MEDICAL UNIVERSITY
„PROF. DR. PARASKEV STOYANOV“ – VARNA

FACULTY OF MEDICINE
SECOND DEPARTMENT OF INTERNAL DISEASES
ES HAEMATOLOGY

Svilena Angelova Atanasova, MD

ROLE OF SELECTED PLASMA MICRORNAS AS DIAGNOSTIC AND PROGNOSTIC BIOMARKERS IN MYELODYSPLASTIC SYNDROME

THESIS SUMMARY

of PhD dissertation

for obtaining the educational and scientific degree „Doctor”
Scientific specialty “Hematology and blood transfusion”

Scientific supervisor:
Prof. Ilina Micheva, MD, PhD

Varna, 2025 г.

The dissertation consists of 176 pages and includes 29 tables and 19 figures. The literature review contains 290 references.

The dissertation was discussed in an open session of the Departmental Council of the Second Department of Internal Medicine at Medical University "Prof. Dr. Paraskev Stoyanov" – Varna (protocol № 2/19.02.2025). It has been accepted and forwarded for defense before a Scientific panel composed of:

External Members:

1. *Prof. Janet Grudeva-Popova, M.D., PhD*
2. *Prof. Gueorgui Balatzenko, M.D., PhD*
3. *Assoc. Prof. Ivan Kindekov, M.D., PhD*

Internal Members:

4. *Assoc. Prof. Trifon Chervenkov, M.D., PhD*
5. *Assoc. Prof. Eleonora Dimitrova Gospodinova, M.D., PhD*

Back-up Members:

1. Assoc. Prof. Antonio Antonov, M.D., PhD
2. Prof. Nikolay Conev, M.D., PhD

The public defense of the dissertation will take place on 28.05.2025г.

The materials of the defense are available at the library of the Medical University "Prof. Dr. Paraskev Stoyanov" - Varna and has been published on the website of MU – Varna.

CONTENTS

ABBREVIATIONS:

AIC – Akaike Information Criterion
APC – Adenomatous Polyposis Coli
AUC – Area Under the Curve
BCL2 – B-Cell Lymphoma 2
ERK1/2 – Extracellular Signal-Regulated Kinases 1/2
FoxO – Forkhead Box O
GATA1 – GATA Binding Protein 1
GLUT – Glucose Transporter
HR – Hazard Ratio
MCL1 – Myeloid Cell Leukemia 1
OR – Odds Ratio
PI3K/AKT – Phosphoinositide 3-Kinase / Protein Kinase B
Pseudo-R^2 – Pseudo Coefficient of Determination
PTEN – Phosphatase and Tensin Homolog
R-IPSS – Revised International Prognostic Scoring System
ROS – Reactive Oxygen Species
SMAD2 – Mothers Against Decapentaplegic Homolog 2
SMAD4 – Mothers Against Decapentaplegic Homolog 4
TET2 – Ten-Eleven Translocation 2
TGF-β – Transforming Growth Factor Beta
TGFBR1 – Transforming Growth Factor Beta Receptor 1
TGFBR2 – Transforming Growth Factor Beta Receptor 2
VEGF – Vascular Endothelial Growth Factor
VIF – Variance Inflation Factor
Wnt – Wingless-Related Integration Site (signaling pathway)

I. Introduction

Myelodysplastic syndromes (MDS) are a heterogeneous group of clonal hematopoietic disorders characterized by ineffective hematopoiesis, peripheral cytopenias, and an increased risk of progression to acute myeloid leukemia. Diagnosing MDS remains a clinical challenge, as it requires a complex integration of clinical, morphological, and genetic findings. Many patients with cytopenia have genetic alterations or dysplasia that may mimic MDS but fall short of meeting established diagnostic criteria. Conditions such as clonal cytopenia of undetermined significance (CCUS) and clonal hematopoiesis of indeterminate potential (CHIP) represent diagnostic gray zones where clonality is present without definitive evidence of MDS. Additionally, factors such as germline mutations, concurrent medical conditions, or treatments (e.g., cytotoxic drugs or immune-modulating therapies) can further complicate the diagnostic process by inducing dysplastic changes or cytopenias. Moreover, some chromosomal abnormalities are considered diagnostic of MDS even in the absence of significant dysplasia, while others are insufficient on their own, contributing to diagnostic ambiguity. These challenges highlight the need for repeated and comprehensive diagnostic evaluations, including bone marrow sampling, advanced flow cytometry, and molecular testing, to reliably differentiate MDS from other overlapping conditions. Moreover, MDS is a highly heterogeneous disease, with patient survival ranging from only a few months to several years. This significant variability in clinical course underscores the need for identifying reliable prognostic biomarkers that can support individualized therapy and improve patient outcomes.

Furthermore, a deeper understanding of MDS pathogenesis could facilitate the development of reliable diagnostic and prognostic biomarkers, ultimately contributing to earlier and more accurate disease detection.

MicroRNAs (miRNAs) are small, non-coding RNA molecules, typically 19–25 nucleotides in length, that regulate gene expression by binding to the 3' untranslated regions of target messenger RNAs (mRNAs), leading to translational repression or degradation. MicroRNAs have emerged as critical regulators in the pathogenesis of MDS, influencing key cellular processes such as differentiation, proliferation, and apoptosis. Dysregulation of miRNA expression has been reported in MDS, with distinct miRNA profiles identified in patients compared to healthy individuals. These alterations in miRNA expression are associated with various disease aspects, including hematopoietic dysfunction, chromosomal abnormalities, and disease progression. Several studies have demonstrated the potential utility of miRNAs as diagnostic and prognostic biomarkers, offering insights into disease stratification and risk assessment. Despite the growing body of evidence supporting the role of miRNAs in MDS, their exact functional implications and clinical applications require further exploration. While most studies focus on bone marrow-derived miRNAs, the potential of circulating miRNAs as minimally invasive biomarkers remains a promising area of research. Advances in high-throughput sequencing and bioinformatics tools have facilitated the identification of novel miRNA signatures, which could enhance diagnostic accuracy and patient stratification. Despite the growing interest in microRNAs as potential biomarkers for MDS, the clinical significance of specific miRNAs remains under investigation.

II. Aims and objectives

1. Aims

The aim of this study is to assess the expression levels of five selected microRNAs—miR-22, miR-144, miR-16, miR-451a, and let-7a—in patients with MDS, to perform a comparative analysis of the results in high- and low-risk MDS groups as well as between MDS patients and healthy controls, and to determine their prognostic significance.

2. Objectives

In line with the stated aims, the following main objectives have been formulated:

1. Selection of a cohort of MDS patients and assessment of their demographic and clinical-laboratory characteristics, as well as selection of healthy controls with demographic characteristics similar to the patient group.
2. Analysis of plasma samples from patients and healthy controls, and measurement of the expression levels of the five selected microRNAs in both groups.
3. Comparison of the expression levels of the five investigated microRNAs between MDS patients and healthy controls.
4. Analysis of the five microRNAs in relation to patients' demographic characteristics.

5. Analysis of the five selected microRNAs between different MDS subtypes and according to the R-IPSS risk stratification.
6. Correlation analysis to examine relationships between the five chosen microRNAs and various hematological and biochemical parameters in MDS patients.
7. Determination of the diagnostic value of the five selected microRNAs.
8. Evaluation of the predictive role of the five microRNAs for distinguishing high- vs. low-risk MDS according to the R-IPSS, and for different MDS types.
9. Assessment of the prognostic value of the five plasma microRNAs in MDS.

III. Materials and methods

1. Facilities and Resources for the Dissertation

- Clinic of Hematology, UMHAT "Sveta Marina" – Varna
- Central Clinical Laboratory, UMHAT "Sveta Marina" – Varna
- Immunological Laboratory, UMHAT "Sveta Marina" – Varna
- Genetics Laboratory, UMHAT "Sveta Marina" – Varna

2. Study Participants

This scientific investigation was prospective and included 50 participants—40 patients with a confirmed diagnosis of myelodysplastic syndrome and 10 healthy controls. The study was conducted at the Clinic of Hematology, UMHAT “Sveta Marina” Varna, following a positive evaluation by the the Commission for Ethics of Scientific Research at the Medical University “Prof. Dr. Paraskev Stoyanov” – Varna. All participants were enrolled after providing written informed consent.

3.Patient Selection

Patients were enrolled based on the study’s inclusion and exclusion criteria.

3.1. Inclusion Criteria:

• Age over 18 years

• Patients diagnosed with myelodysplastic syndrome according to the 2016 WHO criteria

•Expressed willingness to participate in the study, documented by a signed informed consent form

3.2. Exclusion Criteria:
• Patients under 18 years of age
• Patients not meeting the inclusion criteria
• Expressed unwillingness to participate in the study

4. Selection of Healthy Controls

The control group included 10 clinically healthy individuals, matched to the patient group by demographic characteristics, all over 18 years of age, who provided informed consent.

5. Routine Clinical Investigations. Sample Collection and Storage. Analytical Methods

All patients underwent a clinical examination, including medical history and physical assessment. Routine laboratory tests were performed, which included peripheral blood counts, lactate dehydrogenase (LDH), beta-2 microglobulin, ferritin, and erythropoietin.

5.1. Specific Methods of Investigation
To perform a specific analysis of the levels of the selected five microRNAs, a venous blood sample was collected in two EDTA tubes (3 mL each) and one serum Vacutainer (8 mL). The samples were centrifuged at 1,500 rpm for 15 minutes at 4 °C, and the plasma was separated, aliquoted into 500 µL portions, and stored at −80 °C until analysis.

MicroRNAs were isolated from 200 µL of blood plasma using the commercial miRNeasy Serum/Plasma Kit (QIAGEN, Germany), following the manufacturer's protocol. For

normalization purposes, 3.5 µL (1.6×10^8 copies/µL) of the control microRNA C. elegans miR-39 was added to each sample. The samples were eluted in 25 µL of RNase-free water. The isolation procedure included the following steps:

1. The samples were thawed at room temperature (15–25 °C).
2. One thousand microliters (5 volumes) of QIAzol Lysis Reagent was added to each sample, followed by vortexing to homogenize.
3. The tube containing the lysate was incubated at room temperature (15–25 °C) for 5 minutes. Then, 3.5 µL (1.6×10^8 copies/µL) of control microRNA C. elegans miR-39 was added to each sample.
4. The lysate was treated with 200 µL of chloroform (an amount equal to the volume of the initial sample), and the tubes were vortexed for 15 seconds to ensure thorough mixing.
5. The tube was incubated at room temperature for 2–3 minutes.
6. The sample was centrifuged for 15 minutes at 12,000 × g at 4 °C. After centrifugation, the sample separated into three phases: a clear, aqueous upper phase containing RNA; a white interphase; and a lower, red organic phase.
7. Six hundred microliters of the upper aqueous phase were transferred to a new tube, avoiding transfer of material from the interphase. Next, 900 µL (1.5 volumes) of 100% ethanol was added, and the mixture was homogenized by pipetting.
8. The sample was transferred in portions (up to 700 µL each) onto an RNeasy MinElute column placed in a 2 mL collection tube, then centrifuged at 8,000 × g for 15

seconds at room temperature. The flow-through was discarded. This step was repeated until the entire sample had passed through the column.

9. The column membrane was washed with 700 µL of RWT buffer. The column was then centrifuged for 15 seconds at 8,000 × g, and the flow-through was discarded.
10. Five hundred microliters of RPE buffer were added onto the RNeasy MinElute column. The column was centrifuged for 15 seconds at 8,000 × g to wash it, and the flow-through was discarded.
11. Five hundred microliters of 80% ethanol were added onto the RNeasy MinElute column. The column was then centrifuged for 2 minutes at 8,000 × g to wash the spin-column membrane. The collection tube was discarded afterward.
12. The RNeasy MinElute column was placed in a new 2 mL collection tube. To dry the membrane, the column was centrifuged at maximum speed (25,000 × g) for 5 minutes with the cap open. The collection tube was discarded.
13. The RNeasy MinElute column was placed in a new 1.5 mL tube. Elution of the microRNAs was performed by adding 25 µL of RNase-free water directly to the center of the spin-column membrane, followed by centrifugation at maximum speed (25,000 × g) for 1 minute.

Reverse Transcription

The eluted microRNAs were subjected to reverse transcription using the commercial miRCURY LNA RT Kit, following the manufacturer's protocol. The reaction mixture consisted of 1.0 µL of eluted RNA and 9.0 µL of master mix, containing 2.0 µL

of 5× reaction buffer, 6.0 µL of RNase-free water, and 1.0 µL of 10× enzyme mix. The reaction was carried out at 42 °C for 60 minutes, followed by 5 minutes at 95 °C to inactivate the enzyme.

Quantitative PCR

Quantitative measurement of the microRNAs was performed by real-time quantitative polymerase chain reaction (Real-Time PCR) using the commercial miRCURY LNA SYBR Green PCR Kit (200) and (600), along with ready-to-use primers for the target microRNAs (miRCURY LNA miRNA PCR Assay), according to the manufacturer's protocol. The reaction was carried out with 3.0 µL of cDNA, diluted 30-fold, combined with 7.0 µL of master mix (5.0 µL 2× miRCURY SYBR® Green Master Mix, 0.05 µL ROX dye, 1.0 µL primer mix, and 0.95 µL RNase-free water) in a total volume of 10 µL. Each reaction was run in triplicate for six target microRNAs in 384-well plates. The miRCURY LNA miRNA PCR Assay primers (Catalog No. 339306, QIAGEN, Germany) used were as follows (GeneGlobe reference number in parentheses): cel-miR-39-3p (YP00203952), hsa-miR-144-3p (YP00204754), hsa-miR-22-3p (YP00204606), hsa-let-7a-5p (YP00205727), hsa-miR-16-5p (YP00205702), hsa-miR-451a (YP02119305). The thermal cycling parameters were as follows: initial hold at 95 °C for 2 minutes to activate the enzyme; 40 cycles of 10 seconds at 95 °C, followed by 60 seconds at 56 °C with fluorescence detection. A melting curve analysis was then performed to confirm amplification specificity: initial denaturation at 95 °C for 15 seconds, cooling to 60 °C for 60 seconds, then ramping up to 95 °C at +0.05 °C per second with fluorescence monitoring.

