Prognostic Significance of WT1 Expression Level Threshold in Acute Myeloid Leukemia Patients Receiving Hematopoietic Stem Cell Transplantation: Meta-Analysis

Dandan Wang¹, Ling Qin¹,*, Boya Li¹, Tong Li¹

¹. Department Of Hematopathology, The First Affiliated Hospital, and College of Clinical Medicine of Henan University of Science and Technology, Luoyang, China, 471003

*Corresponding author:

Ling Qin

Department Of Hematopathology, The First Affiliated Hospital, and College of Clinical Medicine of Henan University of Science and Technology, Luoyang, China, 471003

Email: lingqindoc@163.com
Abstract

Objective: The WT1 gene is considered as a poor prognostic factor for acute myeloid leukemia (AML) after Allogeneic hematopoietic stem cell transplantation (Allo-HSCT). However, the effect of the expression threshold of WT1 on the prognosis is controversial, which is evaluated in this meta-analysis. Methods: Relevant studies about the expression threshold of WT1 on the prognosis of AML after Allo-HSCT were searched in online databases. Data were extracted from them and analyzed by Stata16.0 software. Results: Five studies involving 739 patients were screened out, including 433 cases experimental group and 306 cases control group. The experimental group and control group were compared for 1-year disease-free survival rate (DFS) [RR=1.19, 95%CI (1.03, 1.38), \( P =0.02 \)] and 4-year DFS [RR=1.18, 95%CI (0.98, 1.42), \( P =0.09 \)]. The experimental group was lower than the control group in 1-year DFS, and there was no statistical significance in 4-year DFS. 1-year overall survival rate (OS) [RR=1.06, 95%CI (0.92, 1.23), \( P =0.40 \)] and 4-year OS [RR=1.16, 95%CI (1.03, 1.32), \( P =0.02 \)], suggesting that the experimental group had a lower 4-year OS than the control group, and 1-year OS had no statistical significance. Conclusions: High WT1 expression is unfavorable to the prognosis of AML patients undergoing Allo-HSCT. A threshold of 250 copies/104ABL of WT1 may be the best value for predicting the poor prognosis in these patients.

Key words: WT1 threshold; Acute Myeloid Leukemia ; Hematopoietic stem cell transplantation; Prognosis; Meta-analysis
Introduction

Acute myeloid leukemia (AML), a group of heterogeneous diseases with different clinical outcomes, involves with cytogenetic and genetic lesions of hematopoietic stem cells or progenitor cells [1], increasing incidence with age [3] as the most common acute leukemia [2], for which the most effective treatment is Allogeneic Hematopoietic Stem Cell Transplantation (Allo-HSCT) [4]. However, the recurrence of patients with myeloid malignancies such as AML often causes the failure of hematopoietic stem cell transplantation, therefore, it is required to perform early monitoring of minimal residual leukemia (MRD) after Allo-HSCT [5].

First, Wilms tumor gene (WT1) involved with AML is expressed in all blast cells of acute leukemia patients, as WT1 nucleoprotein is detected in most of these blast cells [6]. More and more evidence shows that due to high expression of WT1 in various types of AML, WT1 can be used as a useful marker for monitoring MRD [7]. In addition, WT1 has a low expression in normal bone marrow, a high expression at the diagnosis of AML patients and after effective treatment, even a higher expression again before clinical recurrence [8], with higher sensitivity and specificity [5]. Although most studies have acknowledged the predictive value of WT1, few prognostic interventions have been based on WT1 expression levels [7]. Therefore, this study aims to comprehensively evaluate the influence of WT1 on the clinical
prognosis of patients after Allo-HSCT by Meta-analysis method for providing a reference for clinical medical work.

**Materials and methods**

Research type: Randomized controlled trials (RCT) or cohort study. Research subject: (1) AML patients receiving Allo-HSCT after chemotherapy induction and immune intervention; (2) No limitation of the age. Detection of WT1 expression: (1) WT1 expression was detected in bone marrow samples early after transplantation; (2) According to the European Leukemia Network (ELN), real-time quantitative PCR (RT-qPCR) was used to determine the mRNA levels of WT1 and ABL, and the latter (104 copies of ABL) was used as an internal reference to normalize the expression level of WT1 Interventions: According to the expression level of WT1, it was divided into high expression and low expression. Categorized by the threshold of WT1, AML patients after HSCT were further divided into experimental group (250 copies /104ABL of WT1 as the threshold) and control group (100 copies/104ABL of WT1 as the threshold). Outcomes: 1-year, 2-year, 4-year and 7-year recurrence rate (RR) and overall survival rate (OS) in high expression and low expression, and 1-year and 4-year disease-free survival rate (DFS) and overall survival rate (OS) in experimental group and the control group. Exclusion criteria: (1) Literatures with incomplete data or indicators that cannot be extracted; (2) Repeated publications; (3) Literatures not published in Chinese or English language; (4) Review, conference, abstract, case report, etc.

