Influence of pre-transplant minimal residual disease on prognosis after Allo-SCT for patients with acute lymphoblastic leukemia: systematic review and meta-analysis

Zhenglei Shen†, Xuezhong Gu†, Wenwen Mao, Lifen Yin†, Ling Yang‡, Zhe Zhang‡, Kunmei Liu, Lilan Wang and Yunchao Huang

Abstract

Background: This meta-analysis was performed to explore the impact of minimal residual disease (MRD) prior to transplantation on the prognosis for patients with acute lymphoblastic leukemia (ALL).

Methods: A systematic search of PubMed, Embase, and the Cochrane Library was conducted for relevant studies from database inception to March 2016. A total of 21 studies were included.

Results: Patients with positive MRD prior to allogeneic stem cell transplantation (allo-SCT) had a significantly higher rate of relapse compared with those with negative MRD (HR = 3.26; P < 0.05). Pre-transplantation positive MRD was a significant negative predictor of relapse-free survival (RFS) (HR = 2.53; P < 0.05), event-free survival (EFS) (HR = 4.77; P < 0.05), and overall survival (OS) (HR = 1.98; P < 0.05). However, positive MRD prior to transplantation was not associated with a higher rate of nonrelapse mortality.

Conclusions: Positive MRD before allo-SCT was a predictor of poor prognosis after transplantation in ALL.

Trial registration: Not applicable.

Keywords: Acute lymphoblastic leukemia, Allogeneic stem cell transplantation, Minimal residual disease

Background

Acute lymphoblastic leukemia (ALL) is a hematologic malignancy of bone marrow featured by the overproduction of immature lymphoblasts [1]. It represents 75–80% of childhood acute leukemias and 20% of all leukemias in adults, with approximately 6000 cases diagnosed every year in the United States [1, 2]. Despite evolving treatment protocols, the relapse rate is approximately 15–20% in ALL, and the cure rate is much lower after relapse [3]. These relapses are due to the persistence of residual malignant cells, namely minimal residual disease (MRD), that cannot be detected by the morphological examination of the bone marrow [4]. Great efforts have been made to standardize MRD quantification using real-time polymerase chain reaction (PCR) of immunoglobulin and T-cell receptor (TCR) gene rearrangements, real-time PCR-based detection of fusion gene transcripts [e.g., breakpoint cluster region-Abelson (BCR-ABL)] or breakpoints, and flow cytometric immunophenotyping [5, 6]. MRD allows a more precise assessment of treatment efficacy and reduction of leukemic burden [7]. It has important prognostic and therapeutic implications for adults and children with ALL [8, 9]. The UKALL 2003 trial suggested that MRD risk stratification was helpful in adjusting the treatment intensity [10].

Allogeneic stem cell transplantation (allo-SCT) is the preferred treatment for adults with relapsed disease and children with high-risk relapses [11, 12].
The SCT mortality due to relapse is 30–40% in adults and children. The treatment-related mortality is also 30–40% in adults but lowers in children [13, 14]. A body of evidence indicated a direct correlation of the likelihood of relapse after transplant with the MRD status before transplantation [15, 16]. However, this significant association was not observed in some studies [17–20]. Also, patients’ age, detection methods, and adjustment of clinical covariates largely varied among different studies [16]. Additionally, the impact of MRD on overall survival (OS) and nonrelapse mortality (NRM) remained uncertain. Therefore, this systematic review and meta-analysis was conducted to explore the impact of MRD prior to transplantation on the prognosis for patients with ALL.

Methods
Search strategy and inclusion criteria
The meta-analysis was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis statement [21]. Studies in PubMed, Embase, and the Cochrane Library were searched from the database inception to March 2016, using the following text and/or medical subject heading terms: (1) “acute lymphoblastic leukemia” or “acute lymphoblastic leukaemia”; (2) “minimal residual disease”; (3) “relapse” or “relapse-free survival” or “leukemia-free survival” or “leukaemia-free survival” or “disease-free survival” or mortality; and (4) transplantation. The search was restricted to publications in the English language. The references of included studies were screened for potentially missing records. This systematic review with meta-analysis was not registered in a trial registry.

Studies considered for inclusion were as follows: (1) reported the detection of bone marrow MRD prior to allo-SCT in patients with ALL; (2) had no limitation in terms of the age of included patients; (3) cohort study, prospective or retrospective; (4) published in English; and (5) presented data on the main outcomes of relapse, relapse-free survival (RFS), event-free survival (EFS), and/or NRM. Disease-free survival (DFS) and leukemia-free survival (LFS) were interpreted as synonymous with RFS.