All reactions were run on a QuantStudio Dx (Applied Biosystems, USA), and the cycle threshold (Ct) for each sample was recorded. The relative expression of the target microRNAs was calculated using the ΔΔCt method, normalized to the reference microRNA C. elegans miR-39.

The relative concentration of the investigated target microRNAs was determined by the ΔΔCt method [276], using C. elegans miR-39 as a normalization control and a reference sample—which was the arithmetic mean Ct of all individuals—calculated in Microsoft Office Excel 2016 and presented as a ratio relative to the reference sample.

6. Statistical Analysis

The study data were processed using specialized statistical software (IBM SPSS Statistics v.24, GraphPad Prism version 10.4.1, Python v.3.8 and Google Sheets) to evaluate the diagnostic and prognostic roles of the selected plasma microRNAs in MDS. The following methods were employed:

Descriptive Statistics: Used to summarize the main characteristics of the data by calculating means, standard deviations, medians, and percentage distributions.

Shapiro–Wilk Test: Applied to check for normality of the data distribution in a given group. It determines whether the distribution significantly deviates from a normal distribution. A p-value > 0.05 indicates the data follow a normal distribution, whereas $p \leq 0.05$ indicates the distribution is not normal.

Mann–Whitney Test: A nonparametric test used to compare medians between two independent groups. It is suitable when the data do not follow a normal distribution. This test evaluates

whether there is a significant difference in the ranks of values between the two groups.

Effect Size (r): Calculated in the Mann–Whitney test to assess the practical significance of the differences between groups. The r value is interpreted as small ($r < 0.3$), moderate (r between 0.3 and 0.5), or large ($r > 0.5$). This measure provides additional information on the magnitude of the group differences, even when statistical significance is not observed.

Correlation Analysis: Spearman's Correlation Coefficient: Used to evaluate the relationship between two quantitative variables, regardless of data distribution. This method measures the direction and strength of the monotonic relationship between variables. The results are interpreted according to the following correlation strengths:

- Weak: 0.1–0.3
- Moderate: 0.3–0.5
- Marked: 0.5–0.7
- Strong: 0.7–0.9
- Very strong: 0.9–1.0

Regression Analyses:

- **Simple Linear Regression:** Applied to evaluate the linear relationship between one independent and one dependent variable. The results include determination coefficients (R^2), which show the extent to which one variable explains the variation in another.
- **Logistic Regression:** Used to assess the probability of a binary event occurring based on one or more independent variables.

- **Cox Regression Analysis:** Used to evaluate the relationship between the levels of the investigated microRNAs and time to an event (e.g., death) in MDS patients. This semiparametric analysis does not require an assumption about the shape of the baseline hazard function, making it flexible and suitable for survival data. The Cox model computes hazard ratios (HR) indicating the effect of each independent variable on the risk of the event. It is applicable to censored data (where the event has not occurred for all participants) and is widely used to assess the prognostic significance of biomarkers, clinical characteristics, and other factors.
- **LASSO Analysis:** Employed to assess the predictive value of the investigated microRNAs and to identify the most significant among them. The method addresses multicollinearity among microRNAs, which can affect the stability and interpretation of the model. LASSO enables model reduction by eliminating less significant variables, thereby identifying the key microRNAs that contribute to diagnostic and prognostic accuracy.

ROC Analysis: A Receiver Operating Characteristic (ROC) analysis was performed to assess the diagnostic accuracy of the microRNAs. The area under the curve (AUC) was calculated, along with optimal cut-off values based on the Youden Index.

Kaplan–Meier Curves: Used to evaluate time-to-event outcomes, such as survival.

The results were graphically presented using:

- **Box Plots:** Used to compare the distributions of microRNAs across different groups, providing information on the median, interquartile range (IQR), and potential outliers.
- **Scatter Plots:** Used to visualize correlation relationships.
- **ROC Curves:** Used to evaluate the diagnostic accuracy.

All statistical tests were conducted at a significance level of $\alpha = 0.05$. The results were interpreted using p-values and confidence intervals.

IV Results

1. Demographic and Clinical-Laboratory Characteristics of the Patients

A total of 40 patient samples and 10 healthy controls were included. Just over one-third of the patients were female, and about two-thirds were male: 15 (37.5%) women and 25 (62.5%) men (Figure 1).

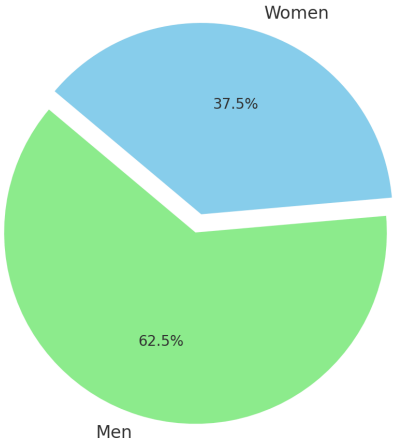


Figure 1. Distribution of patients with MDS by gender.

The average age of the patients is 71 years (41-88), with 67.5% (n=27) being over 70 years old. The predominant age group is between 70-75 years, accounting for 40% (n=16) of the patients. The average age for women is approximately 70 years (69.93), while for men, it is 71 years (70.88). The age distribution is presented in Figure 2.

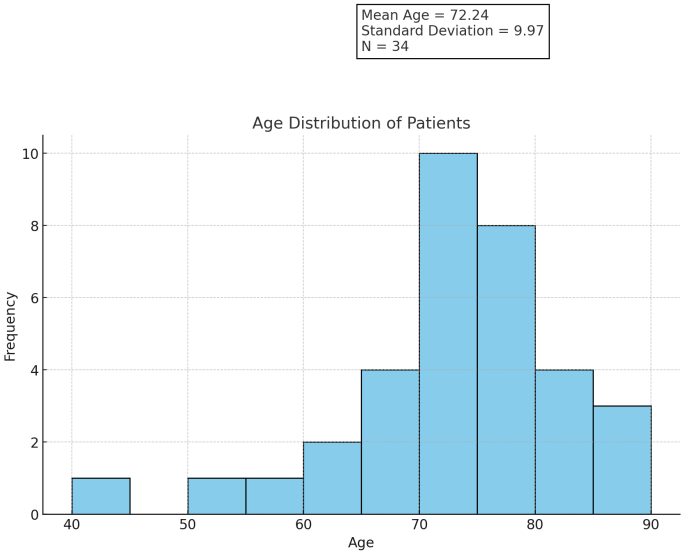


Figure 2. Age Distribution of patients with MDS

The distribution by age and gender shows that the average age for women is approximately 70 years (69.93), while for men it is 71 years (70.88). A hypothesis test was conducted to examine the difference in age between the two genders (Table 1). The research hypothesis assumes that a difference exists between the groups. The null and alternative hypotheses are formulated as follows:

H_0: $\mu_1 = \mu_2$
H_1: $\mu_1 \neq \mu_2$

A significance level of $\alpha = 0.05$ was set. Since the P-value (0.763) is greater than α (0.05), there is no basis to reject H_0 in favor of H_1 at a significance level of $\alpha = 0.05$. This indicates that the difference in average age between men and women is statistically insignificant.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Age	Equal variances assumed	4,422	,042	-,303	38	,763	-,947	3,120	-7,262	5,369
	Equal variances not assumed			-,263	18,904	,796	-,947	3,602	-8,489	6,595

Table 1: Results of the hypothesis test for age difference between the two genders.

According to the 2016 WHO classification, 47.5% (n=19) of patients have MDS with multilineage dysplasia, 10% (n=4) have MDS with ring sideroblasts, 10% (n=4) have MDS with del(5q), and 32.5% (n=13) have MDS with excess blasts (Figure 3).

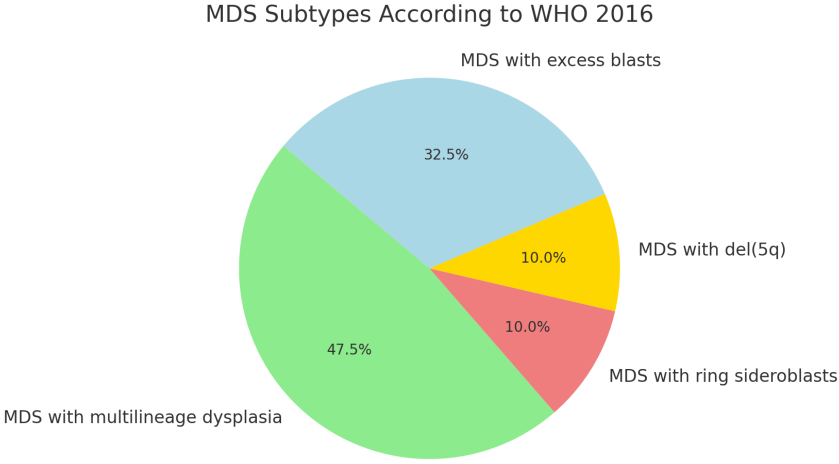


Figure 3: Distribution of Patients by MDS Subtype according to WHO 2016.

According to the risk stratification by R-IPSS, the patients are distributed as follows: very low risk – 5% (n=2), low risk – 25%

(n=10), intermediate risk – 27.5% (n=11), high risk – 22.5% (n=9), and very high risk – 20% (n=8) (Figure 4).

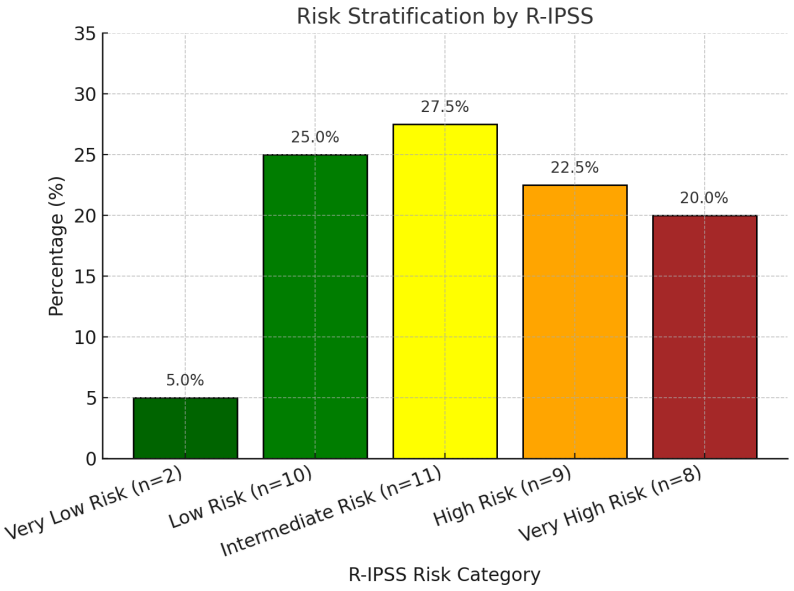


Figure 4: Distribution of patients according to risk stratification by R-IPSS

For statistically significant data analysis, patients were grouped into two risk categories: low-risk MDS, including the R-IPSS groups very low, low, and intermediate risk (≤ 3.5 p), and high-risk MDS, including the R-IPSS groups intermediate (> 3.5 p), high and very high risk.

According to cytogenetic risk, 2.5% (n=1) of patients had a very good risk, 55% (n=22) had a good risk, 27.5% (n=11) had an intermediate risk, 7.5% (n=3) had a poor risk, and 5% (n=2) had a very poor risk (Figure 5).

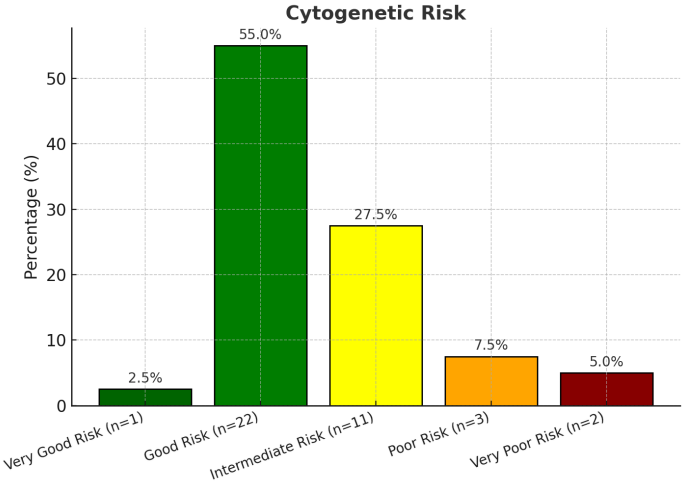


Figure 5: Distribution of patients according to cytogenetic risk.

Regarding the number of cytopenias, 40% (n=16) have only one, 37.5% (n=15) have two, and 22.5% (n=9) have three cytopenias. Concerning the percentage of bone marrow blasts, patients are classified into three groups: <5%, 5-9%, and 10-19%. The first group includes 65% (n=26), the second group 5% (n=2), and the third group 30% (n=12). Based on the treatment received at the time of analysis, patients were divided into: active monitoring 7.5% (n=3), supportive therapy 32.5% (n=13), those receiving therapy aimed at increasing hemoglobin levels 25% (n=18) – erythropoietin, lenalidomide, luspatercept, and those receiving therapy targeting blast reduction in the bone marrow 35% (n=14) – azacitidine and intensive chemotherapy (Figure 6).

