytogenetics) risk factors. In drug development, MRD response has been considered as an early marker of efficacy in clinical studies, with potential use as a surrogate endpoint in registration studies for accelerated drug approval. A full understanding of the prognostic significance of MRD is needed to guide its use across this broad range of settings.

A large number of studies have shown that achievement of MRD-negative status correlates positively with CR duration, reduced risk of relapse, and H SCT success. The use of MRD testing is individualized within a given study protocol, and hence there is wide variation in the test method used, the timing and sensitivity of MRD assessment, the characteristics of the patients, and the treatments used before and after MRD was assessed. We conducted a systematic review to capture the evidence supporting the clinical significance of MRD on clinical outcomes and used this evidence to inform a meta-analysis with the aim of quantifying the impact of MRD status on relapse-free survival (RFS) and overall survival (OS). A recent meta-analysis suggested that event-free survival and OS were almost always better in patients with ALL who were MRD negative than in those who were MRD positive. This meta-analysis included studies in children and adults, and in B-cell and T-cell phenotypes. Our meta-analysis focused on adult patients with B-ALL and includes 23 studies, 18 of which were not in the previous meta-analysis, which enabled us to explore the impact of MRD status within clinically important subgroups in this population.

Methods

Systematic literature review

Evidence base

We systematically reviewed published studies according to a prespecified protocol, using a process compliant with the 2009 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. PubMed and Embase databases were searched (Online Supplementary Table S1) for studies in humans published in English between 1 January, 1995 and 1 March, 2016. Additional searches were conducted for congress proceedings [American Society of Clinical Oncology (ASCO), American Society of Hematology (ASH), European Hematology Association (EHA), and ESMO] published between 2012 and 2016.

Screening and data extraction

Titles or abstracts identified from the initial searches were evaluated against the prespecified inclusion criteria (Online Supplementary Methods). The full texts were obtained for all studies deemed eligible, and for studies whose eligibility was unclear during the title/abstract screening. The full texts were independently screened by two reviewers to confirm which studies should be included. Any discrepancies were resolved by a third reviewer. Data extraction is described in the Online Supplementary Methods.

Meta-analysis

Evidence base

Studies identified from the systematic literature review were excluded from the meta-analysis if RFS was not reported separately for patients who were MRD-negative and MRD-positive in status, or if insufficient data were reported to calculate a hazard ratio according to the methods outlined in Tierney et al. For studies that included patients with B-ALL and T-cell ALL (T-ALL), hazard ratios were calculated only for the patients with B-ALL, if it was possible to do so.

Outcomes

The primary outcome was RFS – alternatively termed disease-free, event-free or leukemia-free survival in some publications. In studies that provided RFS definitions, RFS was mostly measured from CR until relapse or death. Other studies that provided a definition measured RFS from H SCT or from the start of treatment. For the purposes of the meta-analysis, we used the measure of RFS as provided in each study publication. The primary analysis set was based on studies that reported RFS according to MRD status. The secondary outcome was OS, determined from those studies in the primary analysis set that also reported OS according to MRD status.

Statistical analysis

The analysis of hazard ratios used unadjusted measures of treatment effect, where available. Meta-analysis was performed using the random effects model. Heterogeneity between studies was assessed statistically using the F and Cochran Q tests. Meta-regression was performed to investigate the relationships between covariates (Ph status, median follow-up time, MRD cut-off sensitivity threshold, post-MRD treatment, disease stage, sex and age) and study-level hazard ratios.

Results

Search results

The systematic literature search identified 1,899 records (1,252 full papers and 647 congress abstracts) for
title and abstract screening. An additional four full papers were included that were not identified during the systematic literature searches, giving a total of 1,903 records (Figure 1). Following screening of the title and abstract of each record, 278 records were included in the full text screen. Finally, 33 full papers and 21 congress abstracts were covered in the systematic review; these included 8,820 patients with B-ALL, 5,979 of whom had available MRD data. The study and patients’ characteristics, together with clinical outcomes, are summarized in Online Supplementary Tables S2–S7. In total, 32 records (23 full texts and 9 congress abstracts) were eligible for the meta-analysis; of these, 25 records were included in the primary analysis set; the remaining papers were excluded because either the study populations overlapped with another, more recent, publication, and there were no unique data for inclusion in the meta-analysis, or inconsistencies in data reporting meant that it was not possible to calculate hazard ratios. A list of studies identified in the systematic review but not included in the meta-analysis, with reasons, is provided in Online Supplementary Table S8.

Figure 1. PRISMA flow diagram. *Includes four full text articles that were not identified during the systematic literature searches. †Protocol amendment. Includes disease-free survival, leukemia-free survival and duration of complete response or progression-free survival. ALL: acute lymphoblastic leukemia; B cell ALL/B-ALL: B-cell acute lymphoblastic leukemia; HR: hazard ratio; MRD: minimal residual disease; MRD+: minimal residual disease-positive status; MRD−: minimal residual disease-negative status; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; RFS: relapse-free survival; T cell ALL/T-ALL: T-cell acute lymphoblastic leukemia.
Characteristics of studies and patients in the meta-analysis