Study selection and quality assessment
Two independent reviewers (ZLS and XZG) screened the citations for inclusion based on titles and abstracts. Multiple studies involving the same cohort of patients (or duplicate patient populations) were identified and combined. Only the most recent or comprehensive study was selected to avoid double-counting. The Newcastle–Ottawa Scale (NOS) for nonrandomized studies was used to assess the quality of included studies. The items included patient selection (4 points), comparability of cohorts (maximum 2 points), and outcome assessment (maximum 3 points), with a total of 9 points [22]. The data extraction and quality assessment were conducted independently by two authors (LFY and LY). The information was examined and adjudicated independently by an additional author (WWM) referring to the original studies.

Statistical analysis
Time-to-event data were most appropriately analyzed using the hazard ratios (HRs). Thus, the HR and its 95% confidence interval CI) were used as summary effect estimates for these outcomes. Adjusted HRs were directly extracted from the results of multivariate analysis using the Cox regression model. When missing data regarding adjusted HR were encountered, it was indirectly estimated using the Kaplan–Meier (KM) curves guided by the method of Tierney et al [23]. The random-effects model was used for meta-analysis. The heterogeneity between the included studies was assessed using the Cochran Q test and the I² statistics. A P value less than 0.1 or I² values >50% was regarded as heterogeneity [24]. Subgroup analyses were performed based on the following clinical variables: study design (retrospective or prospective), region (USA, Europe, or Asia), population (adults or children), MRD assay modality [PCR or flow cytometry (FC)], and source of effect estimates (adjusted by multivariate analysis or unadjusted from KM curves). A sensitivity analysis was performed by excluding the studies one by one. Publication bias was visually explored using funnel plots and statistically assessed using Egger’s and Begg’s tests. All data were synthesized using the STATA 12.0 software and “metan” package (StataCorp LP, TX, USA). Two-tailed P ≤ 0.05 was considered significant for all statistical analyses.

Results
Literature search
A total of 418 studies were identified, including 221 from PubMed, 158 from Embase, and 39 from the Cochrane Library. The 158 duplicate studies were discarded. Also, 64 reviews or meta-analyses and 146 studies on irrelevant topics were removed. Fifty full-text studies were assessed for eligibility. Moreover, 4 studies of autologous stem cell transplantation and 19 studies that did not present the association between pre-transplant MRD and outcomes were discarded. Twenty-seven studies were included in qualitative synthesis. As 6 studies had insufficient data, 21 studies were finally pooled into the meta-analysis. The flow diagram of study selection is shown in Fig. 1.
Study characteristics and quality appraisal

The characteristics of 21 included studies are shown in Table 1. These articles were published between 1998 and 2016, including 9 retrospective studies and 12 prospective studies. The sample size ranged from 29 to 522. Seven studies enrolled mainly adult patients, 11 studies included mainly pediatric patients, and 3 studies comprised a mixture of adults and children. Six studies used FC and 15 studies employed PCR to detect MRD. The quality assessment of included studies is shown in Additional file 1: Table S1. The NOS score of included studies ranged from 5 to 9, and the important items included a representative of MRD (+) patients, comparability, and adequate follow-up duration. Seven studies did not enroll representative patients with ALL, including two studies of relapsed ALL, [25, 26] two studies of Philadelphia-positive ALL, [19, 27] one study of Philadelphia-negative ALL, [28] and two studies of high-risk ALL [18, 29]. For the item of comparability, the score was deducted when the study did not present sufficient adjusted effect estimates [14, 18, 28–33]. Only seven studies with 5-year outcomes in the follow-up were considered as the adequate follow-up [14, 18, 34–38].

Relapse

Twelve studies investigated the association between pre-transplantation MRD and cumulative incidence of relapse [18, 19, 27, 29, 31–34, 36, 37, 39, 40]. HRs from five studies were derived from the KM curves, [18, 29, 31–33], and HRs from seven studies were obtained from the multivariate analysis [19, 27, 34, 36, 37, 39, 40]. Patients with positive MRD prior to allo-SCT had a significantly higher rate of relapse compared with those with negative MRD (HR = 3.26; 95% CI 2.23–4.75, \( P < 0.05 \)) (Fig. 2). Moderate heterogeneity was revealed (\( I^2 = 46\% \); \( P < 0.05 \)). Subgroup analyses were conducted according to the following variables: population, design, region, detection method, and adjustment of HR. The pooled results remained statistically significant for all stratified analyses (except for the single Asian study), suggesting the robustness of the relationship. The heterogeneity was low or nonsignificant for the subgroups of adult patients, retrospective studies, United States of America (USA), FC, and adjusted HR, indicating that these factors might explain the observed heterogeneity (Table 2). A sensitivity analysis was further conducted. After each study was sequentially excluded from the pooled analysis, the conclusion was not affected by the exclusion of any specific study.