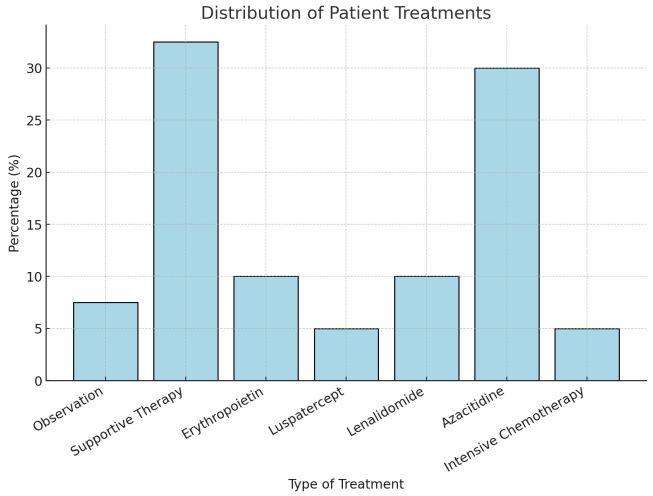


Figure 6: Distribution of patients according to the received treatment.

At the time of analysis, 50% (n=20) of the patients were alive, while 50% (n=20) had deceased (Figure 7). The median overall survival was reached at 35 months.

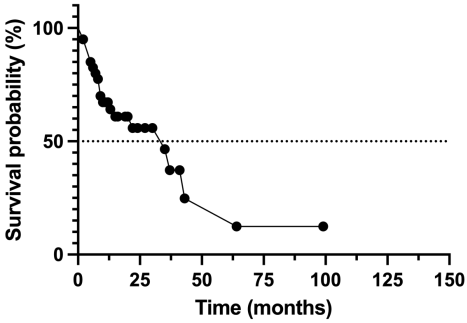


Figure 7: Overall survival of patients.

All demographic and clinic-laboratory data of the patients are presented in Table 2.

Characteristics	N(%) / Mean
Age	70,5 (44-88)
Sex (M:F)	25:15 (62,5:37,5)
WHO classification MDS with multilineage dysplasia MDS with ring sideroblasts MDS with del(5q) MDS excess blasts	 19 (47,5) 4 (10) 4 (10) 13 (32,5)
R-IPSS Very low Low Intermediate High Very high	 2 (5) 10 (25) 11 (27,5) 9 (22,5) 8 (20)
Cytogenetic risk Very good Good Intermediate Poor Very poor Unknown	 1 (2,5) 22 (55) 11 (27,5) 3 (7,5) 1 (2,5) 2 (5)
Cytopenias One	 16 (40)

Two Three	15 (37,5) 9 (22,5)
Marrow blasts <5% 5-9% 10-19%	 26 (65) 2 (5) 12 (30)
Treatment Observation Supportive therapy Therapy aimed at increasing Hb Erythropoietin Luspatercept Lenalidomide Therapy targeting blast reduction Azacitidine Intensive chemotherapy	 3 (7,5) 13 (32,5) 10 (25) 4 (10) 2 (5) 4 (10) 14 (35) 12 (30) 2 (5)
Hemoglobin (g/L)	79,43 ± 17,52 (28-115)
Reticulocytes (%)	2,495 ± 2,113 (0,32 - 9,06)
Leukocytes ($x10^9$/L)	4,01 ± 2,188 (1,54 - 10,32)
Neutrophils ($x10^9$/L)	2,012 ± 1,553 (0,18 - 7,79)
Thrombocytes ($x10^9$/L)	170,8 ± 164,3 (15 - 859)
LDH (U/L)	467,2 ± 329,6 (187 - 1977)
Beta2-microglobulin (mg/L)	3,413 ± 1,518 (1,4 - 8,1)

Erythropoietin (U/l)	320,3 ± 286,5 (23,5 - 751)
Ferritin (µg/L)	840,2 ± 626,7 (11,31 - 3172)

Table 2: Demographic and clinic-laboratory characteristics of patients with MDS.

2. Analysis of microRNA levels in patients and healthy controls

To assess whether the levels of the studied microRNAs in patients with MDS and healthy controls follow a normal distribution, the Shapiro-Wilk test was conducted.

Shapiro-Wilk Test

To evaluate the diagnostic potential of selected plasma microRNAs in MDS patients, a comparative analysis was performed between the two groups—MDS patients and healthy controls. Initially, the Shapiro-Wilk test was applied to determine whether the levels of the studied microRNAs in both groups follow a normal distribution (Table 3).

The Shapiro-Wilk test is a statistical method used to test the hypothesis that a given sample originates from a normally distributed population. It is particularly suitable for small to medium sample sizes and is one of the most sensitive tests for normality.

- **Null hypothesis (H_0):** The data follows a normal distribution.
- **Alternative hypothesis (H_1):** The data does not follow a normal distribution.

If the p-value is less than α = 0.05, the null hypothesis is rejected, indicating that the distribution is not normal.

MicroRNA	Група	W	p-стойност	
MiR-22	Patients	0.6917	<0.0001	No
	Controls	0.8907	0.1726	Yes
MiR-144	Patients	0.8383	<0.0001	No
	Controls	0.9079	0.2666	Yes
MiR-16	Patients	0.898	0.0017	No
	Controls	0.9088	0.2729	Yes
Let-7a	Patients	0.2235	<0.0001	No
	Controls	0.9571	0.7526	Yes
MiR-451a	Patients	0.8845	0.0007	No
	Controls	0.9218	0.3721	Yes

Table 3: Results of the Shapiro-Wilk test for normal distribution of data in MDS patients and healthy controls.

In the analysis of the normality of the distribution of microRNA levels using the Shapiro-Wilk test, it was found that all microRNAs in patients showed significant deviations from a normal distribution ($p<0.05$). At the same time, the values in healthy controls passed the normality test ($p>0.05$). This difference can be explained by several factors. First, the small sample size in healthy controls (n=10) may not be sufficiently

representative of the entire population, making it more difficult to detect deviations from normality. Additionally, healthy controls are typically a more homogeneous group, contributing to more stable and closely clustered microRNA values.

The presence of a non-normal distribution in at least one of the groups requires the use of non-parametric tests for comparison. These tests do not assume normality and are suitable for analyzing data with skewness, deviations, or small sample sizes. Based on these results, non-parametric methods were applied for the subsequent analysis of microRNA levels between patients and healthy controls.

In addition to the statistical analysis using the Mann-Whitney U test, the effect size was also calculated to assess the practical significance of the differences between patients and healthy controls. The effect size (r) is calculated using the formula:

$$r = Z \div \sqrt{N}$$

where Z is the Z-value from the Mann-Whitney test, and N is the total number of observations in both groups. The coefficient r indicates the strength of the difference between the groups, following an interpretation similar to correlation coefficients: $r < 0.1$ is considered a minimal effect, $0.1 \leq r < 0.3$ a small effect, $0.3 \leq r < 0.5$ a medium effect, and $r \geq 0.5$ a large effect. This approach provides additional information that complements the results of the statistical test, allowing for a more detailed interpretation of differences in microRNA levels between patients and healthy controls.

2.1 MicroRNA-22

The first analyzed microRNA was miR-22 (Figure 8). The median level of miR-22 in the patient group was 1.294, while in the control group, it was 1.030. The difference between the two groups was tested using the Mann-Whitney test, which did not detect a statistically significant difference (p=0.4086).

The effect size (*r*), calculated based on the *Z*-value and the total number of observations (*N* = 50), was −0.267, indicating a small effect. These results suggest that the differences in miR-22 levels between MDS patients and healthy controls are both statistically and clinically insignificant.

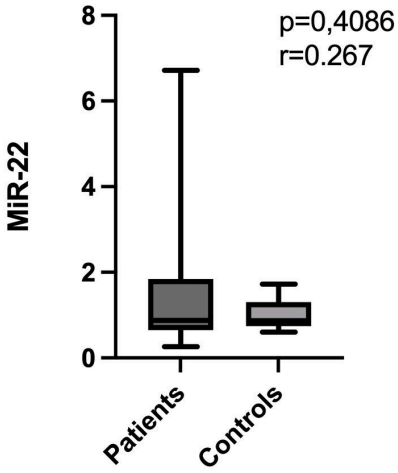


Figure 8: Comparison of miR-22 levels between patients and healthy controls.

2.2 MicroRNA-144

For miR-144, the median expression level in MDS patients is 0.802, whereas in healthy controls, it is 1.933 (Figure 9). This indicates that miR-144 levels are significantly lower in patients

compared to healthy controls. The Mann-Whitney test detected a statistically significant difference between the two groups (p=0.0003), highlighting the potential role of miR-144 as a diagnostic biomarker for MDS.

Additionally, the effect size (r = −1.08) indicates an extremely large effect. The negative sign of r reflects that patient values are lower than those of the controls. This demonstrates the high clinical significance of the differences in miR-144 expression between the two groups.

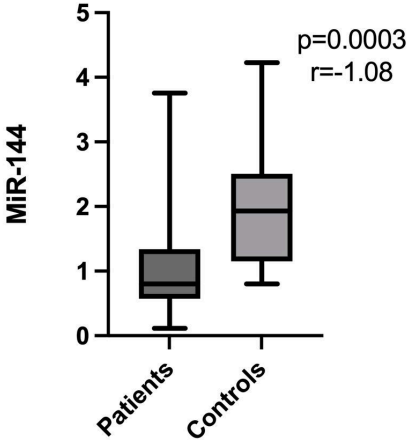


Figure 9: Comparison of miR-144 levels between patients and healthy controls.

2.3 MicroRNA-16

The next analyzed microRNA is miR-16 (Figure 10). The results show a statistically significant difference in expression levels between MDS patients and healthy controls. The

median miR-16 levels in MDS patients are 0.7855, whereas in healthy controls, they are significantly higher at 1.764.

The Mann-Whitney test reported *p*=0.0001, confirming that the difference between the two groups is statistically significant. The effect size (*r*=-1.13) is extremely large. The negative sign of *r* indicates that the values in patients are significantly lower compared to healthy controls. This highlights the clinical significance of the observed differences.

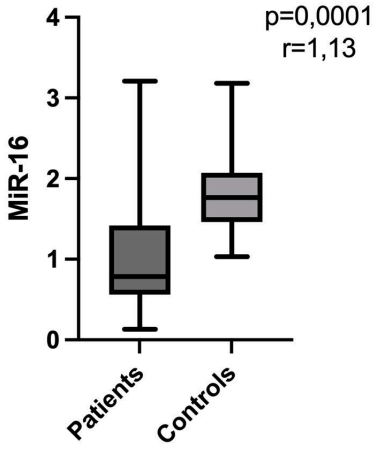


Figure 10: Comparison of miR-16 levels between patients and healthy controls.

2.4 Let-7a

For let-7a, the median expression level in MDS patients is 0.692, whereas in healthy controls, it is significantly higher at 1.669 (Figure 11). One value from the patient group was removed as it was identified as an extreme outlier. The

Mann-Whitney test reported a p-value of 0.0002, confirming that let-7a levels are significantly lower in MDS patients.

The effect size (r=−1.12) is extremely large, emphasizing the significance of the differences and the potential role of let-7a as a diagnostic biomarker.

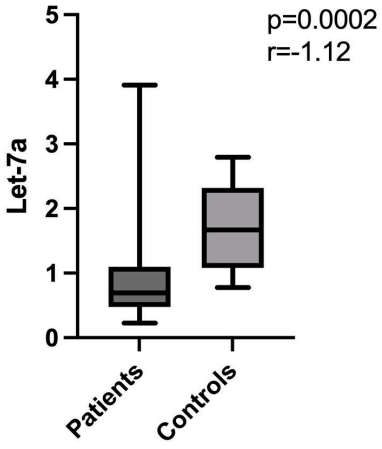


Figure 11: Comparison of let-7a levels between MDS patients and healthy controls.

2.5 MicroRNA-451a

The last analyzed microRNA was miR-451a (Figure 12). The results show a significant difference in the expression levels of miR-451a between MDS patients and healthy controls. The median miR-451a level in MDS patients was 0.8215, while in healthy controls, it was significantly higher at 2.132. The Mann-Whitney test reported p<0.0001, confirming that the difference between the two groups is statistically significant. Furthermore, the effect size (r=−1.28) was exceptionally large,

clearly indicating the clinical significance of the differences in miR-451a levels between the two groups.

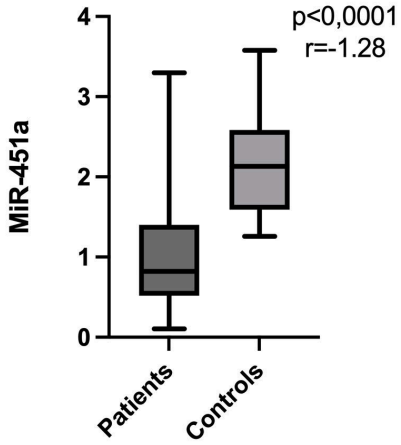


Figure 12: Comparison of miR-451a levels between patients and healthy controls.

Tables 4 and 5 present the summarized results of the Mann-Whitney analysis which compared the microRNA levels between MDS patients and healthy controls.

MicroRNA	Median (Patients)	Median (Controls)	P-value	Effect size (r)	
MiR-22	1.294	1.03	0.4086	-0.267	Not

					significant
MiR-144	0.802	1.933	0.0003	-1.08	Extremely large effect
MiR-16	0.7855	1.764	0.0001	-1.13	Extremely large effect
Let-7a	0.692	1.669	0.0002	-1.12	Extremely large effect
MiR-451a	0.8215	2.132	<0.0001	-1.28	Extremely large effect

Table 4: Results of the Mann-Whitney test comparing miRNA levels between the patient group and healthy controls.