**Disease stage and treatment**

Key characteristics of the studies included in the meta-analysis (Table 1) were considered in order to provide context when analyzing clinical outcomes according to MRD status.\(^{20,31–36,43–56}\) Most studies assessing MRD status in adults with B-ALL were published from 2011 onwards, with the exception of a prospective study in Germany\(^{15}\) and a Polish study.\(^{17}\) Studies were mainly conducted in

| First author     | Year | Total number of patients (with MRD data) | Ph status | MRD status | MRD method | MRD level | MRD test location | ALL phenotype | MRD timing | Disease stage at MRD | Pre-MRD tx | Post-MRD tx | Risk group |
|------------------|------|------------------------------------------|-----------|------------|------------|-----------|-------------------|---------------|-------------|---------------------|------------|-------------|------------|
| Lussana\(^{46}\) | 2016 | 106 (73)                                 | Pos       | NA         | PCR        | 10\(^{-5}\) | NA                | B-cell        | CR1         | Targeted            | HSCT       | NA          |            |
| Gökbuget\(^{41}\) | 2015 | 116 (112)                                | Neg       | CR1        | PCR        | 10\(^{-4}\) | Central           | B-cell        | CR1         | Targeted            | Mix        | NA          |            |
| Chiaretti\(^{42}\) | 2015 | 63 (60)                                  | Pos       | CR1        | PCR        | NA        | NA                | B-cell        | CR1         | Targeted            | Mix        | NA          |            |
| Jabbour\(^{43}\)  | 2017 | 78 (78)                                  | NA        | CR2        | Flow (6 color) | 10\(^{-6}\) | Local           | B-cell        | CR2 or later | Targeted            | Mix        | NA          |            |
| Nishikawa\(^{44}\) | 2016 | 432 (432)                                | Pos       | CR1        | PCR        | 10\(^{-5}\) | Local           | B-cell        | CR1         | Targeted            | Mix        | NA          |            |
| Ravandi\(^{45}\)  | 2016 | 340 (260)                                | Mix       | CR1        | Flow (6 color) | 10\(^{-4}\) | Local           | B-cell        | CR1         | Targeted            | Mix        | NA          |            |
| Bassan\(^{47}\)   | 2014 | 159 (106)                                | Neg       | CR1        | PCR        | 10\(^{-4}\) | NA                | Mix (79% B-cell) | CR1         | Chemo             | Mix        | Mix         |            |
| Beldjord\(^{48}\)  | 2014 | 860 (423)                                | Neg       | CR1        | PCR        | 10\(^{-4}\) | Central          | B-cell        | CR1         | Chemo             | Mix        | Mix         |            |
| Gökbuget\(^{49}\)  | 2012 | 1648 (580)                               | Neg       | CR1        | PCR        | 10\(^{-4}\) | Central          | Mix (60% B-cell) | CR1         | Chemo             | Mix        | Mix         |            |
| Holowiecki\(^{50}\) | 2008 | 131 (116)                                | Neg       | CR1        | Flow (3 color) | 10\(^{-6}\) | Central          | Mix (75% B-cell) | CR1         | Chemo             | Mix        | Mix         |            |
| Patel\(^{51}\)    | 2010 | 161 (161)                                | Neg       | CR1        | PCR        | 10\(^{-4}\) | NA                | B-cell        | CR1         | Chemo             | Mix        | Mix         |            |
| Bassan\(^{52}\)   | 2014 | 304 (141, [98 included in the analysis]) | Neg       | CR1        | PCR        | 10\(^{-4}\) | NA                | Mix (70% B-cell) | CR1 or later | Targeted            | Mix        | Mix         |            |
| Gökbuget\(^{53}\)  | 2014 | 189 (73)                                 | Neg       | CR1        | PCR        | 10\(^{-4}\) | Central          | B-cell        | CR2 or later | Targeted            | Mix        | Mix         |            |
| Banchano\(^{54}\)  | 2014 | 197 (185)                                | Pos       | CR1        | PCR        | NA       | Local            | B-cell        | Pre-HSCT    | Targeted            | HSCT       | NA          |            |
| Giebel\(^{55}\)   | 2010 | 123 (123)                                | Neg       | CR1        | Mix        | 10\(^{-4}\) | Local           | B-cell        | Pre-HSCT    | Targeted            | HSCT       | NA          | High       |
| Tucunduva\(^{56}\) | 2014 | 98 (98)                                  | Pos       | CR1        | Mix        | 10\(^{-4}\) | Local           | B-cell        | CHS         | Targeted            | Mix        | Mix         | NA         |
| Wassmann\(^{57}\)  | 2005 | 27 (27)                                  | Pos       | CR1        | PCR        | 10\(^{-4}\) | NA                | B-cell        | Post-HSCT   | Targeted            | NA         | NA          |            |
| Weng\(^{58}\)     | 2013 | 125 (106)                                | Mix       | CR1        | Flow (6 color) | 10\(^{-4}\) | Local           | B-cell        | CR1         | Chemo             | Mix        | NA          |            |
| Yanada\(^{59}\)   | 2008 | 100 (85)                                 | Pos       | CR1        | PCR        | 10\(^{-4}\) | Central          | B-cell        | CR1         | Targeted            | Mix        | NA          |            |
| Wetzel\(^{60}\)   | 2014 | 34 (34)                                  | Pos       | CR1        | PCR        | NA       | Central          | B-cell        | Post-HSCT   | Targeted            | NA         | NA          |            |
| Yoon\(^{61}\)     | 2016 | 173 (169)                                | Pos       | CR1        | PCR        | 10\(^{-4}\) | Central          | B-cell        | CR1         | Targeted            | NA         | NA          |            |
| Lim\(^{62}\)      | 2016 | 82 (78)                                  | Pos       | CR1        | PCR        | 10\(^{-4}\) | Central          | B-cell        | CR1         | Targeted            | Mix        | NA          |            |
| Short\(^{63}\)    | 2016 | 202 (122)                                | Pos       | CR1        | PCR        | 10\(^{-4}\) | NA                | B-cell        | CR1         | Targeted            | Targeted   | NA          |            |