RFS

Data on RFS were directly obtained from seven studies, [19, 27, 34, 35, 38–40] or indirectly calculated from
| Author (year)          | Design     | No. of patients | Age (year) | Detection method | Endpoint               | Source of HR                  | Adjusted covariates                                                                 | Follow-up                      |
|------------------------|------------|-----------------|------------|------------------|------------------------|------------------------------|------------------------------------------------------------------------------------|-------------------------------|
| Knechtli et al. (1998) | Retrospective | 64              | 1.3–17     | PCR IG/TCR       | 2-year EFS             | Multivariate analysis        | Philadelphia chromosome positive, pre-BMT relapse                                   | Median: 35 months             |
| Bader et al. (2002)    | Retrospective | 41              | 1.5–17.8   | PCR IG/TCR       | 5-year EFS             | KM curve                    | NA                                                                                 | Median: 5.75 years            |
| Sramkova et al. (2007) | Prospective | 36              | 1.1–19     | PCR IG/TCR       | 4-year EFS             | KM curve                    | NA                                                                                 | Median: 26 months             |
| Spinelli et al. (2007) | Prospective | 37              | 18–63      | PCR IG/TCR, or PCR fusion transcript | 3-year CIR, 3-year OS | KM curve                    | NA                                                                                 | Median: 23 months             |
| Bader et al. (2009)    | Prospective | 91              | 3–226      | PCR IG/TCR       | 4-year CIR, 4-year EFS | Multivariate analysis        | Sex, age at relapse, remission status, time point of relapse, immunophenotype, site of relapse, stem-cell donor, T-cell depletion, time to transplantation, GVHD, MRD load before stem-cell transplantation | Median: 3.4 years             |
| Elorza et al. (2010)   | Prospective | 31              | 0.8–12     | FC               | 2-year EFS             | KM curve                    | NA                                                                                 | Median: 9 months              |
| Lankester et al. (2010)| Prospective | 48              | Mean: 8    | PCR IG/TCR       | 5-year CIR, 5-year EFS | KM curve                    | NA                                                                                 | Median: 61.5 months           |
| Doney et al. (2011)    | Retrospective | 161             | 18–62      | FC               | 5-year CIR, 5-year RFS | Multivariate analysis        | Disease status, year of transplantation, CMV serology mismatch                      | NA                            |
| Bachanova et al. (2012)| Prospective | 86              | 1–63       | FC               | 2-year CIR, 3-year DFS, 3-year OS | Multivariate analysis        | Disease status, ALL subtype, sex, time from diagnosis to transplantation, CMV status | Median: 39 years              |
| Ruggeri et al. (2012)  | Retrospective | 170             | <1–17      | PCR IG/TCR, or fusion transcripts | 4-year CIR, 4-year LFS, 4-year OS, 4-year NRM | Multivariate analysis        | Age, median year of transplant, cytogenetic risk group, TBI-based conditioning regimen, number of HLA disparities | Median: 48 months             |
| Mizuta et al. (2012)   | Prospective | 100             | 15–64      | PCR BCR-ABL      | 3-year CIR, 3-year OS, 3-year DFS, 3-year NRM | Multivariate analysis        | Age, donor status, chromosome abnormality, stem-cell source, performance status, BCR-ABL subtype, WBC, CD20 | Median: 31 months             |
| Sanchez-Garcia et al. (2013) | Retrospective | 102             | Mixed      | FC               | 5-year EFS, 5-year OS, 5-year RFS | Multivariate analysis        | Age, disease status, stem-cell source, time from onset to HSCT, GVHD                | Median: 60.8 months           |
| Balduzzi et al. (2014) | Prospective | 82              | <1–20      | PCR IG/TCR       | 5-year CIR, 5-year EFS | Multivariate analysis        | Disease status, donor type, HLA compatibility, GVHD                                | Median: 49 years              |
| Gandemer et al. (2014) | Prospective | 122             | ≤18        | PCR IG/TCR       | 5-year CIR, DFS, OS   | Multivariate analysis        | Sex, antithymocyte globulins, CNS location, risk stratification                    | Median: 34.8 months           |
| Tucunduva et al. (2014) | Retrospective | 98              | 18–66      | PCR or FC BCR-ABL | 3-year CIR, 3-year LFS | Multivariate analysis        | Age, sex, cytomegalovirus, disease status, transplantation method, conditioning, antithymocyte globulin, use of TKI, graft | Median: 36 months             |
| Author (year)     | Design     | No. of patients | Age (year) | Detection method | Endpoint                   | Source of HR   | Adjusted covariates                                                                 | Follow-up          |
|------------------|------------|-----------------|------------|------------------|----------------------------|----------------|-------------------------------------------------------------------------------------|-------------------|
| Logan et al. (2014) | Retrospective | 29              | 16–67      | PCR              | DFS                        | KM curve       | NA                                                                                  | Minimum: 3 years   |
| Zhou et al. (2014) | Retrospective | 149             | 18–70      | FC               | 2-year OS, 2-year PFS      | Multivariate analysis | Age, disease status, allotype, cell type                                       | NA                |
| Bar et al. (2014)  | Retrospective | 160             | 06–62      | FC               | 3-year CR, 3-year OS       | Univariate analysis | None                                                                             | Median: 40.6 months |
| Bader et al. (2015) | Prospective | 113             | 0–18       | PCR IG/TCR       | 3-year CR, 3-year EFS, 3-year NRM | Multivariate analysis | Disease status, immunophenotype, time of relapse, T-cell depletion                        | Range: 3.4–6.5 years |
| Sutton et al. (2015) | Prospective | 81              | < 18       | PCR IG/TCR       | 5-year CR, 5-year LFS, 5-year OS | Multivariate analysis | Sex, age, T-ALL, BCR-ABL1, hyperdiploidy > 50 or ETV6-RUNX1, BCP-other, IKZF1 mutation status, CR > 1, MSD, cord blood donor, mitoxantrone chemotherapy, TBI CY TT conditioning, in vitro T-cell depletion, ATG, GVHD | Median: 4.8 years  |
| Dhédin et al. (2016) | Prospective | 522             | 15–55      | PCR IG/TCR       | 3-year RFS, 3-year OS      | KM curve       | NA                                                                                  | Median: 3.5 years   |