Statistical Hypotheses and Effect Size for Two Independent Samples		**Groups: patients and healthy controls**
MiR-22	statistical significance	**insignificant difference**
	effect size	**small**
MiR-144	statistical significance	**significant difference**
	effect size	**very large**
MiR-16	statistical significance	**significant difference**
	effect size	**very large**

Let-7a	statistical significance	**significant difference**
	effect size	**very large**
MiR-451a	statistical significance	**significant difference**
	effect size	**very large**

Table 5: Schematic Representation of the Mann-Whitney Test Results and the Determined Effect Size.

3. Determination of Threshold Values for MicroRNAs with Statistically Significant Differences between MDS Patients and Healthy Controls

For the microRNAs that showed a statistically significant difference between MDS patients and healthy controls, a ROC (Receiver Operating Characteristic) analysis was performed, and threshold values were determined using Youden's Index. The goal was to identify the optimal diagnostic threshold for distinguishing MDS patients from healthy controls.

Youden's Index is a commonly used method for finding the optimal cut-off value that maximizes the difference between sensitivity and specificity. It is calculated using the formula:

$$\text{Youden Index} = \text{Sensitivity} + \text{Specificity} - 1$$

3.1 Threshold Value Analysis for miR-144

For miR-144, which demonstrated a significant difference between the two groups and a large effect size, the ROC analysis determined an optimal cut-off value of 1.727. At this

cut-off, sensitivity was 92.5%, and specificity was 70% (Figure 13).

The area under the curve (AUC) was 0.8538 (95% CI: 0.7274 - 0.9801), with a p-value of 0.0006.

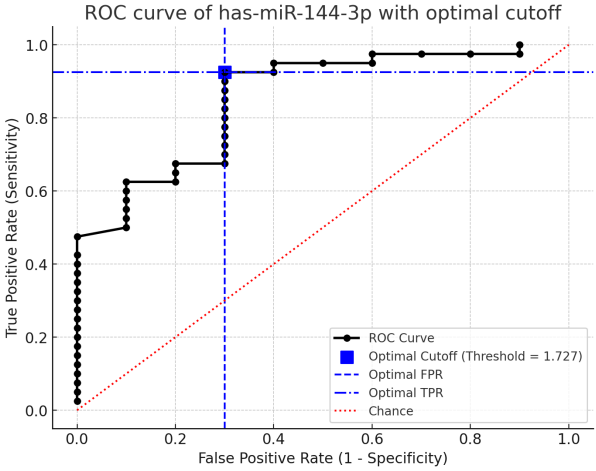


Figure 13: ROC analysis for miR-144 and determination of the diagnostic cut-off.

3.2 Threshold Value Analysis for miR-16

A ROC curve was generated for miR-16 with an AUC of 0.8675 (0.7646 - 0.9704), p = 0.0004, demonstrating good discriminative ability of the model (Figure 14). In the analysis of miR-16, the Youden's Index was calculated with a value of 0.625 for two possible cut-offs – 0.9955 and 1.508. To achieve a better balance between sensitivity and specificity, the cut-off of 1.508 was selected, providing a higher sensitivity (82.5%) while maintaining a specificity of 80%, compared to the cut-off of 0.9955, which had 100% specificity but a lower sensitivity

(62.5%). Higher sensitivity is particularly important in the diagnosis of rare diseases such as MDS.

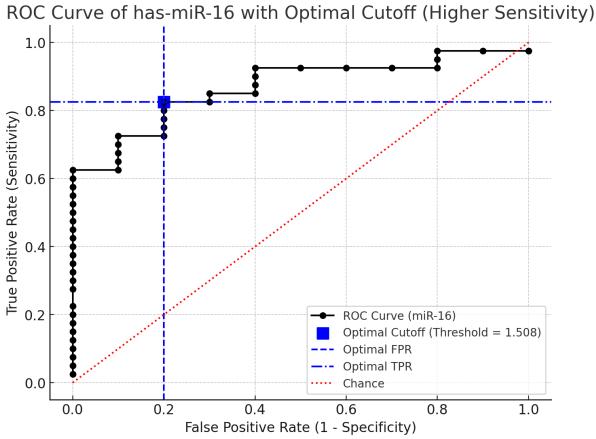


Figure 14: ROC analysis for miR-16 and determination of the diagnostic cut-off.

3.3 Threshold Value Analysis for let-7a

For the let-7a analysis, a ROC curve was generated with an AUC of 0.8615 (CI 0.7536 - 0.9695), p = 0.0005 (Figure 15). The Youden's Index was calculated for each cut-off value, with a maximum value of 0.6179. The optimal cut-off was determined to be 0.9775, providing a sensitivity of 71.79% and a specificity of 90.0%.

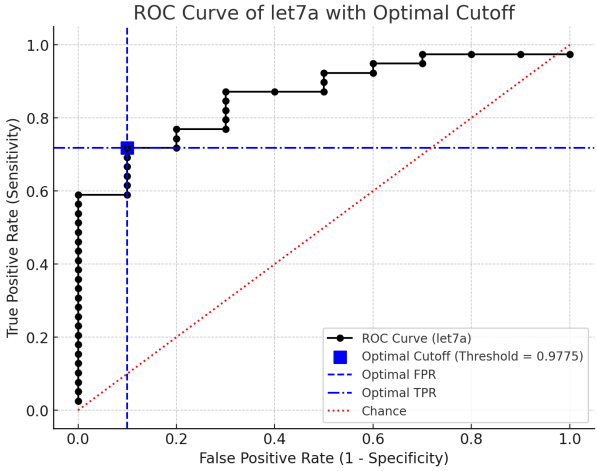


Figure 15: ROC analysis for let-7a and determination of the diagnostic cut-off.

3.4 Threshold Value Analysis for miR-451a

For miR-451a, which showed the most significant difference in expression between the two groups, a ROC curve was generated with an AUC of 0.9175 (CI 0.8354 - 0.996) and $p < 0.0001$ (Figure 16). The Youden's Index was calculated for each cut-off value, with the maximum index being 0.7. The optimal cut-off was determined at 1.161, where sensitivity was 70.0% and specificity was 100.0%. This cut-off ensures high specificity, which is crucial for minimizing false-positive results.

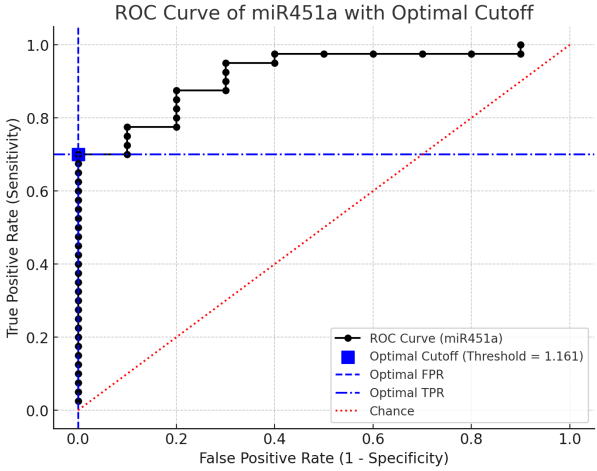


Figure 16: ROC analysis for miR-451a and determination of the diagnostic cut-off.

4. Comparative Analysis by Demographic Characteristics:

4.1 Comparative Analysis by Gender

A comparison of miRNA levels between men and women was conducted to determine whether there are gender-based differences in the studied miRNAs. Statistical analysis did not reveal any significant differences in the levels of any of the miRNAs between men and women (Table 6). The results of this analysis are presented in a table that reflects the mean values and standard deviations for each miRNA in both groups, as well as the corresponding p-values, confirming the absence of significant differences. These findings indicate that gender does not substantially influence the levels of the

studied miRNAs, allowing the data to be pooled for subsequent analyses.

MicroRNA	Gender	Sample size	Mean ± SD	p-value
MiR-22	Male	25	1,357 ± 1,299	0,6633
	Female	15	1,190 ± 0,8924	
MiR-144	Male	25	0,9850 ± 0,5125	0,9284
	Female	15	1,005 ± 0,8713	
MiR-16	Male	25	0,9548 ± 0,5356	0,7547
	Female	15	1,022 ± 0,7567	
Let7a	Male	25	0,8350 ± 0,4645	0,5886
	Female	15	0,9567 ± 0,9274	
MiR-451a	Male	25	0,9184 ± 0,5193	0,6201
	Female	15	1,022 ± 0,7898	

Table 6: Results of Comparative Analysis by Gender

4.2 Analysis of miRNA Levels in Relation to Age

To assess the potential relationship between miRNA levels and patient age, a correlation analysis was performed using Spearman's correlation coefficient. The results showed weak to moderate negative correlations between age and the levels of all examined miRNAs (Table 7). A statistically significant correlation was found only for **let-7a-5p** ($r=-0.4569$,

p=0.0035), suggesting that its levels decrease with increasing age.

For the remaining miRNAs, the correlations were not statistically significant (p>0.05), although the correlation coefficients indicated a tendency toward a negative relationship. This suggests that age is not a significant factor influencing the expression of these miRNAs in the studied groups.

MicroRNA	Correlation coefficient (ρ)	95% CI	P-value	Significance
MiR-22	-0.2894	-0.5578 до 0.03378	0.0701	Not significant
MiR-144	-0.2228	-0.5067 до 0.1048	0.167	Not significant
MiR-16	-0.2968	-0.5634 до 0.02575	0.0629	Not significant
Let-7a	-0.4569	-0.6803 до -0.1557	**0.0035**	Significant
MiR-451a	-0.2682	-0.5418 до 0.05677	0.0943	Not significant

Table 7: Results of Correlation Analysis Between miRNAs and Patient Age

5. Comparative Analysis Between miRNA Levels and Different Disease Characteristics

5.1 Analysis of miRNA Levels in Different MDS Subtypes

To ensure more reliable statistical processing, given the small overall number of study participants, MDS patients were

divided into two groups: those with MDS with multilineage dysplasia, ring sideroblasts, and del(5q), and those with MDS with excess blasts. The comparison of the studied miRNA levels between these MDS subtypes revealed that only let-7a exhibited a statistically significant difference between the two groups ($p=0.0250$), with a large effect size ($r=-0.653$), highlighting its potential role as a biomarker (Table 8).

For the remaining miRNAs—miR-22, miR-144, miR-16, and miR-451a—no statistically significant differences were identified ($p>0.05$). However, for miR-144 ($r=-0.319$) and miR-16 ($r=-0.308$), the effect size indicated a moderate association. This suggests that with a larger sample size, these miRNAs may demonstrate a significant difference between the subtypes. In contrast, miR-22 and miR-451a showed a small effect size ($r=-0.291$ and $r=-0.274$, respectively), indicating a weaker trend toward differentiation.

MicroRNA	Median (MLD, 5q, RARS)	Median (RAEB)	P-value	Effect size (r)	Interpretation
MiR-22	0.779	0.884	0.3307	-0.291	Small to moderate effect
MiR-144	0.775	0.873	0.2793	-0.319	Moderate effect
MiR-16	0.779	0.832	0.2998	-0.308	Moderate effect
Let-7a	0.579	0.851	0.025	-0.653	Large effect
MiR-451a	0.77	0.87	0.3602	-0.274	Small to moderate effect

Table 8: Results of the comparative analysis between the levels of the five microRNAs and different MDS subtypes.

5.2 Analysis of microRNAs according to patient risk stratification based on the R-IPSS scale

Patients were divided into two groups based on their risk stratification according to the R-IPSS scale. The low-risk group included patients with very low, low, and intermediate risk ≤ 3.5 points, while the high-risk group comprised patients with intermediate risk > 3.5 points, high, and very high risk according to R-IPSS.

The comparison of microRNA levels showed that let-7a and miR-451a exhibited statistically significant differences between the two groups (p=0.0172 and p=0.0449, respectively) (Table 9). The levels of these microRNAs were significantly lower in the low-risk group compared to the high-risk group, with large effect sizes (r=−0.753 for let-7a and r=−0.62 for miR-451a), emphasizing their clinical significance.

The other microRNAs (miR-22, miR-144, and miR-16) did not show statistically significant differences (p>0.05), but their effect sizes (r=−0.49, r=−0.51, r=−0.56, respectively) suggest a moderate to strong association. This indicates that with a larger sample size, these microRNAs may also demonstrate significant differences between the groups.

MicroRNA	Median (Low risk)	Median (High risk)	P-value	Effect size (r)	Interpretation
MiR-22	0.758	0.902	0.1132	-0.49	Moderate to large effect
MiR-144	0.731	0.923	0.0968	-0.51	Large effect
MiR-16	0.761	0.96	0.0696	-0.56	Large effect
Let-7a	0.532	0.785	0.0172	-0.753	Large effect
MiR-451a	0.714	0.894	0.0449	-0.62	Large effect

Table 9: Results of the comparative analysis of the five microRNAs and risk classification according to R-IPSS.

5.3 Analysis of microRNAs according to cytogenetic risk

This analysis included only patients with intermediate and good cytogenetic risk, as the number of patients with very good, poor, and very poor risk was below three, which did not allow for a reliable statistical evaluation.

Comparison of microRNA levels revealed statistically significant differences for miR-144 (p=0.0287) and miR-451a (p=0.0481), with a large effect size (r=−0.53) and a medium to large effect size (r=−0.48), respectively (Table 10). This highlights the potential clinical significance of these microRNAs in distinguishing cytogenetic risk groups.

For miR-16 (p=0.0679), the effect size (r=−0.45) was medium to large, suggesting that these differences could become statistically significant with a larger sample size. MiR-22 (r=−0.29) showed a small to moderate effect, whereas let-7a had no significant difference (p=0.8759, r=−0.046) and practically no clinical relevance.