Included in subgroup analysis

| Ravandi\(^{44,d}\) | 2016 | 340 (260)                                | Mix       | CR1        | Flow (6 color) | 10\(^{-4}\) | Local           | B-cell        | CR1         | Targeted            | Mix        | NA          |            |
| Gökbuget\(^{41}\) | 2012 | 1648 (580)                               | Neg       | CR1        | PCR        | 10\(^{-4}\) | Central          | Mix (60% B-cell) | CR1         | Chemo             | Mix        | Standard/high |            |
| Holowiecki\(^{47}\) | 2008 | 131 (116)                                | Neg       | CR1        | Flow (3 color) | 10\(^{-5}\) | Central          | Mix (75% B-cell) | CR1         | Chemo             | HSCT/ Standard/high | Mix        |            |
| Patel\(^{57}\)    | 2010 | 161 (161)                                | Neg       | CR1        | PCR        | 10\(^{-4}\) | NA                | B-cell        | CR1         | Chemo             | HSCT       | NA          |            |

<CR1 65%, CR2 34%, CR3 2%;<CR1 81%, CR2 19%;<CR1 78%, CR2 22%; Study included in primary analysis with additional subgroups available for inclusion in subgroup analyses. Minimal residual disease was assessed at multiple time points, both x3 months and x3 months after induction. ALL, acute lymphoblastic leukemia; Central: central laboratory; CR1: first complete remission; CR2: second complete remission; CR3: third complete remission; Chemo: chemotherapy; Flow: flow cytometry; HSCT: hematopoietic stem-cell transplantation; Local: local laboratory; Mix: mixed; MRD: minimum residual disease; NA: not available; Neg: negative; PCR: polymerase chain reaction; Ph: Philadelphia chromosome; Pos: positive; R/R: relapsed and/or refractory; Targeted: targeted agent (e.g., tyrosine kinase inhibitor, blinatumomab, inotuzumab); tx: treatment.>
Europe or East Asia, and almost all studies prospectively assessed MRD; only two studies included patients exclusively in second CR (CR2) or later, and the rest were in patients in first CR (CR1) at the time of MRD assessment (of these, three studies included a minority of patients in CR2 or later in combination with patients in CR1, and these were categorized as CR1 for the purposes of the meta-analysis). Approximately half were in Ph-positive patients (n=11), and these patients had typically received chemotherapy plus a tyrosine kinase inhibitor before their MRD assessment (Online Supplementary Table S2). Most patients with Ph-negative ALL received chemotherapy, with or without targeted agents. As mentioned, there were only two studies in which all patients were in CR2 or later, and these patients had received treatment with blinatumomab or inotuzumab ozogamicin, with or without chemotherapy, before MRD was assessed. Two studies, in Ph-positive patients, examined the impact of MRD status following transplantation. Treatment received after MRD assessment varied across, and within, the studies; in most studies (n=14), post-MRD survival outcomes were reported for the mix of post-MRD treatments (targeted therapy, chemotherapy and H SCT) (Table 1). Six studies reported outcomes separately for patients who received an H SCT after their MRD assessment; a further two studies reported outcomes separately for post-MRD treatment with a targeted agent.

**Minimal residual disease assessment methodology and timing**

MRD was most commonly assessed at a central laboratory (n=10); other studies used a local reference laboratory (n=7) or did not specify whether centralized assessment was performed (n=6). A range of methodologies were used to assess MRD, although most (n=17) were polymerase chain reaction (PCR) based (Table 1). All but one of the studies of patients with Ph-negative ALL used PCR-based methodologies; three specified that they looked for immunoglobulin or T-cell receptor gene rearrangements (Online Supplementary Tables S2 and S3). The remaining study used multiparameter flow cytometry (Online Supplementary Table S2). For patients who had Ph-positive ALL, all studies (n=11) used BCR-ABL as a marker of MRD; most (n=9) used realtime quantitative PCR, one study used molecular testing and one study used fluorescence in-situ hybridization. The most commonly used sensitivity limit was 1 in 10⁶ cells for all methodologies, and sensitivities of 1 in 10⁸ cells were only reported for PCR-based detection in studies of patients with Ph-positive status (Table 1). The three studies that included patients with mixed Ph status or did not report Ph status all used multiparameter flow cytometry (22%) (Online Supplementary Tables S2 and S3). The number of colors used ranged from three to eight, and no trends in the number of colors used were observed over time in the studies.

MRD measurements were generally taken within 3 months of induction treatment (n=14) (Online Supplementary Tables S2 and S3).

**Clinical outcomes according to minimal residual disease status**

**Relapse-free survival**

Figure 2 shows the meta-analysis results for the 23 studies included in the primary analysis. The overall results show improved RFS for patients who achieved MRD negativity (random effects hazard ratio (HR)=2.34; 95% confidence interval (CI): 1.91–2.86), and the effect was seen consistently across all studies.

The F value was 59%, indicating moderate-to-high heterogeneity between the studies in the primary analysis. This effect was anticipated given the known variations between studies in the design, patients’ characteristics and treatments received. For this reason, a predefined set of subgroups was examined to explore whether MRD negativity had a consistent effect in important subgroups of patients.