**All Acute lymphoblastic leukemia, BMT Bone marrow transplantation, CIR Cumulative incidence of relapse, CNS Central nervous system, DFS Disease-free survival, EFS Event-free survival, FC Flow cytometry, GVHD Graft-versus-host disease, HR hazard ratio, IG Immunoglobulin genes, MRD Minimal residual disease, MSD Matched sibling donor, NRM non-relapse mortality, OS Overall survival, RFS Relapse-free survival, TBI total body irradiation, TCR T-cell receptor genes, TKI Tyrosine-kinase inhibitor**
three studies [28, 31, 33]. Pre-transplantation positive MRD was a significant negative predictor of RFS (HR = 2.53; 95% CI 1.67–3.84; \( P < 0.05 \)) (Fig. 3). Statistically significant heterogeneity was revealed (\( I^2 = 74.1\%; P < 0.05 \)).

Subgroup analyses were performed, and the results are shown in Table 3. Notably, the pooled results in the subgroups of Asian population and unadjusted HRs were nonsignificant. The heterogeneity remained moderate to high among all subgroups, indicating that none of the stratifying variables could explain the high heterogeneity.

In a sensitivity analysis, the conclusion was not affected by the exclusion of any specific study after each study was sequentially excluded from the pooled analysis.

### EFS

Eight studies explored the data on EFS, with four on adjusted HRs [25, 26, 35, 41] and four on unadjusted HRs [14, 30, 31, 36]. All studies were conducted in Europe. The pooled analysis revealed a significant correlation of positive MRD before allo-SCT with worse EFS.

| Study                          | Risk Ratio | RR     | 95%–CI          |
|-------------------------------|------------|--------|-----------------|
| Adding Spinelli,2007 (k=1)    |            | 3.91   | [1.17; 13.08]   |
| Adding Lankeser,2010 (k=2)    |            | 1.93   | [0.62; 6.01]    |
| Adding Eleorza,2010 (k=3)     |            | 2.83   | [0.85; 9.44]    |
| Adding Doney,2011 (k=4)       |            | 2.27   | [1.18; 4.35]    |
| Adding Ruggeri,2012 (k=5)     |            | 2.12   | [1.36; 3.31]    |
| Adding Bachanov,2012 (k=6)    |            | 2.26   | [1.46; 3.50]    |
| Adding Mizuta,2012 (k=7)      |            | 2.32   | [1.52; 3.55]    |
| Adding Balsuzzi,2013 (k=8)    |            | 3.10   | [1.82; 5.27]    |
| Adding Gandemer,2014 (k=9)    |            | 3.12   | [1.92; 5.08]    |
| Adding Tucunduva,2014 (k=10)  |            | 3.05   | [1.98; 4.68]    |
| Adding Bar,2014 (k=11)        |            | 3.07   | [2.10; 4.49]    |
| Adding Logan,2014 (k=12)      |            | 3.26   | [2.23; 4.75]    |

**Random effects model**

![Forest plot showing the association between pre-transplant MRD and relapse after allo-SCT](image)