**Search strategy**
The study based on PubMed, The Cochrane Library, Embase, CNKI, Wanfang and China Biomedical Literature Database (CBM) aims to searched for the effect of WT1 on the prognosis of AML after HSCT, with the retrieval time from the establishment of the database to December 2020, as well as references included in the study are traced back to supplement relevant literature. The retrieval of subject words, carried out by the combination with free words, is adjusted according to the characteristics of each database. In English, "WT1" and "hematopoietic transplantation" are used as search terms, while "WT1" and "hematopoietic stem cell transplantation" in Chinese.

**Literature screening and data extraction**

Two researchers were independently responsible for literature screening, data extraction and cross-check. Any disagreement was solved by the third research through discussion or consultation. Briefly, literatures were initially screened by reviewing the title, followed by the abstract and full-text determined based on inclusion and exclusion criteria. Missing data were obtained by contacting the first author through e-mail or telephone if necessary. The following data were mainly extracted: (1) Basic information, including the first author, country, publication date, etc. (2) Basic characteristics of subjects, including sample size, age, 250 copy threshold, 100 copy threshold, etc. (3) Outcomes: 1-year, 2-year, 4-year and 7-year RR and OS in high expression and low expression, and 1-year and 4-year DFS and OS in experimental group and the control group. If the survival curve was obtainable in literature, the survival index and recurrence index corresponding to each point were yielded using Engauge Digitizer 10.8 software. (4) Key elements of bias risks.
Risk assessment on bias in included studies

Two researchers independently evaluate and cross-check the risk of bias in the included studies. In terms of differences, they discuss and resolve, and refer to a third party for adjudication if necessary. The RCT bias risk assessment is performed using the RCT bias risk assessment tool recommended in the Cochrane Manual 5.1.0., while the risk of bias in the included cohort studies is evaluated by the Newcastle-Ottawa Scale (NOS). The specific contents of the assessment included:

1. Selection, 1) Represent activeness of the exposed cohort; 2) Selection of non-exposed cohort; 3) Ascertainment of exposure; 4) Demonstration that outcome of interest was not present at the beginning of study.
2. Comparability: 1) Comparability of cohort on the basis of the design or analysis.
3. Outcome: 1) Assessment on outcome; 2) follow-up long enough for outcomes to occur; 3) Adequacy of follow-up of cohorts. The full mark of comparability between exposure and non-exposure is 2, and the remaining sub-items are all scored 1. The NOS scale scores on a scale of 9, 0-4 stands for low quality studies, 5-6 for medium quality research and 7-9 for high quality research.

Statistical analysis

Stata 16.0 software is used for Meta analysis. Risk ratio (RR) with 95% confidence interval (95%CI) is used as effect size for classification variables. The heterogeneity among the included results is analyzed by $\chi^2$ test (the test level was $\alpha=0.1$), and the magnitude of heterogeneity is quantitatively determined by $I^2$. $P<0.10$ or $I^2\geq50\%$ indicates significant heterogeneity. If there is no statistical heterogeneity among the results, the fixed-effect model will be used for Meta-analysis. If there is statistical
heterogeneity among the study results, the random effects model will be used for
Meta-analysis, and the sensitivity analysis conducted by excluding references one by
one. All publication offsets are first judged by funnel plot, and then Egger test is used
to evaluate the possibility of offsets. P<0.05 indicates that the difference is
statistically significant. TSA 0.9 software was used for sequential analysis to evaluate
the sample size of meta-analysis and verify the accuracy of the results.

Results

Literature screening and meta-analysis results

A total of 1460 literatures were initially screened out. After thoroughly reviewing
abstracts and full-text, a total of 739 patients in 5 cohort studies were involved in the
present study [9-13], including 217 patients with high expression and 522 patients
with low expression. Classified by the expression threshold of WT1, there were 433
cases in the experimental group with the threshold of 250 copies/10^4 ABL, and 306
cases in the control group with the threshold of 100 copies/10^4 ABL. The literature
screening process and their information were shown in Figure 1.

Basic characteristics of the included studies and quality
evaluation of included studies

The basic characteristics of included studies and WT1 expression are shown in Table
1 and Table 2. Those 5 included studies are all cohort studies. Literature quality score
shows that high quality of the included studies, as shown in Table 3.
### Table 1: Characteristics of the included studies