All subgroups showed an improved RFS in patients who achieved MRD negativity; no significant differential subgroup effects were observed (Figure 3). Although it should be noted that for some subgroups there were very few studies available, which contributed to overlapping confidence intervals, some trends were seen. The prognostic value of MRD appeared stronger in patients who received chemotherapy than in those who received targeted therapy before MRD assessment (HR=2.98; 95% CI: 2.12–4.20 and HR=0.90; 95% CI: 1.53–2.36, respectively), but the confidence intervals overlap. The subgroup that received HSCT before MRD assessment had a hazard ratio of 5.19 (95% CI: 1.95–13.8), but it should be noted that this is based on one study, and the confidence intervals are very wide. The prognostic value of MRD also seemed to be stronger in patients who received chemotherapy after their MRD assessment (HR=6.52; 95% CI: 2.43–17.5) than in those who received HSCT after MRD assessment (HR=1.73; 95% CI: 1.27–2.37), but again it should be noted that the data for chemotherapy after MRD assessment are from only one study.

The subgroup analysis was repeated separately according to Ph status and timing of MRD assessment (Online Supplementary Figures S1 and S2). There was no difference in RFS improvement for patients who achieved MRD negativity between studies in the Ph-positive (HR=2.04; 95% CI: 1.53–2.73) and Ph-negative (HR=2.46; 95% CI: 2.02–2.98) groups. Hazard ratios in all subgroups favored MRD negativity, regardless of Ph status, with no apparent difference between Ph-positive and Ph-negative cases in any subgroup (Online Supplementary Figure S1). Similarly, there was no difference in RFS improvement for patients who achieved MRD negativity at early timepoints (within 3 months from induction: HR=2.60; 95% CI: 2.05–3.31) and later timepoints (more than 3 months from induction: HR=2.23; 95% CI: 1.67–2.97). Hazard ratios were more favorable in most subgroups at the early MRD timepoint compared with the later MRD timepoint group, but the subgroup confidence intervals overlapped between the early and later MRD timepoints (Online Supplementary Figure S2).

**Overall survival**

The meta-analysis results for the 14 studies included in the secondary (OS) analysis are summarized in Figure 4. The overall results show improved OS for patients who achieved MRD negativity (HR=2.19; 95% CI: 1.63–2.94). Individually, the results were consistently in favor of MRD negativity across all the studies with the exception of that by Bachanova et al., who reported a hazard ratio of 0.94, but with a confidence interval that crossed the null value (95% CI: 0.65–1.35). Four other studies also had confidence intervals that crossed the null value. The F value was greater for the OS analysis (67%) than for the RFS analysis (59%), again confirming the extent of
the heterogeneity in the studies in terms of design, populations and treatment protocols. Reassuringly, a predefined analysis of key subgroups of interest showed a consistent improvement in OS in patients who achieved MRD negativity (Figure 5). As for RFS, no significant differential subgroup effects were seen. The prognostic value of MRD appeared to be greater in patients who received HSCT before MRD assessment (HR=8.02; 95% CI: 2.32–27.7) than those assessed after chemotherapy (HR=3.01; 95% CI: 2.08–4.37) or targeted therapy only (HR=1.65; 95% CI: 1.24–2.20). The prognostic value of MRD was less notable in patients who received HSCT treatment after MRD assessment (HR=1.24; 95% CI: 0.86–1.78) than those who received alternative interventions (HR=2.50; 95% CI: 1.88–3.33). As for RFS, the number of studies in some subgroups was small, and many of the confidence intervals overlapped.

**Meta-regression**

Meta-regression was explored to further investigate the heterogeneity between studies. The analysis was limited by the small number of studies that provided information for certain covariates (i.e. those used to define the subgroups), and for this reason we focused on RFS for which more studies were available. Heterogeneity remained even after adjustment for almost all covariates, as reflected by a significant within-group Q-statistic (Qw) (Online Supplementary Table S9). The exception was risk group; however, interpretation of the relevance of this covariate is limited by its inclusion in only three studies.

**Discussion**

This systematic review and meta-analysis found that in adults with B-ALL, achieving MRD negativity was consistently associated with better survival outcomes than those of patients with MRD-positive status. In addition, the benefit of achieving MRD negativity was evident in all the

---

**Figure 2. Forest plot of relapse-free survival hazard ratios for all studies included in the primary analysis.** CI: confidence interval; HR: hazard ratio; MRD: minimal residual disease; MRD neg: minimal residual disease-negative status; MRD pos: minimal residual disease-positive status.
subgroups that we examined, highlighting its relevance for assessing prognosis and measuring treatment efficacy. There were some preliminary indications that MRD status may have a greater effect on outcomes in certain patient subgroups than in others. These data should be interpreted with caution, since no significant differential subgroup effects were seen, and most of the confidence intervals overlapped.