### Table 2 Subgroup analyses for the outcome of relapse

| Subgroups          | No. of studies | HR (95% CI)   | \( I^2 \) (\( P \) value) |
|--------------------|----------------|---------------|---------------------------|
| Population         |                |               |                           |
| Children           | 5              | 3.30 (1.48–7.36) | 73.1% (0.01) |
| Adult              | 4              | 2.90 (1.85–4.57) | 8.5% (0.35)  |
| Design             |                |               |                           |
| Prospective        | 7              | 4.21 (1.93–9.16) | 62.2% (0.01) |
| Retrospective      | 5              | 2.73 (1.96–3.79) | 7.1% (0.37)  |
| Region             |                |               |                           |
| Europe             | 7              | 3.21 (1.80–5.71) | 61.3% (0.02) |
| USA                | 4              | 3.27 (2.00–5.36) | 24.6% (0.26) |
| Asia               | 1              | 7.34 (0.54–99.59) | –                       |
| Detection method   |                |               |                           |
| PCR                | 7              | 3.82 (1.94–7.53) | 62.0% (0.02) |
| FC                 | 4              | 3.10 (1.92–5.02) | 17.0% (0.31) |
| Adjustment         |                |               |                           |
| Adjusted           | 7              | 3.24 (2.10–5.00) | 32.8% (0.18) |
| Crude              | 5              | 3.43 (1.59–7.38) | 64.6% (0.02) |
| Competing risk framework |          |               |                           |
| Yes                | 10             | 2.96 (2.22–3.96) | 52.9% (0.02) |
| No                 | 2              | 2.40 (1.45–3.98) | 0.0 (0.39)   |
No significantly low heterogeneity was shown ($I^2 = 30.5\%; P > 0.1$).

In the subgroup analyses, the pooled results remained significant in all subgroups (Table 4). No heterogeneity was shown for the subgroups of pediatric patients, retrospective studies, and FC. In a sensitivity analysis, the conclusion did not change substantially by removing any single study after excluding the included studies one by one.

**OS**

Ten studies showed data on OS outcome [19, 28, 29, 32, 35, 37–40, 42]. The adjusted HRs were directly obtained from four studies [19, 35, 37, 42] and indirectly derived from KM curves in six studies [28, 29, 32, 38–40]. The pooled data demonstrated that patients with ALL having positive MRD prior to allo-SCT had a significantly unfavorable OS (HR = 1.98; 95% CI 1.40–2.80; $P < 0.05$) (Fig. 5). A significant heterogeneity was revealed ($I^2 = 67.2\%; P < 0.01$).

In stratified analyses, the pooled data for the subgroups of adult patients and Asian population showed no statistical significance (Table 5). The heterogeneity was low in the three subgroups of the population and the subgroups of American or Asian studies. In a subgroup analysis, the exclusion of any specific study did not alter the pooled conclusion significantly.

**NRM**

Three studies were eligible [19, 25, 40]. Positive MRD prior to transplantation was not associated with a higher rate of NRM (HR = 1.24; 95% CI 0.79–1.96; $P = 0.35$) (Fig. 6). A high heterogeneity was revealed ($I^2 = 69.4\%; P < 0.05$). Subgroup analyses or publication test was not conducted due to a limited number of studies. The three studies were excluded one by one, and the conclusion did not change significantly.

**Publication bias**

The publication bias for the outcomes of relapse, RFS, EFS, and OS was assessed. Notably, the funnel plots were asymmetrical for the outcome of relapse (Fig. 7a) and RFS (Fig. 7b). In contrast, the plots were symmetrical for the outcomes of EFS (Fig. 7c), and OS (Fig. 7d). When statistically assessed using the Egger’s test, the publication bias was statistically significant for relapse ($P = 0.02$) and RFS ($P = 0.02$), but not for EFS ($P = 0.20$) or OS ($P = 0.49$). Further, the results of Begg’s tests were examined, showing no publication bias for relapse ($P = 0.06$), RFS ($P = 0.21$), EFS ($P = 0.39$), and OS ($P = 0.86$).
Discussion

The prognostic value of pre-transplant MRD was demonstrated in several ways. First, the collected data revealed that patients with positive MRD before allo-SCT had a higher cumulative incidence of relapse in the follow-up. In accordance, the RFS was much shorter for the MRD (+) patients. The EFS was a composited outcome defined as the time from allo-SCT to the first occurrence of relapse or death. Positive MRD before allo-SCT was also an independent negative predictor of EFS. Furthermore, patients with positive pre-transplant MRD were proved to have a significant dismal OS. To the best of my knowledge, this meta-analysis was the first to appraise the role of MRD assessment in the pre-transplant setting in patients with ALL.