MicroRNA	Median (Good risk)	Median (Intermediate risk)	P-value	Effect size (r)	Interpretation
MiR-22	0.916	0.804	0.2484	-0.29	Small to moderate effect
MiR-144	0.916	0.633	0.0287	-0.53	Large effect
MiR-16	0.9955	0.601	0.0679	-0.45	Moderate to large effect
Let-7a	0.712	0.785	0.8759	-0.046	Minimal effect
MiR-451a	0.8975	0.517	0.0481	-0.48	Moderate to large effect

Table 10: Results of the comparative analysis between the levels of the five microRNAs and cytogenetic risk.

6. Correlation analysis to examine the relationships between selected microRNAs and various hematological and biochemical parameters in patients with MDS.

To study the relationship between two variables that do not meet the requirements for normal distribution or are on ordinal scales, the most appropriate method is the Spearman correlation coefficient. The Spearman correlation coefficient measures the strength and direction of a monotonic relationship between two variables, regardless of whether the relationship is linear.

If the Spearman correlation coefficients are statistically significant at a significance level of $\alpha = 0.05$ with a two-tailed critical region, this indicates a statistically significant relationship between the studied variables.

To determine the strength of this relationship, the absolute values of statistically significant correlation coefficients are interpreted as follows:

- [0.1; 0.3) – weak correlation
- [0.3; 0.5) – moderate correlation
- [0.5; 0.7) – substantial correlation
- [0.7; 0.9) – strong correlation
- [0.9; 0.99) – very strong correlation
- [0.99; 1] – functional or nearly functional correlation

Spearman correlation is primarily used for ranked data or when variables are not normally distributed. In this context, it provides a more accurate assessment of the relationship between two variables, particularly for skewed distributions or data that are not strictly quantitative.

6.1 Spearman Correlation Analysis Between Selected microRNAs

A Spearman correlation analysis was conducted to examine the relationships between the selected microRNAs. The absolute values of the correlation coefficients indicate varying degrees of relationship strength between the variables. The correlations are presented in **Table 11**.

		MiR-22	MiR-144	MiR-16	Let-7a	MiR-451a
MiR-22	*r* *p*	1	0,7870 <0,0001	0,8336 <0,0001	0,6018 <0,0001	0,7553 <0,0001
MiR-144	*r* *p*	0,7870 <0,0001	1	0,9246 <0,0001	0,5203 0,0007	0,8860 <0,0001
MiR-16	*r* *p*	0,8336 <0,0001	0,9246 <0,0001	1	0,5771 <0,0001	0,9559 <0,0001
Let-7a	*r* *p*	0,6018 <0,0001	0,5203 0,0007	0,5771 0,0001	1	0,5389 0,0004
MiR-451	*r*	0,7553	0,8860	0,9559	0,5389	1

a	*p*	<0,0001	<0,0001	<0,0001	0,0004	

Table 11: Results of the Correlation Analysis Between the Levels of Individual microRNAs.

Results indicate that all microRNAs exhibit a statistically significant positive correlation. A significant correlation (with a correlation coefficient between 0.5 and 0.7) was observed between let-7a and the other four microRNAs: miR-22, miR-144, miR-16, and miR-451a. A strong correlation (with a correlation coefficient between 0.7 and 0.9) was identified between miR-22 and miR-144, miR-16 and miR-451a, as well as between miR-144 and miR-451a. A very strong correlation (with a correlation coefficient between 0.9 and 0.99) was established between miR-16 and miR-144, as well as between miR-16 and miR-451a. All correlations have positive values, indicating that an increase in the levels of one microRNA is associated with an increase in the levels of others. These relationships are visualized through scatter plots, allowing a clear view of the direction and strength of the correlations (Figure 17), while a schematic representation of the summarized results is presented in Table 12.

		MiR-22	MiR-144	MiR-16	Let-7a	MiR-451a
MiR-22	Correlation		**Strong**	**Strong**	**Significant**	**Strong**

	Significance		Significant	Significant	Significant	Significant
MiR-144	Correlation	Strong		Very strong	Significant	Strong
	Significance	Significant		Significant	Significant	Significant
MiR-16	Correlation	Strong	Very strong		Significant	Very strong
	Significance	Significant	Significant		Significant	Significant
Let-7a	Correlation	Significant	Significant	Significant		Significant
	Significance	Significant	Significant	Significant		Significant
MiR-451a	Correlation	Strong	Strong	Very strong	Significant	
	Significance	Significant	Significant	Significant	Significant	

Table 12: Schematic representation of the results from the correlation analysis between the levels of individual microRNAs.

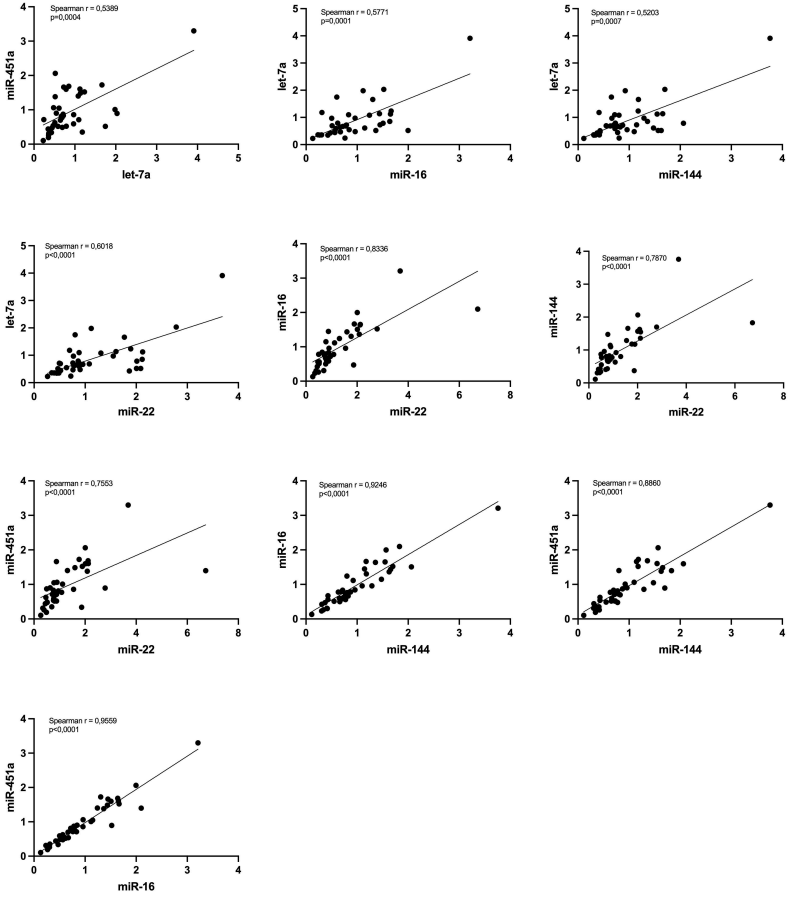


Figure 17: Scatter plots showing the results of the correlation analysis between the levels of individual microRNAs.

Simple linear regression was used to assess the relationship between the levels of selected microRNAs, aiming to determine whether there is a linear dependence between these variables. Table 13 presents the results of the analysis, where: R^2, represents the percentage of variation in the

dependent variable that can be explained by the independent variable (miRNA), b_0 (Intercept) - indicates the predicted value of the dependent variable when the independent variable (miRNA) is zero, b_1 (Regression coefficient) - shows how the dependent variable changes with each unit increase in the independent variable (miRNA), CI b_1 (95% Confidence Interval for Slope) - represents the range of values within which the true slope value is expected to lie with 95% confidence, p-value for b_1 - determines whether the regression coefficient b_1 is statistically significant, Standard Error (SE) of b_1 - evaluates the precision of the b_1 coefficient estimate.

Independent variable	Dependent variable	R^2	b1 (Slope)	b0 (Y-intercept)	CI b1	p	SE
MiR-22	miR-144	0,48	0,3958	0,4801	0,2605 - 0,5311	<0,0001	0,06684
MiR-22	miR-16	0,58	0,4096	0,4497	0,2957 - 0,5235	<0,0001	0,5625
MiR-22	let-7a	0,46	0,6046	0,1834	0,3875 - 0,8217	<0,0001	0,1071
MiR-22	miR-451a	0,36	0,3258	0,5353	0,1831 - 0,4685	<0,0001	0,07049
MiR-144	miR-16	0,87	0,8748	0,1117	0,7623 - 0,9872	<0,0001	0,05553
MiR-144	let-7a	0,49	0,7167	0,1859	0,4712 - 0,9623	<0,0001	0,1212
MiR-144	miR-451a	0,80	0,8496	0,1139	0,7096 - 0,9897	<0,0001	0,06917

MiR-16	let-7a	0,46	0,7611	0,1579	0,4872 - 1,035	<0,0001	0,1351
MiR-16	miR-451a	0,91	0,9675	0,009118	0,8699 - 1,065	<0,0001	0,04818
Let-7a	miR-451a	0,395	0,5901	0,4253	0,3470 - 0,8333	<0,0001	0,1200

Table 13: Results of Simple Linear Regression for Assessing the Relationship Between Individual microRNAs.

Based on the results of the conducted regression analysis for the relationship between the five microRNAs, it was established that a strong association (with $R^2 > 0.7$) exists for the pairs miR-144 with miR-16, miR-144 with miR-451a, and miR-16 with miR-451a. These results indicate that a significant portion of the variation in the levels of one microRNA can be explained by the variation in the levels of the other. A moderate association (with R^2 between 0.5 and 0.7) was found between miR-22 and miR-16, suggesting a moderate yet significant relationship between these variables. For the remaining pairs, although a statistically significant correlation was observed, the association was weak ($R^2 < 0.5$), which could be attributed to the influence of other factors not included in the regression model. Additionally, all 95% confidence intervals are narrow, as they are within 10% of the slope value, indicating high precision in the estimation of regression coefficients.

6.2 Correlation Analysis Between microRNAs and Laboratory Parameters

A correlation analysis was also conducted between each microRNA and specific laboratory parameters, namely:

hemoglobin, reticulocytes, leukocytes, neutrophils, platelets, LDH (lactate dehydrogenase), beta-2 microglobulin, ferritin, and erythropoietin levels (Table 14).

The analysis identified the following significant correlations:

- **miR-22** showed a statistically significant moderate correlation with LDH and erythropoietin levels.
- **miR-144** demonstrated a statistically significant moderate correlation with LDH and ferritin levels.
- **miR-16** exhibited a statistically significant moderate correlation with ferritin levels.
- **Let-7a** showed a statistically significant strong correlation with LDH.
- **miR-451a** demonstrated a statistically significant moderate correlation with ferritin.

No significant correlations were found between microRNAs and hemoglobin, reticulocyte count, leukocyte count, neutrophil count, platelet count, or beta-2 microglobulin levels.

Laboratory parameters	**Let7a**		**MiR-16**		**MiR-22**		**MiR-144**		**MiR-451a**	
	r	**p**	**r**	**p**	**r**	**p**	**r**	**p**	**r**	**p**
Hemoglobin (g/L)	0,2317	0,1558	-0,1152	0,4791	-0,1054	0,5174	0,01633	0,9203	-0,1480	0,3620
Reticulocytes (%)	0,2686	0,1307	-0,09811	0,5810	0,1357	0,4442	-0,04263	0,8108	-0,2022	0,2516
Leucocytes (x10^9/L)	0,1508	0,3594	0,05694	0,7271	0,09062	0,5781	0,1118	0,4921	0,1131	0,4870
Neutrophils (x10^9/L)	0,02915	0,8602	-0,02927	0,8577	-0,02608	0,8731	-0,001782	0,9913	0,03171	0,8460

Thrombocytes (x10^9/L)	-0,04333	0,7934	0,1803,	0,2655	0,2062	0,2017	0,2217	0,1692	0,1456	0,3700
LDH (U/L)	**0,5652**	**0,0002**	0,3114,	0,0505	**0,3970**	**0,0112**	**0,3704**	**0,0186**	0,2677	0,0950
Beta2-microglobulin (mg/L)	-0,1766	0,2820	-0,05319	0,7444	0,08063	0,6209	-0,09797	0,5475	-0,1467	0,3664
Erythropoietin (U/l)	0,2317	0,2180	0,3134	0,0861	**0,3592**	**0,0472**	0,3172	0,0821	0,3051	0,0952
Ferritin (µg/L)	0,1169	0,4909	**0,3374**	**0,0383**	0,2610	0,1135	**0,3282**	**0,0443**	**0,3730**	**0,0211**

Table 14: Results of the Correlation Analysis Between microRNA Levels and Selected Laboratory Parameters.

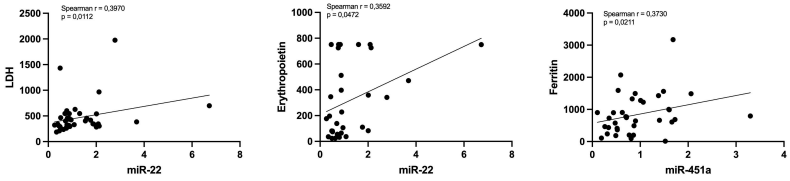
For a deeper understanding of these relationships, the analysis was extended through simple linear regression between the microRNAs and laboratory parameters for which a statistically significant correlation was established (Table 15). This approach allowed for the assessment of the degree of linear dependence and the determination of whether laboratory parameters could be predicted based on microRNA levels.

MiR	Laboratory parameters	R^2	b1	b0	CI b1	p	SE
MiR-22	LDH	0,08	80,82	362,6	-9,184 - 170,8	0,077	44,46
MiR-22	Erythropoietin	0,16	88,06	208,0	10,39 - 165,7	**0,0277**	37.98
MiR-144	LDH	0,05	112,8	355,2	-47,15 - 272,9	0,1615	79,04

MiR-144	Ferritin	0,06	220,7	622,9	-86,05 - 527,4	0,1532	151,2
MiR-16	Ferritin	0,08	283,6	565,6	-36,88 - 604	0,0811	158
Let-7a	LDH	0,1	153,0	326,3	-3,429 - 309,4	0,0550	77,19
MiR-451 a	Ferritin	0,09	287,9	568,7	-29,16 - 605	0,0738	156,3

Table 15: Results of simple linear regression assessing the relationship between microRNAs and specific laboratory parameters.