Our data build upon evidence from a recent meta-analysis of patients with B-ALL or T-ALL by Berry et al. that showed hazard ratios for event-free survival and OS almost uniformly favor MRD negativity, and that this

---

**Figure 3. Forest plot of relapse-free survival hazard ratios by subgroup (random effects model).** B cell ALL/B-ALL: B-cell acute lymphoblastic leukemia; CI: confidence interval; CR1: first complete remission; CR2: second complete remission; Chemo: chemotherapy; Flow: flow cytometry; HR: hazard ratio; HSCT: hematopoietic stem-cell transplantation; MRD: minimal residual disease; MRD neg: minimal residual disease-negative status; MRD pos: minimal residual disease-positive status; N: number of studies; PCR: polymerase chain reaction; Ph: Philadelphia chromosome; SCT: stem-cell transplantation; targeted: targeted agent (e.g., tyrosine kinase inhibitor, blinatumomab, inotuzumab); tx: treatment.
effect is retained within each of the subgroups. The hazard ratio effect size appears slightly larger than in our meta-analysis, which could reflect differences in the studies included or the different analysis approach that was used (with a Bayesian rather than a random effects model). There was only partial overlap in studies between the two meta-analyses, with five studies included in both, meaning that our study adds 18 studies and 2,245 patients to the combined evidence base. Only one of the adult studies in the analysis by Berry et al. was identified as concerning B-ALL, but our efforts to extract data from the B-ALL populations wherever possible allowed us to conduct subgroup analyses specific for this phenotype with 19 studies of RFS and 12 studies of OS. Similarly, we calculated outcomes according to Ph status for a larger number of studies (20 vs. 8 in Berry et al.), enabling a more robust subgroup analysis and allowing further subgroup analyses to be conducted within each cytogenetic category. The lack of overlap between the two meta-analyses was likely due to differences in eligibility criteria: our study allowed for a larger range of methods to calculate hazard ratios (as described in Tierney et al.), only excluding studies with small patient numbers if MRD was examined retrospectively, and included congress abstracts. Our study focused on a narrower population than that of Berry et al., but a high degree of heterogeneity remained; this highlights the limitations of conducting meta-analyses according to MRD status and should encourage future research using pooled analysis of patient-level data.

In general, MRD evaluations are conducted in two different scenarios. The first is in patients who achieve morphological remission after receiving induction therapy, at the time of remission evaluation or shortly after, to evaluate the quality of the response. The second scenario is to use MRD monitoring as a predictor for pending relapse. We examined subgroups according to the timing of MRD assessment, either before or after 3 months from starting induction, or before or after HSCT. MRD assessments up to 3 months from induction were generally taken during or at the end of induction treatment – i.e. before consolidation or intensification therapy – and this subgroup therefore falls into the first scenario. In our analysis, there was little difference in the relative effect of MRD assessed at 5 months or later, and four studies that evaluated outcomes from multiple MRD timepoints showed a beneficial effect of MRD negativity at both earlier and later timepoints. Furthermore, subgroup analysis of RFS according to the timing of MRD assessment showed that there was no difference in RFS improvement for patients who achieved MRD negativity at early timepoints compared to those who achieved it at later timepoints. Measurement of MRD after HSCT may also fall into the second scenario, and the apparently large impact of MRD on RFS and OS in the post-HSCT subgroup may reflect the strong predictive ability of MRD to detect relapse in this setting.

Nevertheless, the timing of MRD measurement is an important consideration if the technique is to be used to support treatment decisions, and evidence exists to support early MRD testing. For example, in a study conducted in the USA that enrolled patients with Ph-negative B-ALL or T-ALL, MRD levels as early as 28 days following the initiation of induction therapy were used to predict out-

Figure 4. Forest plot of overall survival hazard ratios for all studies included in the primary analysis. CI: confidence interval; HR: hazard ratio; MRD: minimal residual disease; MRD neg: minimal residual disease-negative status; MRD pos: minimal residual disease-positive status.
comes. In another study assessing patients with Ph-negative B-ALL or T-ALL, an early MRD response (day 11) was associated with the best prognosis. In addition, in an analysis of patients with Ph-negative ALL from France, Belgium and Switzerland, lack of MRD response 6 weeks after induction initiation could identify patients who would benefit most from HSCT. This highlights the importance of early testing if MRD status is to be used to influence treatment decisions.

Information from MRD testing could help to target highly effective immunotherapies, as well as therapies that are less well tolerated, such as HSCT, to those at high risk of relapse. In a study of Ph-negative patients from the GRAALL-2003 or -2005 trials, transplant was found to

| Disease stage: CR1 CR2 or later | HR [95% CI] | Subgroup N |
|----------------------------------|-------------|------------|
|                                  | 2.33 [1.67, 3.26] | 12 |
|                                  | 1.52 [0.93, 2.48] | 2 |