The prognostic power of pre-transplant MRD, as well as the sources of heterogeneity, was examined by subgroup analyses. A careful statistical process was conducted with caution that several clinical covariates might interact with the role of MRD and potentially cause bias. Thus, the adjusted estimates were preferred. The most relevant studies paid attention to the well-known confounding factors of disease status (CR1 or CR2), age, sex, and genetic mutations. Notably, the pooled HRs from the adjusted multivariate analysis or crude analysis from KM curves uniformly demonstrated a significant association between pre-transplant MRD and the outcomes of relapse, EFS, and OS. Only for RFS, the crude analysis by three studies failed to show the significant correlation. The included studies using PCR assay mostly amplified the rearranged immunoglobulin and TCR gene segment in the leukemic clone. Only few studies tested the fusogenic gene of BCR-ABL1. The FC-based assay, by examining the immunophenotypes, was advantageous in terms of rapid process and readily available results. Subgroup analyses demonstrated that the statistical significance for relapse, RFS, EFS, and OS was estimated using either PCR or FC method. The leukemogenic events were different for pediatric and adult patients with ALL [43]. A higher prevalence of unfavorable genetic subtypes, such as the BCR-ABL1 fusion protein, was observed among older patients [2]. The latest National Comprehensive Cancer Network guideline stated that the OS decreased substantially with increased age for patients with ALL [1]. Interestingly, MRD was not found to be a significant predictor of dismal OS for adult patients with ALL. However, MRD prior to transplantation was a constant negative predictor among the outcomes of relapse, RFS, and EFS, regardless of discrepancies in patients’ ages. Compared with prospective studies, retrospective studies relied on data recall or information from previous medical records. However, a significant difference was not observed by analyzing retrospective or prospective studies alone.

Of note, when assessing the outcome of NRM, the MRD status prior to transplantation did not have a significant role. Myeloablative conditioning for
MRD-positive patients versus reduced intensity conditioning for patients with undetected MRD might affect the NRM outcomes. A small number of studies might limit the statistical power. Additionally, some other factors might outweigh MRD in predicting NRM. Previous evidence suggested that younger adults had reduced post-transplant mortality. Myeloablation might be not feasible in patients older than 35 years because a higher toxicity was more commonly seen in these recipients [26, 44]. Thus, MRD might only be a subordinate factor for this outcome.

This meta-analysis had several strengths. It included 21 studies with a total of 2323 patients around the world. Comprehensive prognostic outcomes, including relapse, RFS, EFS, OS, and NRM, were evaluated. The prognostic value of MRD was appraised according to different assaying modalities, populations, and study designs. Both adjusted and crude data were presented. Largely, the association between MRD and outcomes remained stable among subgroups. It was confirmed that the detection of MRD was of considerable importance in identifying patients with poor outcome after allo-SCT. MRD was advocated to be a useful molecular biomarker for accurate triage of patients’ pre-transplantation and preemptive escalation of post-transplant interventions [15].

This study also had several shortcomings. Several studies collected data retrospectively. Some studies had small sample sizes, which might have reduced the statistical power. Patients’ disease status of NRM might have biased this relationship. The inclusion criteria were heterogeneous, and patients were treated using varied chemotherapy protocols. The timing and duration of follow-up were inconsistent. Also, the definitions of assay-specific thresholds and the lack of one universal detection method or testing target were heterogeneous among the included studies. No consensus was reached regarding the standardization of MRD measurement. Furthermore, it failed to give strong justification for providing a quantitative assessment of the influence of pre-transplant MRD. A multitude of confounding factors, such as the use of tyrosine kinase inhibitors, the pre-transplant remission type (CR1 or CR2), donor source, and graft-versus-leukemia, were not sufficiently adjusted in many studies when analyzing the impact of MRD. In fact, these factors were even inconsistent within an individual study [25, 40]. Even for studies that reported adjusted HRs, the degree of adjustment largely varied. The subgroup findings should be considered as exploratory, and thus would need to be tested in original studies. Finally, this study was conducted with summary

| Study                        | Hazard Ratio | HR (95% CI) |
|------------------------------|--------------|-------------|
| Adding Spinelli,2007 (k=1)   |              | 1.00 [0.43; 2.32] |
| Adding Bachanova,2012 (k=2)  | 1.45 [0.68; 3.11] |
| Adding Ruggeri,2012 (k=3)    | 1.71 [1.21; 2.43] |
| Adding Mizuta,2012 (k=4)     | 1.66 [1.19; 2.32] |
| Adding Sanchez–Garcia,2013 (k=5) | 1.99 [1.16; 3.43] |
| Adding Gandmre,2014 (k=6)    | 2.25 [1.36; 3.73] |
| Adding Bar,2014 (k=7)        | 2.33 [1.57; 3.45] |
| Adding Zhou,2014 (k=8)       | 2.19 [1.54; 3.12] |
| Adding Sutton,2015 (k=9)     | 2.21 [1.61; 3.03] |
| Adding Dh..edin,2016 (k=10)  | 1.98 [1.40; 2.80] |
| Random effects model         | 1.98 [1.40; 2.80] |