<table>
<thead>
<tr>
<th>The first author</th>
<th>Year</th>
<th>Country</th>
<th>WT1 expression threshold ((/10^4 ABL))</th>
<th>Sample size</th>
<th>Age</th>
<th>Outcome indicator</th>
<th>Quality Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Josep F. Nomde[9]</td>
<td>2018</td>
<td>Spain</td>
<td>100 copies</td>
<td>184</td>
<td>All</td>
<td>DFS, OS</td>
<td>8</td>
</tr>
<tr>
<td>Sarah Pozzi[10]</td>
<td>2013</td>
<td>Italy</td>
<td>100 copies</td>
<td>122</td>
<td>15-69</td>
<td>OS</td>
<td>9</td>
</tr>
<tr>
<td>Anna Candoni[11]</td>
<td>2017</td>
<td>Italy</td>
<td>250 copies</td>
<td>122</td>
<td>18-70</td>
<td>DFS, OS</td>
<td>8</td>
</tr>
<tr>
<td>Anna Candoni[12]</td>
<td>2017</td>
<td>Italy</td>
<td>250 copies</td>
<td>54</td>
<td>20-69</td>
<td>DFS, OS</td>
<td>8</td>
</tr>
</tbody>
</table>

Abbreviations: AML, Acute Myelogenous Leukemia; RR, Recurrence Rate; OS, Overall Survival; DFS, Disease Free Survival.
### Table 2: Characteristics of WT1 expression

<table>
<thead>
<tr>
<th>Included studies</th>
<th>Experimental group (250 copies /104 ABL)</th>
<th>Control group (100 copies /104 ABL)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>high expression</td>
<td>low expression</td>
</tr>
<tr>
<td>Anna Candoni2017[12]</td>
<td>21</td>
<td>33</td>
</tr>
<tr>
<td>Silvia Park2020 [13]</td>
<td>56</td>
<td>201</td>
</tr>
</tbody>
</table>

Abbreviations: RR, Recurrence Rate; OS, Overall Survival.
Table 3: Results of quality assessment using the Newcastle-Ottawa Scale for case-control studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Selection</th>
<th>Comparability</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Representativeness of the exposed cohort</td>
<td>Selection of non-exposed cohort</td>
<td>Ascertainment of exposure</td>
</tr>
<tr>
<td>Josep F. Nomde[9]</td>
<td>★ ★ ★ ★ ★</td>
<td>★ ★ ★</td>
<td>★</td>
</tr>
<tr>
<td>Sarah Pozzi[10]</td>
<td>★ ★ ★ ★ ★</td>
<td>★ ★ ★</td>
<td>★</td>
</tr>
<tr>
<td>Anna Candoni[12]</td>
<td>★ ★ ★ ★ ★</td>
<td>★ ★ ★</td>
<td>★</td>
</tr>
<tr>
<td>Silvia Park [13]</td>
<td>★ ★ ★ ★ ★</td>
<td>★ ★ ★</td>
<td>★</td>
</tr>
</tbody>
</table>
Abbreviations: ★, That is lscor
Results of Meta-analysis

WT1 expressed recurrence rate

RR was reported in four studies [10-13], and the random-effect model showed that compared with low expression, AML patients receiving Allo-HSCT with high expression had significantly higher 1-year RR [relative risk (RR) = 5.09, 95%CI (3.30, 7.83), \( P \leq 0.001, I^2 = 0.00\% \)], and 2-year RR [RR = 4.09, 95%CI (2.71, 6.18), \( P \leq 0.001, I^2 = 24.83\% \)], 4-year RR [RR = 3.27, 95%CI (2.42, 4.42), \( P \leq 0.001, I^2 = 0.00\% \)], and 7-year RR (RR = 3.33, 95%CI (2.49, 4.46), \( P \leq 0.001, I^2 = 0.00\% \)] (Figure 2A-D). It is suggested that AML patients after Allo-HSCT with WT1 high expression had a poor prognosis.

WT1 expression overall survival rate

OS was reported in four studies [10-13], and the random-effects model showed that the 1-year OS [RR=0.73, 95%CI (0.57, 0.94), \( P=0.02, I^2=70.96\% \)], 2-year OS [RR=0.59, 95%CI (0.45, 0.77), \( P \leq 0.001, I^2=53.53\% \)], 4-year OS [RR = 0.54, 95%CI (0.38, 0.75), \( P \leq 0.001, I^2 = 55.88\% \)] and 7-year OS [RR = 0.51, 95%CI (0.37, 0.70), \( P \leq 0.001, I^2 = 35.13\% \)] were significantly higher low expression than high expression (Figure 3A-D). It is suggested that AML patients after Allo-HSCT with low expression had a better prognosis.

The disease-free survival rate of experimental group and control group

Three studies reported DFS [9,11,12], and the random-effects model showed that the
1-year DFS was significantly lower in experimental group than that of control group [RR = 1.19, 95%CI (1.03, 1.38), \( P = 0.02 \)], while no significant difference was detected in the 4-year DFS [RR = 1.18, 95%CI (0.98, 1.42), \( P = 0.09 \)] (Figure 4A, B).

**The overall survival of experimental group and control group**

OS was reported in four studies [9-12], and compared in experimental group and control group. The random-effects model showed that the 4-year OS in experimental group was significantly lower than that of control group [RR = 1.16, 95%CI (1.03, 1.32), \( P = 0.02 \)], while no significant difference was detected in the 1-year OS between experimental group and control group [RR = 1.06, 95%CI (0.92, 1.23), \( P = 0.40 \)] (Figure 5A, B). It is suggested that the threshold of WT1 at 250 copies was an appropriate threshold for predicting the poor prognosis of AML after Allo-HSCT.