| Timing of MRD rel HSCT: after HSCT before HSCT | 6.10 [2.47, 15.1] | 2 |
| MRD level: $10^{-4}$ $10^{-5}$ | 2.48 [1.93, 3.18] | 9 |
| Ph status: mixed Ph negative Ph positive | 1.52 [1.14, 2.01] | 2 |
| Phenotype: B-cell mixed | 3.40 [1.20, 9.59] | 2 |
| Post MRD tx: mixed SCT targeted therapy | 2.55 [1.93, 3.37] | 5 |
| Pre MRD tx: HSCT only chemo only targeted therapy | 1.84 [1.15, 2.94] | 6 |
| Risk group: high risk standard risk | 2.16 [1.54, 3.03] | 12 |
| MRD testing location: central local | 2.42 [1.64, 3.56] | 2 |
| Timing of MRD: $\leq$ 3 months from induction $> 3$ months from induction | 3.89 [1.21, 12.5] | 2 |
| MRD methodology: flow PCR Overall | 3.40 [1.20, 9.59] | 2 |
| 2.55 [1.93, 3.37] | 5 |
| 1.84 [1.15, 2.94] | 6 |
| 2.16 [1.54, 3.03] | 12 |
| 2.42 [1.64, 3.56] | 2 |
| 3.40 [1.20, 9.59] | 2 |
| 2.55 [1.93, 3.37] | 5 |
| 1.84 [1.15, 2.94] | 6 |
| 2.16 [1.54, 3.03] | 12 |
| 2.42 [1.64, 3.56] | 2 |
| 3.40 [1.20, 9.59] | 2 |
| 2.55 [1.93, 3.37] | 5 |
| 1.84 [1.15, 2.94] | 6 |
| 2.16 [1.54, 3.03] | 12 |
| 2.42 [1.64, 3.56] | 2 |
| 3.40 [1.20, 9.59] | 2 |
| 2.55 [1.93, 3.37] | 5 |
| 1.84 [1.15, 2.94] | 6 |
| 2.16 [1.54, 3.03] | 12 |
| 2.42 [1.64, 3.56] | 2 |
| 3.40 [1.20, 9.59] | 2 |
| 2.55 [1.93, 3.37] | 5 |
| 1.84 [1.15, 2.94] | 6 |
| 2.16 [1.54, 3.03] | 12 |
| 2.42 [1.64, 3.56] | 2 |
| 3.40 [1.20, 9.59] | 2 |
| 2.55 [1.93, 3.37] | 5 |
| 1.84 [1.15, 2.94] | 6 |
| 2.16 [1.54, 3.03] | 12 |
prolong RFS compared with chemotherapy among those who did not achieve an early MRD response, but was no better than chemotherapy in patients who did achieve an early MRD response. Likewise, the PETHHEMA ALL-AR08 trial used MRD to guide treatment decisions at the end of consolidation, and found that HSCT could be avoided in patients who reached MRD negativity without adversely affecting their prognosis. Similarly, a study using the Northern Italy Leukemia Group trial protocol 10/07 implemented risk-stratification to define a cohort of patients with an MRD response who could achieve good outcomes with conventional maintenance instead of HSCT after induction therapy. These studies provide evidence that MRD can play a role in sparing patients from the risks associated with transplantation, without negatively affecting survival outcomes. New systemic therapies may offer alternatives to HSCT; for example, a phase II study (published in full since the literature review was conducted) found that blinatumomab could induce a complete MRD response in over three-quarters of patients with MRD-positive ALL in first-line treatment; blinatumomab is now approved for use in patients with MRD-positive ALL in the USA. There is a need to establish the role of MRD testing in guiding treatment in each subgroup of patients with ALL and to determine an evidence-based treatment protocol to optimize patients’ outcomes.

Ph-positive and Ph-negative disease differ not just in terms of prognosis, but also in treatment regimens and method of MRD assessment. We analyzed RFS separately for Ph-positive and Ph-negative studies, and the prognostic ability of MRD negativity remained the same regardless of Ph status, with no discernible difference among subgroups. Subdividing the total population according to Ph status led to shrinkage or elimination of certain subgroups, but these results again point to the consistency of effect of MRD negativity across clinically relevant populations. It will be important in future research to further dissect the Ph-like subset of patients (who have a similar gene expression profile to Ph-positive ALL but without the targetable BCR-ABL translocation), given their poor prognosis and persistence of MRD.

The question of how sensitive MRD assays need to be remains to be resolved. In a study of patients with untreated Ph-negative B-ALL or T-ALL, MRD levels of 10−3 or greater were associated with a poor prognosis. In our study, no real differences in the association between MRD level and outcome were seen, regardless of whether the sensitivity threshold was 10−3, 10−4 or 10−5. Therefore, the threshold of 10−3 recommended by ESMO seems appropriate.

The studies included in this systematic review and meta-analysis represent a broad range of patient subgroups and treatment regimens. Consequently, the findings should be generalizable across the whole B-ALL population. Nevertheless, these analyses have some limitations. Given the wide range of treatment regimens, follow-up times and methodologies for assessing MRD, the data should be interpreted with caution. Heterogeneity was high (as measured using I2), but this was expected because it reflects the nature of studies included: they were not randomized and included different populations of patients, methods of analysis, follow-up times and study designs. In addition, treatments administered after MRD assessment, particularly HSCT, would be a major confounding factor when assessing the relationship between MRD and survival outcomes. We investigated sources of heterogeneity by using predetermined subgroup analyses of factors that might affect MRD and survival, and by using meta-regression. The subgroup analyses support the overall conclusion that MRD negativity is associated with better survival outcomes than MRD positivity. Heterogeneity remained high after accounting for key variables, indicating that there could be additional factors that affect the relative effect of MRD status on outcomes, which we were unable to account for in this analysis. Furthermore, for some subgroups very few studies were available, and data were sparse, meaning that the reliability of some of the subgroup results was suboptimal.

In some cases, time-to-event data were missing, and hazard ratios were calculated using published methodologies to account for this. However, the meta-regression for RFS indicated that the Tierney method used to calculate hazard ratios was unlikely to influence the magnitude of the effect of MRD negativity.

It would be of interest to study the impact of MRD on bone marrow relapse. Although RFS is a good proxy for relapse, deaths in remission (e.g., following HSCT) may affect the results. This meta-analysis focused on MRD as a risk factor for systemic relapse and did not examine MRD as a predictor of extramedullary relapse or death from other causes, such as transplant-related mortality; therefore, different rates of death across subgroups from causes other than bone marrow relapse could confound the results. The extent of confounding due to HSCT-related mortality is unclear, however, and only two of the studies included in the meta-analysis comprised solely patients who did not undergo HSCT. Overall, the strong relationship observed in our study, and that by Berry et al., highlights the high prognostic value of MRD, despite this confounding factor.