Based on the conducted regression analysis, no significant relationships were identified between laboratory parameters and microRNAs. Despite the statistically significant correlations, the regression models indicate that laboratory parameters cannot be reliably predicted based on microRNA levels. This could be due to the influence of other factors not included in the analysis, which likely have a stronger impact on these laboratory parameters. The data from the correlation and regression analysis are graphically represented using scatter plots, allowing for visualization of the relationships between variables (Figure 18).



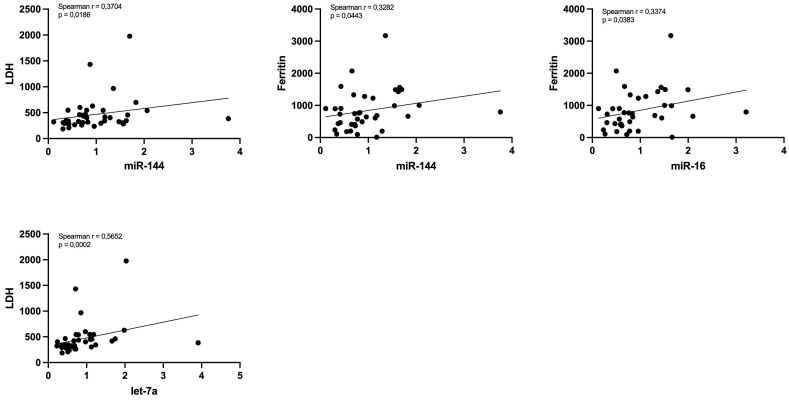


Figure 18: Scatter plots showing the results of the correlation analysis between microRNA levels and specific laboratory parameters.

6.3 Correlation Analysis Between MicroRNA Levels and Bone Marrow Blast Percentage

Additionally, a correlation analysis was conducted between microRNA levels and the percentage of blasts in the bone marrow (Table 16).

	Bone marrow blast percentage (%)	
	Spearman r	**p value**
MiR-22	0,1284	0,4298
MiR-144	0,1179	0,4687
MiR-16	0,1643	0,3109

Let-7a	**0,3622**	**0,0235**
MiR-451a	0,1173	0,4709

Table 16: Results of Correlation Analysis Between MicroRNA Levels and Bone Marrow Blast Percentage.

The analysis identifies a statistically significant, positive, moderate correlation between let-7a levels and the percentage of blasts in the bone marrow (Figure 19). No correlation is observed between the other miRNAs and the percentage of bone marrow blasts. Based on this finding, a simple linear regression was performed, but it did not establish a significant predictive relationship, as the R^2 value was 0.17. The confidence interval for the slope coefficient ranged from 1.059 to 6.902, and the p-value was 0.0089, indicating statistical significance but low predictive value of the model. The data from this analysis are graphically presented using a scatter plot to visualize the relationship.

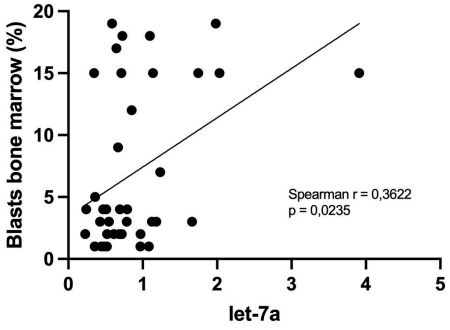


Figure 19: Graphical representation of the correlation analysis results between let-7a levels and the percentage of bone marrow blasts.

7. Evaluation of the Diagnostic Value of Selected miRNAs

The next part of the analysis focuses on evaluating the predictive diagnostic value of the four microRNAs that showed a statistically significant difference between MDS patients and healthy controls. To achieve this, a simple logistic regression analysis was conducted to assess the probability of these microRNAs serving as predictors for distinguishing between the two groups (Table 17). The following parameters were calculated:

- **β0 (Intercept):** This is the intercept of the logistic model, representing the log-odds of belonging to the patient group when the microRNA level is zero.
- **β1 (Slope):** The slope coefficient indicates the change in log-odds when the microRNA level changes by one unit.
- **Odds Ratio (OR):** This is the exponential value of β1, representing the odds ratio for belonging to the patient group. An OR < 1 suggests that an increase in the microRNA level decreases the odds of belonging to the patient group.
- **95% CI (OR):** This is the confidence interval for the odds ratio. If the interval does not include 1, the relationship is statistically significant.
- **P-value (β_1):** This is the p-value for β1, indicating whether the slope is statistically significant. A p-value < 0.05 suggests a significant association.

- **Pseudo-R^2:** This is the equivalent of the coefficient of determination (R^2) for logistic regression and indicates the proportion of variance in the dependent variable explained by the model.
- **Deviance:** Represents the deviation of the model from an ideal fit. Lower deviance values indicate a better model fit.
- **Chi-Square (χ^2):** This is a statistical test for model significance. Higher values and a p-value < 0.05 suggest a significant association.

MiR	β_0	β_1	OR	95% CI (OR)	P-value (β_1)	Pseudo-R^2	Deviance	Chi-Square (χ^2)	P-value (χ^2)
MiR-16	4.016 (2.177 до 6.670)	-1.941 (-3.589 до -0.7598)	0.1436	[0.02763, 0.4677]	0.0055	0.2428	38.16	11.88	0.0006
MiR-451a	4.999 (2.826 до 8.396)	-2.411 (-4.330 до -1.145)	0.08977	[0.01317, 0.3182]	0.0022	0.4193	30.57	19.47	<0.0001
MiR-144	3.548 (1.962 до 5.770)	-1.555 (-2.897 до -0.5846)	0.2111	[0.05520, 0.5573]	0.0074	0.2552	38.67	11.37	0.0007
Let-7a	3.095 (1.68	-1.430	0.2393	[0.07049,	0.0092	0.2048	40.64	8.953	0.0028

	1 до 4.946)	(-2.652 до -0.4638)		0.6289]					

Table 17: Results of the logistic regression analysis assessing the potential of selected microRNAs to distinguish MDS patients from healthy controls.

7.1 Overview of Results for Individual microRNAs

For miR-16, the slope coefficient (β_1) is −1.941, leading to an Odds Ratio (OR) of 0.1436 (95% CI: [0.02763, 0.4677]), indicating that higher levels of this microRNA significantly reduce the likelihood of disease. The p-value for β_1 is 0.0055, confirming a statistically significant relationship.

For miR-451a, the coefficient β_1 is −2.411, and the OR is 0.08977 (95% CI: [0.01317, 0.3182]), also suggesting a strong reduction in the likelihood of being in the patient group with higher levels of this microRNA. The p-value is 0.0022, confirming the significance of this relationship.

Regarding miR-144, with $\beta_1 = -1.555$ and an OR of 0.2111 (95% CI: [0.05520, 0.5573]), it was found that higher levels of this microRNA also reduce the probability of belonging to the patient group. The p-value of 0.0074 indicates a statistically significant association.

Lastly, for let-7a, the coefficient β_1 is −1.430, with an OR of 0.2393 (95% CI: [0.07049, 0.6289]), also demonstrating a significant relationship with a p-value of 0.0092.

In addition to analyzing the regression coefficients, the Pseudo-R^2 values were examined to assess how well the model explains the variation in the dependent variable for each microRNA. The highest Pseudo-R^2 was found for miR-451a (0.4193), indicating that approximately 42% of the variation in the diagnostic model can be explained by this microRNA, confirming its high predictive value. For miR-16, a Pseudo-R^2 of 0.2428 was calculated, explaining around 24% of the model's variation. Although lower than miR-451a, it still suggests moderate predictive capability. MiR-144 and let-7a showed Pseudo-R^2 values of 0.2552 and 0.2048, respectively, meaning around 25% and 20% of the diagnostic model's variation can be explained by these microRNAs. While statistically significant, their predictive values are lower compared to miR-451a.

For all models, the Chi-square test was used to evaluate the statistical significance of the model. For miR-451a, the Chi-square value is 19.47 with $p < 0.0001$, indicating an extremely significant relationship. MiR-16 and miR-144 also showed significant results, with Chi-square values of 11.88 ($p = 0.0006$) and 11.37 ($p = 0.0007$), respectively, confirming the significance of these models. For let-7a, the Chi-square value is 8.953 with $p = 0.0028$, which also indicates a significant relationship but with a lower predictive value compared to the other microRNAs.

Summary of Results

The logistic regression analysis evaluated the relationship between four microRNAs and the likelihood of belonging to the MDS patient group. The analysis demonstrated that all four microRNAs (miR-16, miR-451a, miR-144, and let-7a) have

statistically significant associations. Higher levels of each microRNA were associated with a reduced likelihood of disease.

The Pseudo-R^2 values indicate that the models explain a moderate to substantial portion of data variation, with miR-451a showing the highest predictive value.

The Chi-square test further confirmed the significance of all models, with miR-451a emerging as the strongest predictor, providing the best predictive value among the examined microRNAs.

7.2 Evaluation of the Predictive Value of microRNAs through Regression Analysis

To determine the predictive ability of the five studied microRNAs (miR-22, miR-144, miR-16, let-7a, miR-451a), regression analysis was performed. Due to the observed multicollinearity (high correlation between some microRNAs), which could affect the stability and interpretation of the coefficients in the model, a LASSO (Least Absolute Shrinkage and Selection Operator) analysis was conducted. This method allows the identification of the most significant predictors by reducing the coefficients of less relevant microRNAs to zero and evaluating their combined effect on the probability of belonging to the patient group.

The regression coefficients (β) were analyzed to determine the strength and direction of the relationship between each microRNA and the likelihood of belonging to the patient group. Additionally, Odds Ratios (ORs) were calculated, which indicate how many times the odds of belonging to the patient group change with variations in microRNA levels. Model

performance indicators were also assessed, including AUC (Area Under the Curve) and AIC (Akaike Information Criterion), where lower AIC values indicate a better model fit.

The results showed that miR-22 and miR-451a had the most significant regression coefficients (β), at 0.260 (95% CI: [0.095, 0.419]) and -0.362 (95% CI: [-0.889, 0.053]), respectively (Table 18). This indicates that higher levels of miR-22 were associated with increased odds of belonging to the patient group (OR = 1.297, 95% CI: [1.100, 1.521]), whereas higher levels of miR-451a were associated with reduced odds of being in the patient group (OR = 0.696, 95% CI: [0.411, 1.054]).

The other microRNAs showed a weaker association with patient group classification. MiR-144 had a coefficient β = 0.015 (95% CI: [-0.225, 0.257]) and OR = 1.015 (95% CI: [0.799, 1.293]), indicating minimal influence. MiR-16 had a coefficient β = −0.010 (95% CI: [-0.534, 0.575]) and OR = 0.990 (95% CI: [0.586, 1.777]), suggesting a slight decrease in disease likelihood with higher microRNA levels. Similarly, let-7a had a negative coefficient β = −0.070 (95% CI: [-0.216, 0.039]) and OR = 0.932 (95% CI: [0.806, 1.040]), suggesting a slight reduction in disease risk with higher microRNA levels.

The model performance indicators confirmed its effectiveness: AUC is 0.994 (95% CI: [0.973, 1.0]), indicating excellent discriminatory ability, AIC is 31.109, suggesting a strong model fit and the overall test accuracy is 91.84%.

Variable	Coefficient β (95% CI)	Odds Ratio (95% CI)	
Intercept	0.796 (0.725, 0.866)	2.216 (2.065, 2.377)	AUC: 0.994, AIC: 31.109, 91.84%
MiR-22	0.260 (0.095, 0.419)	1.297 (1.100, 1.521)	
MiR-144	0.015 (-0.225, 0.257)	1.015 (0.799, 1.293)	
MiR-16	-0.010 (-0.534, 0.575)	0.990 (0.586, 1.777)	
Let-7a	-0.070 (-0.216, 0.039)	0.932 (0.806, 1.040)	
MiR-451a	-0.362 (-0.889, 0.053)	0.696 (0.411, 1.054)	

Table 21: Results of LASSO analysis for determining the predictive value of the five studied microRNAs.

Based on the results of the initial LASSO analysis, a decision was made to conduct a simplified analysis focusing only on the two microRNAs with the greatest diagnostic significance—miR-22 and miR-451a. These microRNAs were selected due to their significant coefficients and strong associations with the probability of belonging to the patient group, as demonstrated in the initial model. Conducting the analysis with these two microRNAs allows for the assessment of their combined effect and predictive value in distinguishing patients with MDS from the control group.

In the simplified LASSO analysis, it was determined that both miRNAs remain significant predictors in the model. The

coefficient β for miR-22 is 0.228 (95% CI: [0.095, 0.378]), suggesting that higher levels of this microRNA are associated with increased odds of belonging to the patient group (Table 19). This is confirmed by the OR, which is 1.256 (95% CI: [1.100, 1.460]), indicating a higher risk of disease at elevated levels of miR-22.

On the other hand, miR-451a has a negative coefficient β of -0.378 (95% CI: [-0.509, -0.256]), suggesting that higher levels of this microRNA reduce the likelihood of belonging to the patient group. The OR for this microRNA is 0.685 (95% CI: [0.601, 0.774]), indicating a decreased risk of disease with higher miR-451a levels.

The model's performance indicators remain high: the AUC is 0.994 (95% CI: [0.974, 1.0]), demonstrating the model's excellent ability to distinguish patients from control subjects, while the AIC decreases to 24.916, with the overall accuracy of the model increasing to 95.92%.