**Heterogeneity analysis**

The 1-year, 2-year, 4-year, and 7-year OS analysis of high expression and low expression was performed by introducing the random-effects model. However, \( I^2 \geq 50\% \) suggested heterogeneity. We subsequently explored the possible causes of heterogeneity through subgroup analyses based on the country and threshold of WT1. Subgroup analysis on 2-year OS showed that \( I^2 = 32.95\% \), suggesting that the 2-year OS may be the cause of heterogeneity. In addition, we conducted sensitivity analysis by excluding one literature each time. It is shown that \( I^2 = 14.92\% \) after excluding the literature reported by Candoni et al. [11] that reported 4-year OS, and \( I^2 = 12.44\% \) after excluding the literature reported by Candoni et al. [12] that reported 7-year OS,
indicating that data extracted from literature of Candoni et al. [11] and Candoni et al. [12] were sources of heterogeneity. The heterogeneity of 1-year OS may be related to publication bias. Although there was significant heterogeneity in 1-year, 2-year, 4-year and 7-year OS between high expression and low expression, the above factors affected the reliability of OS through heterogeneity analysis.

**Publication bias**

Stata 16.0 software was used for depicting funnel plots and the Egger test was performed to evaluate publication bias. RR of AML patients after HSCT in high expression and low expression were analyzed, and funnel plots were well symmetric. Egger test results showed that there were no significant differences in the 1-year ($P = 0.143$), 2-year ($P = 0.078$), 4-year ($P = 0.099$) and 7-year RR ($P = 0.238$) of AML patients between the high expression and low expression, suggesting no publication bias. In addition, there were significant differences in the 1-year ($P = 0.002$), 2-year ($P = 0.025$), 4-year ($P = 0.016$) and 7-year OR ($P = 0.036$), suggesting the publication bias in the included studies. It is considered that authors, research sponsors and manuscript editors are confounding factors for data reliability. The funnel plots of disease-free survival rate and overall survival rate in the 250 copies /10^4 ABL experimental group and the 100 copies /10^4 ABL control group showed good symmetry, as shown in Fig. 6A and B, and Fig. 7A and B. There were no significant differences in the 1-year DFS ($P = 0.812$), 4-year DFS ($P = 0.938$), 1-year OS ($P = 0.875$), and 4-year OS ($P = 0.968$), suggesting no publication bias.

**Discussion**
The WT1 gene is located on chromosome 11p13, which was originally identified as a tumor-suppressor gene associated with the Wilms tumor, a childhood kidney cancer. WT1 is overexpressed in various solid cancers and hematological malignant tumors [14]. Moreover, it is overexpressed in bone marrow (BM) or peripheral blood (PB) of more than 80% of AML patients, and is also served as a strong and independent marker of minimal residual disease (MRD) [15]. Multiple measurement methods of WT1 expression level in PB and BM have been proposed as follows [16,17,19-24]. (1) Ratio method: Real-time quantitative PCR (RT-qPCR), with ABL gene as the internal reference in the unit of 104 ABL copies; (2) Logarithmic method: Logarithmic order of WT1 gene decline; (3) Standard curve method: The mRNA level of WT1 detected in 1 μg K562 cells was defined as 1 to normalize its level in samples. In our included papers, ratio method was used to measure the expression level of WT1 in BM. Besides, real-time quantitative polymerase chain reaction (RT-qPCR) was used to measure mRNA levels of WT1 and ABL according to the European Leukemia Network (ELN), and the mRNA level of WT1 was normalized to 104 ABL copies as the unit. Here, ratio method was used to evaluate the effect of WT1 expression level on the prognosis of AML patients receiving Allo-HSCT. Recently, quantitative analysis of the WT1 gene combined with multi-parameter flow cytometry has significantly improved the sensitivity in predicting the risk of AML recurrence after Allo-HSCT, which has been applied in analyzing the prognosis of AML patients after Allo-HSCT [18].