In four studies, it was not possible to separate the B-cell and T-cell populations; however, the majority of patients in these studies had B-ALL (66% to 79%), and there was no difference in the prognostic value of MRD between the B-ALL and mixed studies for either RFS or OS. It should also be noted that our analysis only included two studies in CR2 or later; however, results in this subgroup were consistent with the overall findings.

In conclusion, this systematic review and subsequent meta-analysis have generated support for the use of MRD as a prognostic marker in the management of patients with ALL. Overall, and in all subgroups analyzed, MRD negativity was associated with better survival outcomes than MRD positivity. This finding appeared relatively unaffected by variation in the timing or sensitivity threshold of the MRD assay applied, though other factors – such as the window of opportunity for treatment before relapse – may need to be considered when refining MRD assessment methodology to optimize its value in practice. Detection of MRD offers a promising clinical tool that may help clinicians to harness the potential of emerging targeted therapies for ALL.

Acknowledgments

The authors would like to thank James O’Kelly PhD of Amgen Ltd who provided editorial support, and Kim Allcott PhD of Oxford PharmaGenesis who assisted with the systematic literature review and provided editorial support. This research was funded by Amgen.
Inhibition therapy for adults with acute lymphoblastic leukemia: results of more than 1500 patients from the international ALL trial: MRC UKALL XG, DOG E2993. Blood. 2005;106(12):3760–3767.

Thomas X, Boiron JM, Huguet E, et al. Outcome of treatment in adults with acute lymphoblastic leukaemia: results of the LALA-94 trial. Clin Oncol. 2004;22(20):4075–4086.

Xie X, Xie H, Wood BL, et al. Relation of clinical response and minimal residual disease and their prognostic impact on outcome in acute myeloid leukaemia. J Clin Oncol. 2015;33(1):1258–1264.

Gökbuget N, Kneba M, Raft T, et al. Adult patients with acute lymphoblastic leukaemia and molecular failure display a poor prognosis and are candidates for stem cell transplantation and targeted therapies. Blood. 2012;120(9):1856–1867.

Scheuring U, Pfleiderer H, Wassmann B, et al. Serial minimal residual disease (MRD) analysis as a predictor of response duration in Philadelphia-positive acute lymphoblastic leukaemia (Ph+ALL) during imatinib treatment. Leukemia. 2003;17(9):1700–1706.

Bassan R, Intermesoli T, Scattolin A, et al. Minimal residual disease assessment and risk-based therapy in acute lymphoblastic leukaemia. Clin Lymphoma Myeloma Leuk. 2017;17(5):52–59.

Brown FA, Shah B, Advani A, et al. NCCN clinical practice guidelines in oncology with NCCN evidence blocks. Acute lymphoblastic leukemia. Version 1.2018 – December 13, 2018. Available from: https://www.nccn.org/professionals/physician_gls/pdf/acute_ligneous_blocks.pdf (Accessed 31 January 2019).

Short NJ, Jabbour E, Alfibar M, et al. Recommendations for the assessment and management of measurable residual disease in adults with acute lymphoblastic leukaemia: a consensus of North American experts. Am J Hematol. 2019;94(2):257–265.

European Medicines Agency. Guideline on the use of minimal residual disease as a clinical endpoint in multiple myeloma studies. 2018. Available from: https://www.ema.europa.eu/documents/scientific-guideline/draft-eu-guide-use-minimal-residual-disease-clinical-endpoint-multiple-myeloma-studies_en.pdf (Accessed 25 October 2018).

Food and Drug Administration. Hematologic malignancies: regulatory considerations for assessment of minimal residual disease in evaluation of drug and biological products for treatment – guidance for industry. 2018. Available from: https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM563355.pdf (Accessed 25 October 2018).

Brüggemann M, Kotreva M. Minimal residual disease in adult ALL: technical aspects and implications for clinical interpretation. Blood Adv. 2017;1(25):2455–2466.

Berry DA, Zhou S, Higley H, et al. Association of minimal residual disease with clinical outcome in pediatric and adult acute lymphoblastic leukemia: a meta-analysis. JAMA Oncol. 2015;1(7):e170580.

Librati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. Ann Intern Med. 2009;150:b270.

Tiemyer JF, Stewart LA, Gherzi D, Burdett S, Sydes MR. Practical methods for incorporating summary time-to-event data into meta-analysis. Trials. 2007;8:16.

Bachanova V, Marks DJ, Zhang MJ, et al. Ph+ ALL patients in first complete remission have similar survival after reduced intensity and myeloablative allogeneic transplantation: impact of tyrosine kinase inhibitor and minimal residual disease. Leukemia. 2014;28(8):658–665.

Giebel S, Stella-Holowiecka B, Krawczyk-Kulis M, et al. Status of minimal residual disease determines outcomes of autologous hematopoietic SCT in adult ALL. Bone Marrow Transplant. 2010;45(5):1095–1101.

Tucunduva L, Ruggeri A, Sanz G, et al. Impact of minimal residual disease on outcomes after umbilical cord blood transplantation for adults with Philadelphia-positive acute lymphoblastic leukaemia: an analysis on behalf of Eurocord, Cord Blood Committee and the Acute Leukaemia working party of the European Group for Blood and Marrow Transplantation. Br J Haematol. 2014;166(5):749–757.