Variable	Coefficient β (95% CI)	Odds Ratio (95% CI)	Model (AUC/AIC/ Accuracy)
Intercept	0.799 (0.710, 0.876)	2.224 (2.034, 2.402)	AUC: 0.994, AIC: 24.916, 95.92%
MiR-22	0.228 (0.095, 0.378)	1.256 (1.100, 1.460)	
MiR-451a	-0.378 (-0.509, -0.256)	0.685 (0.601, 0.774)	

Table 19: Results from the simplified LASSO analysis including miR-22 and miR-451a

To confirm the results of the simplified LASSO analysis, a multivariable logistic regression analysis was performed, with the two most significant microRNAs—miR-22 and miR-451a.

Additionally, a multicollinearity assessment was conducted by calculating the Variance Inflation Factor (VIF) to ensure that there was no significant overlap between the two microRNAs in the model, which could affect the stability and interpretation of the results.

The results of the multivariable logistic regression analysis, incorporating miR-22 and miR-451a, confirmed their significant joint predictive effect. The regression coefficient β for miR-22 was 10.31 (95% CI: [3.180, 35.99]), indicating a strong positive association with the probability of belonging to the patient group (Table 20). On the other hand, miR-451a had a negative coefficient β of -8.912 (95% CI: [-22.32, -3.946]), suggesting that higher levels of this microRNA significantly reduce the likelihood of belonging to the patient group.

The model's performance evaluation demonstrated excellent discriminatory ability, with an AUC of 0.995 (95% CI: [0.982, 1.000]), indicating high accuracy in distinguishing patients from healthy controls. The AIC was further reduced to 13.73, and the overall accuracy of the model reached 96%, confirming its strong fit.

In the multicollinearity analysis, the calculated VIF values for miR-22 and miR-451a were close to 1 (1.183), indicating no significant multicollinearity between the two microRNAs and confirming the stability of the model.

Variable	Coefficient β (95% CI)	Odds Ratio (95% CI)	Model (AUC/AIC/ Accuracy)	VIF
Intercept	4.397 (-3.567, 12.361)	81.20 (0.09993, 4.77e+06)	AUC: 0.995, AIC: 13.73, 96.00%	—
MiR-22	10.31 (3.180, 35.99)	29,934 (24.04, 4.28e+15)		1.183
MiR-451a	-8.912 (-22.32, -3.946)	0.0001347 (2.03e-10, 0.01933)		1.183

Table 23: Results of the multivariable logistic regression analysis for determining the predictive value of miR-22 and miR-451a in distinguishing patients from healthy control**s.**

In conclusion, the conducted analyses highlight the significant role of miR-22 and miR-451a as predictors of belonging to the group of patients with myelodysplastic syndrome. The initial LASSO analysis, which included all microRNAs, identified these two as the most significant, while subsequent simplified and multivariable analyses confirmed their strong predictive value. MiR-22 shows a positive association with disease probability, whereas miR-451a has an inverse relationship, suggesting a protective effect. The models demonstrate excellent accuracy and stability, supporting the diagnostic significance of these two microRNAs.

8. Evaluation of the Predictive Role of microRNAs in Differentiating High- and Low-Risk MDS According to R-IPSS

8.1 LASSO Analysis of the Five microRNAs

The next part of the analysis focuses on the predictive ability of microRNAs to distinguish between high-risk and low-risk MDS according to R-IPSS. The initial analysis indicated that only let-7a and miR-451a showed a significant difference between the two groups. To determine whether other microRNAs also hold predictive value for this distinction, a LASSO analysis was conducted. This method allows for the identification of additional microRNAs that may play a significant role in risk stratification in MDS.

The LASSO analysis results revealed different regression coefficients (β) and corresponding ORs for each microRNA. The coefficient for let-7a was 1.275 (95% CI: [0.000, 5.617]), with a corresponding Odds Ratio (OR) of 3.579 (95% CI: [1.000, 57.461]), suggesting that higher levels of let-7a increase the likelihood of belonging to the high-risk group, implying an association between elevated let-7a levels and higher-risk MDS (Table 21). The remaining microRNAs (miR-22, miR-144, miR-16, and miR-451a) showed varying trends in their coefficients and ORs, but all had wide confidence intervals, indicating uncertainty in their predictive roles.

The model's performance indicators demonstrated moderate discriminative ability, with an AUC of 0.712 (95% CI: [0.500, 0.931]) and an overall accuracy of 64.1%. The AIC of the model was 55.841, suggesting moderate model fit. The

model's intercept was estimated with a coefficient of 0.617 (95% CI: [-0.211, 2.344]) and a corresponding OR of 1.854 (95% CI: [0.810, 10.421]).

Variable	Coefficient β (95% CI)	Odds Ratio (95% CI)	
MiR-22	-0.278 [-2.562, 1.331]	0.757 [0.112, 5.544]	
MiR-144	0 [-2.954, 3.692]	1 [0.055, 16.924]	
MiR-16	0 [-11.963, 3.109]	1 [0.0002, 72.111]	AUC: 0.712 (95% CI: [0.500, 0.931]), Accuracy: 64.1%, AIC: 55.841
Let-7a	1.275 [0.000, 5.617]	3.579 [1.000, 57.461]	
MiR-451a	0.237 [-1.571, 8.580]	1.267 [0.174, 1124.977]	
Intercept	0.617 [-0.211, 2.344]	1.854 [0.810, 10.421]	

Table 21: Results of LASSO Analysis for Determining the Predictive Ability of the Five microRNAs in Differentiating High- and Low-Risk MDS.

8.2 Univariate Logistic Regression Analysis of let-7a

Since the results from the LASSO analysis indicated that only let-7a has a potential predictive role in distinguishing between high- and low-risk MDS, a simple logistic regression analysis was conducted. The objective of this analysis was to assess the independent effect of let-7a on the probability of belonging to either the high-risk or low-risk group and to determine its significance and predictive value in this context.

The results demonstrate a significant effect of this microRNA on the model (Table 22). The slope coefficient β_1 for let-7a was

calculated as 2.154 (95% CI: [0.4085, 4.636]), with a p-value of 0.0447, indicating a statistically significant association. The standard error (SE) was 1.073, and the corresponding Odds Ratio (OR) was 8.620 (95% CI: [1.505, 103.1]), suggesting that higher levels of let-7a increase the likelihood of belonging to the high-risk group.

The model's performance evaluation showed an AUC of 0.7065 (95% CI: [0.5436, 0.8694]), with a p-value of 0.03, indicating a moderate ability of the model to differentiate between high- and low-risk MDS. The Pseudo-R^2 was 0.1441, suggesting that 14.41% of the variation in the dependent variable could be explained by the model. The Chi-squared (χ^2) value was 6.702, with a p-value of 0.0096, confirming the model's significance.

Variable	β (95% CI)	SE	OR (95% CI)	p	χ^2	p (χ^2)	AUC (95% CI)	p (AUC)	R^2	Dev
Intercept (β0)	-1.292 (-3.083, 0.1826)	0.8227	0.2747 (0.04584, 1.200)	—	—	—	—	—	—	—
Let-7a (β1)	2.154 (0.4085, 4.636)	1.073	8.620 (1.505, 103.1)	0.0447	6.702	0.0096	0.7065 (0.5436, 0.8694)	0.03	0.1441	46.1

Table 22: Results from simple logistic regression analysis of let-7a for differentiating high- and low-Risk MDS.

The simple logistic regression analysis confirms that let-7a has a significant predictive role in differentiating between high- and low-risk MDS, with higher levels associated with an increased likelihood of belonging to the high-risk group. The model demonstrates moderate discriminatory ability and a significant

association, supporting the potential predictive value of let-7a in MDS risk stratification.

9. Evaluation of the Predictive Value of microRNAs for differentiating Different Types of MDS

9.1 LASSO Analysis of the Five microRNAs

The next part of the analysis focuses on determining the predictive value of microRNAs regarding different types of MDS. Due to the limited number of patients, they were divided into two groups—patients with MDS with a low percentage of blasts (including cases of MDS with multilineage dysplasia, MDS with ring sideroblasts, and MDS del(5q)) and patients with MDS with excess blasts (RAEB I and II). Since only one microRNA showed different levels between these two groups, an initial LASSO analysis was performed on all studied microRNAs to evaluate their predictive role in differentiating between different MDS types.

The results of the LASSO analysis indicate that, among the five examined microRNAs, only let-7a has a significant predictive role (Table 23). The regression coefficient (β) for let-7a is 1.547 (95% CI: [0.031, 3.064]), with an Odds Ratio (OR) of 4.70 (95% CI: [1.03, 21.41]). This suggests that higher levels of let-7a are associated with increased odds of belonging to the high-blast percentage group.

The other microRNAs (miR-22, miR-144, miR-16, and miR-451a) show weaker associations and wide confidence intervals for their coefficients and OR, indicating a lack of significant predictive value.

Model performance was assessed using AUC, which is 0.74 (95% CI: [0.56, 0.90]), indicating good discriminatory ability. The AIC value is 52.14, and the overall model accuracy is 74%, suggesting moderate effectiveness in distinguishing between patients with low and high percentages of blasts.

Variable	Coefficient β (95% CI)	Odds Ratio (95% CI)	
MiR-22	-1.526 [-3.715, 0.663]	0.22 [0.02, 1.94]	
MiR-144	0.235 [-2.107, 2.578]	1.27 [0.12, 13.17]	
MiR-16	2.348 [-2.284, 6.981]	10.47 [0.10, 1075.80]	AUC: 0.74 (95% CI: [0.56, 0.90]), AIC: 52.14, Accuracy: 74%
Let-7a	1.547 [0.031, 3.064]	4.7 [1.03, 21.41]	
MiR-451a	-1.419 [-4.741, 1.902]	0.24 [0.01, 6.70]	
Intercept	-0.736 [-1.509, 0.037]	0.48 [0.22, 1.04]	

Table 23: Results of the LASSO analysis for determining the predictive ability of the five microRNAs in distinguishing between MDS with a low percentage of blasts and MDS with a high percentage of blasts.

9.2 Simple Logistic Regression Analysis of let-7a

A simple logistic regression analysis was conducted, focusing solely on let-7a, to assess its independent predictive role in distinguishing patients with MDS with a low versus high

percentage of blasts. This analysis aimed to provide a more precise evaluation of its effect on the likelihood of belonging to one of the two groups.

The analysis revealed a statistically significant association between let-7a levels and the MDS subtype (Table 24). The slope coefficient (β1) for let-7a was estimated at 1.782 (95% CI: [0.3984, 3.681]), with a standard error of 0.8356 and a p-value of 0.0329. The corresponding Odds Ratio (OR) was 5.943 (95% CI: [1.489, 39.69]), indicating an increased risk of MDS with a high percentage of blasts at higher levels of let-7a.

The model's performance showed an AUC of 0.7219 (95% CI: [0.5488, 0.8950]), with a p-value of 0.0255, suggesting a moderate ability to differentiate between the groups. The Chi-squared (G squared) test yielded a value of 7.176 with a p-value of 0.0074, confirming the model's significance. The Pseudo-R^2 was 0.1681, indicating that approximately 16.81% of the variation in the dependent variable could be explained by the model, while the model's deviance (Dev) was 42.47.

Variable	β (95% CI)	SE	OR (95% CI)	p	χ^2	p (χ^2)	AUC (95% CI)	p (AUC)	R^2	Dev
Intercept (β0)	-2.275 (-4.074, -0.8609)	0.8115	0.1028 (0.01700, 0.4228)	—	—	—	—	—	—	—
Let-7a (β1)	1.782 (0.3984, 3.681)	0.8356	5.943 (1.489, 39.69)	0.0329	7.176	0.0074	0.7219 (0.5436, 0.8950)	0.0255	0.1681	42.47

Table 24: Results of simple logistic regression analysis of let-7a for differentiation between MDS with low and high blast percentage.

In summary, the analysis demonstrates that let-7a has a significant predictive role in distinguishing between different types of MDS. The LASSO analysis identified let-7a as the strongest predictor among the investigated microRNAs, while the simple logistic regression analysis confirmed this association, showing that higher levels of let-7a are linked to an increased likelihood of belonging to the high-blast MDS group. The model exhibits moderate discriminatory ability and significant predictive value.

10. Evaluation of the Prognostic Value of the Selected Plasma microRNAs

The next part of the analysis aims to determine the prognostic value of each examined microRNA concerning the survival of patients with MDS. To achieve this, an individual analysis approach was chosen, where each microRNA is evaluated separately using the Cox proportional hazards model. This model allows the assessment of how microRNA levels influence the time to event (death), accounting for data censoring.

Cox regression provides hazard ratios (HR), which indicate the relative risk associated with changes in microRNA levels. An HR >1 suggests an increased risk, HR <1 indicates a reduced risk, and p-values determine the statistical significance of the relationship. This approach enables a more in-depth investigation of the individual effect of each microRNA, which

is crucial for evaluating their potential as prognostic biomarkers.

10.1 Cox Proportional Hazards Analysis for miR-22

The analysis of the relationship between miR-22 levels and patient survival using Cox regression revealed that the coefficient (β) is 0.1447, with a standard error of 0.1344 and a 95% confidence interval ranging from −0.1643 to 0.3782.

The calculated HR is 1.156 (95% CI: 0.8485–1.460), indicating a slightly increased risk of an event with higher miR-22 levels. However, this effect is not statistically significant.

The AIC value for the model including miR-22 is 108.5, which is very close to the value for the null model (107.5), suggesting that the inclusion of miR-22 does not significantly improve the prognostic value of the model.

10.2 Cox Proportional Hazards Analysis for miR-144

The analysis of the relationship between miR-144 levels and patient survival using Cox regression revealed that the coefficient (β) is 0.4848, with a standard error of 0.2934 and a 95% confidence interval ranging from −0.1714 to 1.003.

The calculated HR is 1.624 (95% CI: 0.8425–2.726), suggesting a trend toward an increased risk of an event with higher miR-144 levels. However, this effect remains statistically non-significant.