Table 5 Subgroup analyses for the outcome of overall survival

| Subgroups                              | No. of studies | HR (95% CI) | I² (P value) |
|----------------------------------------|----------------|-------------|--------------|
| Population                             |                |             |              |
| Adult                                  | 4              | 1.15 [0.86–1.53] | 0 (0.73) |
| Children                               | 3              | 2.40 [1.49–3.89] | 37.7% (0.20) |
| Mixed adult and children               | 3              | 3.06 [1.98–4.70] | 32.1% (0.23) |
| Design                                 |                |             |              |
| Prospective                             | 6              | 1.68 [1.02–2.77] | 60.1% (0.03) |
| Retrospective                          | 4              | 2.37 [1.55–3.63] | 64.8% (0.04) |
| Region                                 |                |             |              |
| Europe                                 | 5              | 2.05 [1.10–3.80] | 83.0% (< 0.01) |
| USA                                    | 3              | 2.10 [1.50–2.94] | 1.8% (0.36) |
| Asia                                   | 2              | 1.89 [0.99–3.62] | 0 (0.32) |
| Detection method                       |                |             |              |
| PCR                                    | 6              | 1.65 [1.07–2.55] | 61.5% (0.02) |
| FC                                     | 4              | 2.53 [1.58–4.04] | 58.1% (0.07) |
| Adjustment                             |                |             |              |
| Adjusted                               | 4              | 2.58 [1.27–5.26] | 71.1% (0.02) |
| Crude                                  | 6              | 1.70 [1.18–2.44] | 58.7% (0.03) |
statistics rather than with individual data, which might have ignored the impact of some covariates on the outcomes at the patient level. The availability of data from individual patients could resolve this problem and increase the power of meta-analysis.

Future studies should aim to decide how best to use the prognostic information of MRD. Several ways can be considered to improve the outcomes for MRD (+) patients at transplantation. Pre-transplantation treatment with non-cross-resistant agents might be helpful in decreasing the residual malignant clone [14, 40]. Preemptive immunotherapy or chemotherapy might be beneficial during the post-transplantation stage [45, 46]. Lankester et al. preliminarily revealed that alloimmune intervention after allo-SCT was feasible in reducing residual leukemic cells [17]. Further, a randomized trial should be performed on patients with ALL in complete remission who had positive MRD and received either allo-SCT or additional novel chemotherapy before transplantation.

Conclusions
In conclusion, this meta-analysis provided evidence that positive MRD prior to allo-SCT was associated with

![Forest plot showing the association between pre-transplant MRD and non-relapse mortality after allo-SCT](image)

![Funnel plots for the outcomes of relapse (a), relapse-free survival (b), event-free survival (c), and overall survival (d)](image)
higher relapse and poor survival in patients with ALL. Allo-SCT appeared to be insufficient for some patients with positive MRD at transplantation. The findings of this study suggested the rationale for future studies to prevent relapse and improve survival for this group of high-risk patients.

Additional file

Additional file 1: Table S1. Quality assessment of included studies using the Newcastle–Ottawa Scale (maximum score of 9). (DOCX 19 kb).

Abbreviations

ALL: Acute lymphoblastic leukemia; allo-SCT: allogeneic Stem cell transplantation; BCR-ABL: Breakpoint cluster Region-Abelson; DFS: Disease-free survival; EFS: Event-free survival; FC: Flow cytometry; HRs: Hazard ratios; KM: Kaplan–Meier; LFS: Leukemia-free survival; MRD: Minimal residual disease; NRM: Nonrelapse mortality; OS: Overall survival; PCR: Polymerase chain reaction; RFS: Relapse-free survival; TCR: T-cell receptor

Funding

This work was supported by the National Natural Science Foundation of China (Project No. 81360089) and the Applied Basic Research Joint Special and General Program of Yunnan Provincial Science and Technology Department (Project No. 2015FB072).

Availability of data and materials

The datasets used and/or analyzed in this study are available from the corresponding author on reasonable request.

Authors’ contributions

ZLS, XZG, and LFY contributed to conception and design. ZLS, XZG, WWM, LFY, LY, ZZ, KML, LLW, and YCH contributed to acquisition, analysis, and interpretation of data. ZLS, XZG, WWM, LFY, LY, ZZ, KML, LLW, and YCH were involved in drafting the manuscript or revising it critically for important intellectual content. All authors have given final approval of the version to be published.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

1Department of Hematology, The Third Affiliated Hospital of Kunming Medical University, Kunming, China. 2Department of Hematology, The First People Hospital in Yunnan Province, Kunming, China. 3Department of Geriatrics, The Second Hospital of Kunming, Kunming, China. 4Department of Hematology, The Second Affiliated Hospital of Kunming Medical University, Kunming 650031, China. 5Department of Chest Surgery, The Third Affiliated Hospital of Kunming Medical University, Kunming, China.