In 2009, Cilloni et al. [19] defined overexpressed WT1 based on ELN as ≥ 50 copies
ABL copies in peripheral blood samples and ≥ 250 copies / 104 ABL copies in bone marrow samples detected by RT-qPCR. In other studies, ratio method is also used to evaluate the effect of WT1 expression on disease prognosis. Cho et al. [20] tested WT1 in BM samples in patients after Allo-HSCT, and suggested that 250 copies of WT1 is the optimal threshold and 3 months after Allo-HSCT was the optimal time point to detect the copy number of WT1. Ffairia et al. [21] believed that WT1 > 150 copies / 104 ABL copies in BM samples before transplantation is independently associated with RR and survival. In this study, the WT1 threshold was different from the ELN threshold. Sun et al. [22] believed that WT1 > 87 copies / 104 ABL copies in BM samples indicates a poor prognosis. Zhao et al. [23] believed that WT1 > 585 copies /104 ABL copies in BM samples is unfavorable to the prognosis after transplantation. In addition, Mashima et al. [24] found that two consecutive tests of WT1 transcription levels at more than 100 copies during remission may help identify patients at high risk of relapse. threshold of WT1 expression are inconsistent in many studies, which limits its routine clinical application. Two reasons may explain the inconsistent findings. Firstly, AML is a highly heterogeneous disease composed of subgroups with distinct biological characteristics [25]. Secondly, ELN threshold or threshold calculated based on the median of WT1 threshold is adopted in different studies [8]. Due to the lack of relevant guidelines and consensus on the optimal threshold of WT1 expression, this study analyzed the prognostic factors for AML patients after Allo-HSCT by meta-analysis. It is shown that 1-year DFS in the experimental group was significantly lower than that in the control group. The 4-year
OS of the experimental group was significantly lower than that of the control group, suggesting that the threshold of 250 copies/104 ABL of WT1 may be the optimal threshold for predicting the poor prognosis in AML patients after Allo-HSCT. In the heterogeneity analysis, data heterogeneity was found in Candoni et al. [11] and Candoni et al. [12], and 250 copies /104 ABL was defined as overexpression in these two papers. Therefore, considering the accuracy of the results of this study, TSAV0.9 software was used for sequential analysis to evaluate the sample size of meta-analysis. The results are shown in Figure 8A and B, which showed that this study did not reach the expected sample size (TSAn=5064) and there were false positive results. Therefore, 250 copies /104 ABL of WT1 expression as the optimal threshold should be further expanded to provide strong evidence for the impact on prognosis.

Previous studies have not highlighted the comparison between different thresholds of WT1 expression, and its optimal threshold. This study compared the effects of two different WT1 expression thresholds on the prognosis of AML patients after Allo-HSCT, so as to provide evidence for future clinical application. WT1 is overexpressed in a variety of hematopoietic tumors like AML, and previous studies have found that upregulated WT1 is associated with leukemia recurrence [16]. AML patients after Allo-HSCT expressing a high level of WT1 generally have a higher risk of recurrence, which negatively affects the survival. In this study, AML patients with high WT1 expression after Allo-HSCT were followed up at 1, 2, 4 and 7 years. Compared with patients with low WT1 expression of AML after Allo-HSCT, patients with high WT1 expression had a higher recurrence rate and a higher overall survival.
rate, which was similar to the results of previous published studies.

The optimal treatment for AML has not yet been established, and Allo-HSCT is the first-line treatment for low-risk and very low-risk AML patients [26], while post-transplant recurrence is a poor prognostic factor. Multiple studies have shown that monitoring WT1 expression level after Allo-HSCT can predict its recurrence [23]. In the present study, the cut-off value at 250 copies of WT1 was found to be the optimal threshold for predicting the poor prognosis of AML after Allo-HSCT. Monitoring the optimal overexpression threshold of WT1 contributes to the early identification of the recurrence and timely performance of donor lymphocyte infusion in AML patients after Allo-HSCT [25]. In addition, immunotherapy targeting WT1 may have the potential to prevent recurrence and improve survival [27]. Chapuis et al. [28] found that T cell receptor gene therapy targeting WT1 can prevent the recurrence of AML after transplantation. Recent studies have shown that WT1 peptide vaccine can inhibit tumor growth, and postoperative and/or chemotherapy vaccination can significantly reduce disease recurrence in patients with high tumor load [29]. WT1 gene expression may be related to the prognosis of AML with other factors. For example, the transplantation does not achieve disease remission or the presence of high-risk chromosomal abnormalities. Hidaka et al. [30] reported that WT1 expression is higher in INV (16) and T (15;17) AML patients than in T (8;21) AML patients, and FLT3-ITD or NPM1 mutation is independently associated with high WT1 expression in AML patients. A variety of WT1 gene mutations have been reported in solid tumors and AML. Wang et al. [31] believed that WT1 mutation is an
independent adverse prognostic factor in pediatric AML.

Several limitations in this study should be noted. (1) Due to the small sample size, this may reduce the statistical power of the results, and further expansion of the sample size is needed. (2) The lack of patients’ data and the limited number of studies may result in potential heterogeneity that influenced the accuracy of the results. (3) Chinese and English published literatures were searched in this study, but the classification of the threshold was only reported in foreign countries, which may have a certain publication bias.

In conclusion, high expression of WT1 significantly affects the prognosis of AML patients after Allo-HSCT, which is served as an adverse factor for predicting the recurrence. In detail, the threshold at 250 copies/10^4 ABL of WT1 is the optimal threshold for predicting the adverse prognosis of AML after Allo-HSCT. It is recommended to closely monitor the expression level of WT1 based on the calculated optimal threshold of copy numbers, thus predicting the recurrence, reducing adverse events and improving clinical outcomes in AML patients after Allo-HSCT. Due to the small sample size, it is necessary to further expand the sample size to validate our findings.