Nakahashi S, Imai K, Mizuta S, et al. Impact of MRD and TKI on allogeneic hematopoietic cell transplantation for Ph−ALL: a study from the adult ALL WG of the JSHCT. Bone Marrow Transplant. 2013;50:40–50.

Wassmann B, Pfleiderer H, Stadler M, et al. Early molecular response to posttransplantation imatinib determines outcome in MRD+ Philadelphia-positive acute lymphoblastic leukemia (Ph+ ALL). Blood. 2005;106(2):458–463.

Gökbuget N, Kantarjian H, Brüggemann M, et al. An evaluation of molecular response in a phase 2 open-label, multicenter confirmatory study in patients (pts) with relapsed/refractory B-cell acute lymphoblastic leukaemia (r/r ALL) receiving treatment with the BiT E® antibody construct blinatumomab. Blood. 2014;124(21):3704–3709.

DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(5):177–188.

Whitehead A. Meta-Analysis of Controlled Clinical Trials. Chichester, UK: John Wiley & Sons, Ltd, 2005.

Hardy RJ, Thompson SG. Detecting and describing heterogeneity in meta-analysis. Stat Med. 1998;17(8):841–856.

Higgins JF, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557–560.

Gökbuget N, Dombret H, Bonafos M, et al. Long-term outcomes after blinatumomab treatment: follow-up of a phase 2 study in patients (pts) with minimal residual disease (MRD) positive B-cell precursor acute lymphoblastic leukaemia (ALL). Blood. 2015;126(25):680 (abstract).

Chimenti S, Vitale A, Elia L, et al. Multicenter total therapy Gimenia LA 1509 protocol for de novo adult Ph+ acute lymphoblastic leukaemia (ALL) patients. Updated results and refined genetic-based prognostic stratification. Blood. 2015;126(25):81.

Jabbour E, Short NJ, Jorgensen JL, et al. Differential impact of minimal residual disease negativity according to the salvage status in patients with relapsed/refractory B-cell acute lymphoblastic leukemia. Cancer. 2017;123(5):294–302.

Ravandi F, Jorgensen JL, O’Brien SM, et al. Minimal residual disease in required treatment: a parameter flow cytometry is highly prognostic in adult patients with acute lymphoblastic leukaemia. Br J Haematol.
Different molecular levels of post-induction minimal residual disease may predict hematopoietic stem cell transplantation outcome in adult Philadelphia-negative acute lymphoblastic leukemia. Blood Cancer J. 2014;4:e225.

Holowicek J, Krawczyk-Kulis M, Giebel S, et al. Status of minimal residual disease after induction predicts outcome in both standard and high-risk Ph-negative adult acute lymphoblastic leukemia. The Polish Adult Leukemia Group ALL 4-2002 MRD study. Br J Haematol. 2006;132(2):227–237.

Bassan R, Masciulli A, Spinelli O, et al. Enhanced survival in adult Philadelphia-negative acute lymphoblastic leukaemia: final results of the international trial UKALL XII/ECOG2995. Br J Haematol. 2010;148(1):80–99.

Weng XO, Shen Y, Sheng Y, et al. Prognostic significance of monitoring leukaemia-associated immunophenotypes by eight-color flow cytometry in adult B-acute lymphoblastic leukemia. Blood Cancer J. 2015;5:153.

Clinical significance of minimal residual disease quantification in adult patients with standard-risk acute lymphoblastic leukaemia. Blood. 2005;106(2):1116–1125.

Favorable outcomes for older adolescents and young adults (AYA) with acute lymphoblastic leukemia (ALL): early results of U.S. Intergroup Trial C10403. Blood. 2014;124(21):796.

Bruggemann M, Raff T, Flohr T, et al. Clinical significance of minimal residual disease quantification in adult patients with standard-risk acute lymphoblastic leukemia. Blood. 2005;106(2):1116–1125.

Enhanced survival in adult Philadelphia-negative acute lymphoblastic leukemia (Ph− ALL) with an updated pediatric derived minimal residual disease (MRD)/risk-specific treatment strategy: NILG Study 10. Haematologica. 2014;99(Suppl 1):S26.

Variable of complete molecular response on survival in patients with Philadelphia chromosome-positive acute lymphoblastic leukemia treated with imatinib. Haematologica. 2016;101(suppl 1):E856.

Achieving molecular remission before allogeneic stem cell transplantation in adult acute lymphoblastic leukemia: Impact on relapse and long-term outcome. Blood Marrow Transplant. 2016;22(11):1983–1987.

Differential molecular levels of post-induction minimal residual disease kinetics during the tyrosine kinase inhibitor-based chemotherapy courses in adults with Philadelphia-chromosome positive acute lymphoblastic leukemia in first remission. Blood. 2016;128(22):3979.

Ribera JM, Oriol A, Morgades M, et al. Treatment of high-risk Philadelphia chromosome-negative acute lymphoblastic leukemia in adolescents and adults according to early cytologic response and minimal residual disease after consolidation assessed by flow cytometry: final results of the PETHEMA ALL-AR-08 trial. J Clin Oncol. 2014;32(15):1595–1604.

Amgen Inc. BLINCYTO prescribing information. 2018. Available from: https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/125557s013lbl.pdf [Accessed 01 May 2018].

Haematologica | 2019; 104(10)