Received: 23 February 2018 Accepted: 15 July 2018

Published online: 23 July 2018

References

1. Alvamas JC, Brown PA, Aoum P, Ballen KK, Bellam N, Blum W, et al. Acute lymphoblastic leukemia. J Natl Compr Cancer Netw. 2012;10:858–914.
2. Hunger SP, Mullighan CG. Acute lymphoblastic leukemia in children. N Engl J Med. 2015;373:1541–52.
3. Oskarsson T, Soderhall S, Arvidson J, Foretist E, Montgomery S, Brottai M, et al. Relapsed childhood acute lymphoblastic leukemia in the Nordic countries: prognostic factors, treatment and outcome. Haematologica. 2016;101:68–76.
4. Campana D, Pui CH. Detection of minimal residual disease in acute leukemia: methodologic advances and clinical significance. Blood. 1995;85:1416–34.
5. Campana D. Applications of cytometry to study acute leukemia: in vitro determination of drug sensitivity and detection of minimal residual disease. Cytometry. 1994;18:68–74.
6. van der Velden VH, Caazaniga G, Schrauder A, Hancock J, Bader P, Panzer-Grumayer ER, et al. Analysis of minimal residual disease by Ig/TCR gene rearrangements: guidelines for interpretation of real-time quantitative PCR data. Leukemia. 2007;21:604–11.
7. Rubinzl JE, Inaba H, Dalh G, Ribeiro RC, Bowman WP, Taub J, et al. Minimal residual disease-directed therapy for childhood acute myeloid leukaemia: results of the AML02 multicentre trial. Lancet Oncol. 2010;11:543–52.
8. Bruggemann M, Raff T, Neuba M. Has MRD monitoring superseded other prognostic factors in adult ALL? Blood. 2012;120:4470–81.
9. Pui CH, Pei D, Constan-Smith E, Jhea S, Cheng C, Bowman WP, et al. Clinical utility of sequential minimal residual disease measurements in the context of risk-based therapy in childhood acute lymphoblastic leukemia: a prospective study. Lancet Oncol. 2015;16:465–74.
10. 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;14:199–209.
11. Parker C, Waters R, Leighton C, Hancock J, Sutton R, Moorman AV, et al. Effect of mitoxantrone on outcome of children with first relapse of acute lymphoblastic leukaemia (ALL R3): an open-label randomised trial. Lancet. 2010;376:2009–17.
12. Eckert C, Henze G, Seeger K, Hagedorn N, Mann G, Panzer-Grumayer R, et al. Use of allogeneic hematopoietic stem-cell transplantation based on minimal residual disease response improves outcomes for children with relapsed acute lymphoblastic leukemia in the intermediate-risk group. J Clin Oncol. 2013;31:2736–42.
13. Bassan R, Hoelzer D. Modern therapy of acute lymphoblastic leukemia. J Clin Oncol. 2011;29:532–43.
14. Bader P, Hancock J, Kreyenberg H, Goulden N, Niethammer D, Oakhill A, et al. Minimal residual disease (MRD) status prior to allogeneic stem cell transplantation is a powerful predictor for post-transplant outcome in children with ALL. Leukemia. 2002;16:1668–72.
15. Buckley SA, Appelbaum FR, Walter RB. Prognostic and therapeutic implications of minimal residual disease at the time of transplantation in acute leukemia. Bone Marrow Transplant. 2013;48:630–1.
16. Campana D, Leung W. Clinical significance of minimal residual disease in patients with acute leukaemia undergoing haematopoietic stem cell transplantation. Br J Haematol. 2013;162:147–61.
17. Imashuku S, Tsuru K, Matsuyma T, Asami K, Tsuchiya S, Ishi E, et al. Lack of clinical utility of minimal residual disease detection in allogeneic stem cell recipients with childhood acute lymphoblastic leukaemia: multi-institutional collaborative study in Japan. Bone Marrow Transplant. 2003;31:1127–35.
18. Lankester AC, Berinings MB, van Wering ER, Wijkhuis AJ, de Weger RA, Wijnen JT, et al. Preemptive allogeneic intervention in high-risk pediatric acute lymphoblastic leukemia patients guided by minimal residual disease level before stem cell transplantation. Leukemia. 2010;24:1462–9.
19. Mizuta S, Matsuo K, Maeda T, Yuji T, Hatta Y, Kimura Y, et al. Prognostic factors influencing clinical outcome of allogeneic hematopoietic stem cell transplantation following imatinib-based therapy in BCR-ABL-positive ALL. Blood Cancer J. 2012;2:e72.
20. Patel B, Rai L, Buck G, Richards SM, Montuza Y, Mitchell C, et al. Minimal residual disease is a significant predictor of treatment failure in non T-lineage adult acute lymphoblastic leukaemia: final results of the international trial UKALL XII/ECOG2993. Br J Haematol. 2010;148:80–9.
21. Mohler D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6:e1000097.
22. The Newcastle–Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available at: http://wwwohri.ca/programs/clinical_epidemiology/oxford.asp (Accessed: 15th March).
23. Tierney JF, Stewart LA, Ghersi D, Burdett S, Sydes MR. Practical methods for incorporating time-to-event data into meta-analysis. Trials. 2007;8:16.