Acknowledgments

Thanks for the support of translation Center of Henan University of Science and Technology.

Conflict of interests

The authors declare that there is no conflict of interests regarding the publication of
this paper.

**Availability of data**

All data were extracted from the original articles included.

**References**


Figure 1: The Preferred Reporting Items for Meta-Analyses flow chart for study selection

- Literature retrieval (n=1460)

  - Exclusion of literature (n=534)
    - Rechecking (n=276)
    - Has nothing to do with literature (n=258)

  - Title, abstract screening (n=526)

    - Exclusion of literature (n=391)
      - No threshold study (n=320)
        - Review (n=59)
        - Other (n=261)
      - Non-randomized controlled trials (n=7)
      - The subjects were non-acute myeloid leukemia (n=23)
      - No hematopoietic stem cell transplantation intervention (n=17)
      - Other (n=24)

    - The full text retrieval (n=115)

      - Exclusion of literature (n=112)
        - Non-randomized controlled trials (n=27)
        - No hematopoietic stem cell transplantation intervention (n=46)
        - No prognostic indicators could be extracted (n=31)
        - Other (n=8)

      - Divided by 250 copies and 100 copies (n=13)
        - Exclusion of literature (n=8)

      - Divided into 585 copies (n=5)
        - Exclusion of literature (n=5)

      - Divided by 87 copies (n=5)

  - Eligibility

  - Included

  - Final study inclusion (n=5)
**Figure 2A: 1-year recurrence rate**

<table>
<thead>
<tr>
<th>Study ID</th>
<th>high</th>
<th>low</th>
<th>RR</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rel</td>
<td>norel</td>
<td>rel</td>
<td>norel</td>
</tr>
<tr>
<td>Anna Candoni (2017)</td>
<td>13</td>
<td>8</td>
<td>1</td>
<td>32</td>
</tr>
<tr>
<td>Silvia Park (2020)</td>
<td>17</td>
<td>39</td>
<td>13</td>
<td>188</td>
</tr>
</tbody>
</table>

**Overall**

Heterogeneity: $I^2 = 0.00$, $H^2 = 1.00$

Test of $\theta = 0$: $Q(3) = 2.34$, $p = 0.50$

Test of $\theta = 0$: $z = 7.39$, $p \leq 0.001$

Random-effects DerSimonian-Laird model
**Figure 2B: 2-year recurrence rate**

<table>
<thead>
<tr>
<th>Study ID</th>
<th>High</th>
<th>Low</th>
<th>RR with 95% CI</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sarah Pozz(2013)</td>
<td>35</td>
<td>32</td>
<td>4.79 [2.18, 10.54]</td>
<td>21.34</td>
</tr>
<tr>
<td>Anna Candon(2017)</td>
<td>15</td>
<td>6</td>
<td>4.74 [2.51, 8.95]</td>
<td>29.52</td>
</tr>
<tr>
<td>Silvia Park (2020)</td>
<td>18</td>
<td>38</td>
<td>7.88 [2.58, 23.90]</td>
<td>12.03</td>
</tr>
<tr>
<td>Overall</td>
<td>2.69 [1.58, 4.50]</td>
<td>37.11</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: $\tau^2 = 0.04$, $I^2 = 24.83\%$, $H^2 = 1.33$

Test of $\theta = 0$: $Q(3) = 3.99$, $p = 0.26$

Test of $\theta = 0$: $z = 6.71$, $p < 0.001$

Random-effects DerSimonian-Laird model
Figure 2C: 4-year recurrence rate

<table>
<thead>
<tr>
<th>Study</th>
<th>High</th>
<th>Low</th>
<th>RR</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anna Candoni(2017)</td>
<td>24</td>
<td>15</td>
<td>66</td>
<td>3.16 [1.87, 5.34]</td>
</tr>
<tr>
<td>Silvia Park (2020)</td>
<td>21</td>
<td>29</td>
<td>172</td>
<td>2.60 [1.61, 4.19]</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
<td>3.27 [2.42, 4.42]</td>
</tr>
</tbody>
</table>

Heterogeneity: $\chi^2 = 0.00$, $I^2 = 0.00\%$, $H^2 = 1.00$

Test of $\theta_0 = 0$: $Q(3) = 2.83$, $p = 0.42$

Test of $\theta = 0$: $z = 7.70$, $p \leq 0.001$

Random-effects DerSimonian-Laird model
**Figure 2D: 7-year recurrence rate**

<table>
<thead>
<tr>
<th>Study ID</th>
<th>High</th>
<th>Low</th>
<th>RR with 95% CI</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pozzi (2013)</td>
<td>35</td>
<td>32</td>
<td>3.19 [1.68, 6.05]</td>
<td>20.70</td>
</tr>
<tr>
<td>Candoni (2017)</td>
<td>28</td>
<td>13</td>
<td>3.69 [2.23, 6.09]</td>
<td>33.60</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td></td>
<td></td>
<td>2.69 [1.66, 4.36]</td>
<td>36.46</td>
</tr>
</tbody>
</table>

Heterogeneity: $\chi^2 = 0.00$, $I^2 = 0.00\%$, $H^2 = 1.00$

Test of $\theta = 0$: $Q(3) = 2.29$, $p = 0.51$

Test of $\theta = 0$: $z = 8.10$, $p \leq 0.001$

Random-effects DerSimonian-Laird model
Figure S3A: 1-year overall survival

<table>
<thead>
<tr>
<th>Study ID</th>
<th>1-year</th>
<th>RR with 95% CI</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sarah Pozzi (2013)</td>
<td>19</td>
<td>0.92 [0.75, 1.13]</td>
<td>30.21</td>
</tr>
<tr>
<td>Anna Candoni (2017)</td>
<td>66</td>
<td>0.54 [0.38, 0.77]</td>
<td>21.29</td>
</tr>
<tr>
<td>Anna Candoni (2017)</td>
<td>16</td>
<td>0.62 [0.33, 0.83]</td>
<td>16.70</td>
</tr>
<tr>
<td>Silvia Park (2020)</td>
<td>165</td>
<td>0.87 [0.73, 1.04]</td>
<td>31.80</td>
</tr>
</tbody>
</table>

Overall

Heterogeneity: τ² = 0.04, I² = 70.96%, H² = 3.44

Test of θ = 0: Q(3) = 10.33, p = 0.02
Test of θ = 0: z = -2.41, p = 0.02

Random-effects DerSimonian-Laird model
Figure 3B: 2-year overall survival

<table>
<thead>
<tr>
<th>Study ID</th>
<th>High survival</th>
<th>Low survival</th>
<th>RR (95% CI)</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sarah Pozzi (2013)</td>
<td>37 30 40 15</td>
<td></td>
<td>0.76 [0.58, 0.99]</td>
<td>32.36</td>
</tr>
<tr>
<td>Anna Candoni (2017)</td>
<td>13 28 65 16</td>
<td></td>
<td>0.40 [0.25, 0.63]</td>
<td>19.79</td>
</tr>
<tr>
<td>Anna Candoni (2017)</td>
<td>9 12 28 5</td>
<td></td>
<td>0.51 [0.30, 0.84]</td>
<td>17.34</td>
</tr>
<tr>
<td>Silvia Park (2020)</td>
<td>26 30 147 54</td>
<td></td>
<td>0.63 [0.47, 0.85]</td>
<td>30.50</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td>0.59 [0.45, 0.77]</td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: $\chi^2 = 0.04, I^2 = 53.53\%, H^2 = 2.15$
Test of $\theta = 0$: $Q(3) = 6.46, p = 0.09$
Test of $\theta = 0$: $z = -3.88, p \leq 0.001$

Random-effects DerSimonian-Laird model
### Figure 3C: 4-year overall survival

<table>
<thead>
<tr>
<th>4-year study ID</th>
<th>high</th>
<th>low</th>
<th>RR with 95% CI</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>sur</td>
<td>nosur</td>
<td>nosur</td>
<td></td>
</tr>
<tr>
<td>Sarah Pozzi (2013)</td>
<td>29</td>
<td>38</td>
<td>35 20</td>
<td>0.68 [ 0.48, 0.95]</td>
</tr>
<tr>
<td>Anna Cardoni (2017)</td>
<td>10</td>
<td>31</td>
<td>59 22</td>
<td>0.33 [ 0.19, 0.58]</td>
</tr>
<tr>
<td>Anna Cardoni (2017)</td>
<td>6</td>
<td>15</td>
<td>25 8</td>
<td>0.38 [ 0.19, 0.76]</td>
</tr>
<tr>
<td>Silvia Park (2020)</td>
<td>25</td>
<td>31</td>
<td>135 66</td>
<td>0.66 [ 0.49, 0.90]</td>
</tr>
</tbody>
</table>

Overall

Heterogeneity: $\chi^2 = 0.66, \hat{I}^2 = 55.88\%, H_t = 2.27$

Test of $H_0: \phi = 0$: $Q(3) = 6.80, p = 0.08$

Test of $H_0: \phi = 0$: $z = -3.70, p \leq 0.001$

Random-effects DerSimonian-Laird model
### Figure 3D: 7-year overall survival

<table>
<thead>
<tr>
<th>7-year study ID</th>
<th>high sur</th>
<th>nosur</th>
<th>low sur</th>
<th>nosur</th>
<th>RR with 95% CI</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sarah Pozzi(2013)</td>
<td>27</td>
<td>40</td>
<td>35</td>
<td>20</td>
<td>0.63 [0.44, 0.90]</td>
<td>38.08</td>
</tr>
<tr>
<td>Anna Candon(2017)</td>
<td>8</td>
<td>33</td>
<td>41</td>
<td>34</td>
<td>0.36 [0.19, 0.69]</td>
<td>17.61</td>
</tr>
<tr>
<td>Anna Candon(2017)</td>
<td>3</td>
<td>18</td>
<td>20</td>
<td>13</td>
<td>0.24 [0.08, 0.70]</td>
<td>7.60</td>
</tr>
<tr>
<td>Silvia Park (2020)</td>
<td>20</td>
<td>36</td>
<td>127</td>
<td>74</td>
<td>0.57 [0.39, 0.82]</td>
<td>36.71</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.51 [0.37, 0.70]</td>
<td></td>
</tr>
</tbody>
</table>

**Heterogeneity:** $\chi^2 = 0.04, I^2 = 35.13\%, H^2 = 1.54$

**Test of $\theta = 0$:** $Q(3) = 4.62, p = 0.20$

**Test of $\theta = 0$:** $z = -4.19, p \leq 0.001$

Random-effects DerSimonian-Laird model

![Graph showing survival data with 7-year overall survival rates and corresponding weights]
Figure 4A: 1-year disease-free survival

<table>
<thead>
<tr>
<th>1-year study ID</th>
<th>250copies</th>
<th>100copies</th>
<th>Risk Ratio with 95% CI</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anna Candoni (2017.10) Josep F. Nomdedeu(2018)</td>
<td>80</td>
<td>42</td>
<td>100</td>
<td>84</td>
</tr>
<tr>
<td>Anna Candoni(2017.01)Josep F. Nomdedeu(2018)</td>
<td>34</td>
<td>20</td>
<td>100</td>
<td>84</td>
</tr>
</tbody>
</table>

Overall

Heterogeneity: $I^2 = 0.00$, $I^1 = 0.00\%$, $H^2 = 1.00$
Test of $G_0$: $Q(1) = 0.07$, $p = 0.79$
Test of $6_0$: $z = 2.30$, $p = 0.02$

Random-effects DerSimonian-Laird model
### Figure 4B: 4-year disease-free survival

<table>
<thead>
<tr>
<th>Study ID</th>
<th>250 copies</th>
<th>100 copies</th>
<th>Risk Ratio with 95% CI</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anna Candoni (2017.10) Josep F. Nombiedeu (2018)</td>
<td>62</td>
<td>60</td>
<td>80</td>
<td>104</td>
</tr>
<tr>
<td>Anna Candoni (2017.01) Josep F. Nombiedeu (2018)</td>
<td>28</td>
<td>26</td>
<td>80</td>
<td>104</td>
</tr>
</tbody>
</table>

Overall

Heterogeneity: $I^2 = 0.00$, $H^2 = 1.00$

Test of $H = 0$: $Q(1) = 0.01$, $p = 0.92$

Test of $H = 0$: $Z = 1.70$, $p = 0.09$

Random-effects DerSimonian-Laird model
Figure 5A: 1-year overall survival

<table>
<thead>
<tr>
<th>Study ID</th>
<th>250 copies</th>
<th>100 copies</th>
<th>1-year</th>
<th>Risk Ratio with 95% CI</th>
<th>Weight (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anna Candoni (2017.10)</td>
<td>89</td>
<td>33</td>
<td>109</td>
<td>75</td>
<td>1.23 [1.05, 1.45]</td>
</tr>
<tr>
<td>Anna Candoni (2017.01)</td>
<td>38</td>
<td>16</td>
<td>109</td>
<td>75</td>
<td>1.19 [0.96, 1.47]</td>
</tr>
<tr>
<td>Anna Candoni (2017.10) Sarah Pozzi (2013)</td>
<td>80</td>
<td>33</td>
<td>93</td>
<td>29</td>
<td>0.96 [0.83, 1.11]</td>
</tr>
<tr>
<td>Anna Candoni (2017.01) Sarah Pozzi (2013)</td>
<td>38</td>
<td>16</td>
<td>93</td>
<td>29</td>
<td>0.92 [0.76, 1.13]</td>
</tr>
<tr>
<td>Overall</td>
<td>1.06</td>
<td></td>
<td></td>
<td>0.92 [1.23]</td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: $I^2 = 0.01, I^2 = 63.21\%, H^2 = 2.72$

Test of $G = 0$: $Q(3) = 8.15, p = 0.04$

Test of $G = 0$: $z = 0.64, p = 0.40$
Figure SB: 4-year overall survival
Figure 6: Funnel plot of disease free survival rate

A: 1-year
B: 4-year
Figure 7: Funnel plot of overall survival

A: 1-year
B: 4-year
Figure 8: The TSA curve

A: 1-year DFS

TSA is a One-sided Lower graph.
B: 4-year OS

TSA is a One-sided Lower graph.