The AIC value for the model including miR-144 is 107.3, which is slightly lower than the value for the null model (107.5). This suggests that including miR-144 may slightly improve the

model, although it does not demonstrate significant prognostic value.

10.3 Cox Proportional Hazards Analysis for miR-16

The Cox regression analysis for miR-16 revealed a coefficient (β) of 0.5192, with a standard error of 0.3302 and a 95% confidence interval ranging from −0.1817 to 1.126.

The calculated HR is 1.681 (95% CI: 0.8338–3.083), but since the confidence interval includes 1, the effect is not statistically significant.

However, the AIC value for the model including miR-16 is 107.3, which is slightly lower than the AIC of the null model (107.5). This suggests that adding miR-16 to the model provides some prognostic information, although its predictive value remains limited.

10.4 Cox Proportional Hazards Analysis for let-7a

The Cox regression analysis assessing the relationship between let-7a levels and survival revealed a coefficient (β) of 0.3197, with a standard error of 0.3015 and a 95% confidence interval ranging from −0.3778 to 0.8346.

The calculated HR is 1.377 (95% CI: 0.6854–2.304), suggesting a slightly increased risk of an event with higher let-7a levels. However, the effect is not statistically significant.

Furthermore, the AIC value for the model including let-7a is 103.8, which is slightly higher than the AIC of the null model (102.8). This indicates that adding let-7a does not improve the prognostic model.

10.5 Cox Proportional Hazards Analysis for miR-451a

The Cox regression analysis assessing the relationship between miR-451a levels and patient survival revealed a coefficient (β) of 0.5012, with a standard error of 0.3308 and a 95% confidence interval ranging from −0.2075 to 1.103.

The calculated HR is 1.651 (95% CI: 0.8126–3.012), indicating a potentially increased risk of an event with higher miR-451a levels, but the effect is not statistically significant as the confidence interval includes 1.

Furthermore, the AIC value for the model including miR-451a is 107.5, which is identical to the AIC of the null model. This suggests that adding miR-451a does not improve the prognostic value of the model.

The summary of all conducted analyses is presented in Table 25.

MiR	β	SE	HR	95% CI (HR)	AIC
MiR-22	0.1447	0.1344	1.156	0.8485 – 1.460	108.5
MiR-144	0.4848	0.2934	1.624	0.8425 – 2.726	107.3
MiR-16	0.5192	0.3302	1.681	0.8338 – 3.083	107.3
Let-7a	0.3197	0.3015	1.377	0.6854 – 2.304	103.8
MiR-451a	0.5012	0.3308	1.651	0.8126 – 3.012	107.5

Table 25: Results of Cox proportional hazards analysis for the five microRNAs to determine their prognostic value.

10.6 Cox Proportional Analysis for miR-144 and miR-16

Based on the obtained results, a combined model including miR-144 and miR-16 was developed to assess their joint prognostic value for overall survival in patients with MDS (Table 26). These two microRNAs were selected based on their individual results from the Cox regression analysis, which showed a trend toward increased risk with higher expression levels (HR of 1.624 for miR-144 and 1.681 for miR-16, respectively). Additionally, both microRNAs demonstrated lower AIC values (107.3) compared to the null model, suggesting that they may contribute prognostic value to the analysis. The combined model allowed for an evaluation of their independence and synergistic effect on survival outcomes.

The analysis of the combined model with miR-144 and miR-16 yielded the following results. The regression coefficient (β) for miR-144 was 0.2795 with a standard error of 0.8734 and a 95% confidence interval ranging from −1.453 to 2.034. For miR-16, the regression coefficient (β) was 0.2370 with a standard error of 0.9469 and a 95% confidence interval from −1.714 to 2.039. The calculated HR were 1.322 for miR-144 and 1.267 for miR-16, with confidence intervals encompassing 1 (0.2338 – 7.646 for miR-144 and 0.1802 – 7.687 for miR-16), indicating a lack of statistical significance.

The AIC value for the model including both microRNAs was 109.2, which was higher than the AIC for the null model (107.5), suggesting that the inclusion of these two microRNAs does not improve the prognostic value of the model.

MiR	β	Standard error	HR (Hazard Ratio)	95% CI за HR	AIC
MiR-144	0.2795	0.8734	1.322	0.2338 – 7.646	109.2
MiR-16	0.237	0.9469	1.267	0.1802 – 7.687	109.2

Table 26: Results from the Cox Proportional Hazards Model for miR-144 and miR-16.

These results indicate that despite the individual potential of miR-144 and miR-16, their combination in a single model does not add significant prognostic value to the analysis.

10.7 Summary of the Analysis on the Prognostic Value of microRNAs

The analysis of the prognostic value of the studied microRNAs was conducted using Cox regression, evaluating each microRNA individually, followed by a combined model for miR-144 and miR-16. The results from the individual analysis showed that all examined microRNAs, including miR-22, miR-144, miR-16, let-7a, and miR-451a, exhibited hazard ratios (HR) above 1, suggesting a tendency toward an increased risk of events with higher expression levels. However, the confidence intervals for HR included 1 for all microRNAs, indicating a lack of statistical significance.

The combined analysis of miR-144 and miR-16 also did not demonstrate significant prognostic value, as the AIC value for the model (109.2) was higher than that for the null model (107.5), suggesting that adding these microRNAs did not improve the predictive power of the model. These results

emphasize that despite the identified trends, the individual and combined models do not provide sufficient evidence for a significant prognostic effect of the studied microRNAs in the current cohort.

The lack of significance in this analysis may be attributed to the small sample size, which limits statistical power. However, the observed trends for certain microRNAs, such as miR-144 and miR-16, suggest that they may hold prognostic value in a larger cohort or in combination with other clinical markers. Further studies with a larger sample size and multivariable models are necessary to assess their potential role as prognostic biomarkers in patients with MDS.

V. Discussion

1. Introduction

Myelodysplastic syndromes represent a heterogeneous group of clonal disorders characterized by ineffective hematopoiesis, morphological dysplasia in the bone marrow, and an increased risk of progression to acute myeloid leukemia (AML). The severity of the disease varies significantly among patients, ranging from slow-progressing forms to aggressive variants with short survival and a high risk of transformation to AML. This heterogeneity highlights the need for a better understanding of the disease pathogenesis and the identification of reliable diagnostic and prognostic biomarkers.

The pathogenesis of MDS is complex, involving both genetic and epigenetic mechanisms. Genetic alterations, such as chromosomal aberrations and somatic mutations, are common and lead to the disruption of key cellular processes, including cell proliferation, apoptosis, RNA splicing, and signal transduction. In addition to genetic changes, epigenetic mechanisms such as aberrant DNA methylation, histone modifications, and dysregulation of non-coding RNAs play a crucial role in disease development. Among epigenetic regulators, microRNAs hold a significant place in the pathogenesis of MDS. These small non-coding RNA molecules regulate gene expression by suppressing translation or degrading target mRNAs. Altered miRNA expression is associated with various aspects of MDS, including impaired hematopoiesis, enhanced cell proliferation, and suppressed apoptosis.

The present study focuses on the role of five microRNAs—miR-22, miR-144, miR-16, miR-451a, and let-7a—in the context of MDS. By analyzing their expression levels and correlation with clinical and laboratory characteristics of patients, the study aims to provide a better understanding of their significance as diagnostic and prognostic biomarkers.

2. Analysis of microRNA Expression Levels in Different Groups of MDS Patients

2.1 MicroRNA-22

MiR-22 plays a key role in regulating TET2, a tumor suppressor gene that catalyzes the conversion of 5-methylcytosine to 5-hydroxymethylcytosine, an essential step in the DNA demethylation process. Reduced TET2 activity, caused by its negative regulation by miR-22, leads to genome hypermethylation, impaired differentiation of hematopoietic stem cells (HSCs), and the development of MDS. MiR-22 is thought to play a crucial role in hematopoiesis by promoting self-renewal of stem cells and suppressing their differentiation.

The findings of Song et al. demonstrate that high miR-22 levels in HSCs result in a significant reduction in TET2 expression and global 5-hydroxymethylcytosine levels, accompanied by an increase in 5-methylcytosine levels. This leads to an enhanced self-renewal capacity and suppressed differentiation of stem cells, predisposing to MDS and leukemia.

In our study, no statistically significant differences in miR-22 expression levels were found between MDS patients and healthy controls. No differences were observed when comparing low-risk and high-risk patients based on the R-IPSS or among the different MDS subtypes. Although miR-22 levels did not show significant differences in R-IPSS stratification, the effect size suggests a moderate association, indicating that with a larger sample size, these differences might become significant. Additionally, no substantial differences in miR-22 levels were detected concerning cytogenetic risk.

Contrary to our results, a study by Ma et al. reported significantly higher miR-22 levels in the plasma of MDS patients compared to healthy controls, with the most pronounced increase observed in high-risk MDS cases. Although no statistically significant differences were found between low- and high-risk patients in our study, the moderate effect size suggests a possible trend toward a difference in miR-22 levels between these groups. The same study found that increased miR-22 expression was associated with advanced stages of MDS, particularly in patients with excess blasts.

The lack of statistical significance in our study may be attributed to several factors. The limited sample size and heterogeneity of the studied groups may have reduced the statistical power of the analysis. Geographic and ethnic variations among the studied populations could also play a role. These differences highlight the need for larger cohort studies to further investigate the role of miR-22 in MDS.

2.2 MicroRNA-144

MiR-144, together with miR-451a, is highly expressed in erythroid tissues and plays a significant role in hematopoiesis. According to Qian et al., miR-144 regulates AKAP12, a crucial tumor suppressor gene involved in the ERK1/2 signaling pathway. The suppression of AKAP12 by miR-144 leads to increased cell proliferation and reduced apoptosis, which are critical for the pathogenesis of MDS. Inhibition of miR-144 activates the ERK1/2 pathway and enhances apoptosis in MDS cells.

Our study found significantly lower miR-144 levels in MDS patients (median 0.802) compared to healthy controls (median 1.933). Statistical analysis revealed a significant difference between the two groups ($p = 0.0003$) with an extremely large effect size ($r = -1.08$), highlighting the clinical significance of the observed differences. ROC curve analysis determined a cut-off value of 1.727, with high sensitivity (92.5%) and specificity (70%), distinguishing MDS patients from healthy controls.

When patients were stratified based on cytogenetic risk, miR-144 levels were significantly lower in intermediate-risk patients compared to those in the good-risk group ($p = 0.0287$, $r = -0.53$). Although no statistically significant differences were found in R-IPSS risk stratification, the effect size ($r = -0.51$) suggested a moderate trend toward differentiation between low-risk and high-risk patients.

Analysis of MDS subtypes revealed no statistically significant differences in miR-144 levels; however, the effect size ($r = -0.319$) suggests the possibility of significant differences in larger cohorts. These findings emphasize the diagnostic

potential of miR-144 and its role as a biomarker for more precise patient stratification in MDS.

In contrast to our results, which demonstrate lower miR-144 levels in MDS patients compared to healthy controls, the study by Zuo et al. found that miR-144 levels were elevated in MDS patients. Additionally, they observed that miR-144 levels were higher in patients with isolated aberrations such as del(7q)—associated with poor prognosis—and lower in those with del(20q)—linked to favorable prognosis. This finding contradicts our results, where miR-144 levels were higher in patients with good cytogenetic risk.

These discrepancies are likely due to methodological differences, such as the NanoString technology used for microRNA profiling in Zuo et al.'s study, compared to RT-qPCR in our study. Moreover, Zuo et al. focused on specific cytogenetic aberrations, while our study considered grouped cytogenetic risks. These observations underscore the need for further investigations to validate these findings.

2.3 MicroRNA-16

MiR-16 is a key regulator of angiogenesis in MDS through its direct influence on vascular endothelial growth factor (VEGF) [174]. VEGF is a major angiogenic factor that plays a central role in the regulation of hematopoietic stem cells and is associated with tumor proliferation and angiogenesis. Xiong et al. demonstrated that miR-16 directly interacts with the 3' untranslated region of VEGF, suppressing its expression [173]. In experimental models, miR-16 overexpression reduces VEGF secretion, inhibits migration and angiogenesis, whereas miR-16 inhibition has the opposite effect. These findings

suggest a tumor-suppressive role of miR-16 in MDS pathogenesis through VEGF inhibition and angiogenesis modulation. Furthermore, miR-16 targets multiple oncogenes, including BCL2 and MCL1, reinforcing its tumor-suppressive function [279]. Low miR-16 levels may disrupt the balance between apoptosis and cell proliferation, which is a key mechanism in MDS pathogenesis.

In line with these findings, our study identified significantly lower miR-16 levels in MDS patients compared to healthy controls. The median miR-16 level in MDS patients was 0.7855, while in healthy controls, it was significantly higher (1.764). The difference was statistically significant ($p = 0.0001$), highlighting miR-16's clinical relevance as a potential diagnostic biomarker. Similar findings were reported by Zuo et al., who confirmed that miR-16 levels were significantly lower in MDS patients compared to healthy controls [225][232]. ROC analysis demonstrated high diagnostic value for miR-16, with a cut-off value of 1.508, ensuring high specificity (80%) and sensitivity (82.5%).

When comparing miR-16 levels across different MDS subtypes and between patients with low and high risk according to R-IPSS, no statistically significant differences were observed ($p > 0.05$). However, the effect size in the R-IPSS risk group analysis was large, suggesting that a larger sample size might reveal a significant difference in miR-16 levels between low- and high-risk patients. Although our study did not detect such an association, Xiong et al. found that miR-16 levels were significantly lower in high-risk MDS patients compared to low-risk patients [173]. It is important to note that their analysis was performed on CD34+ stem cells from bone marrow, where an inverse relationship between miR-16 and VEGF levels was

observed. Lower miR-16 levels were associated with increased VEGF secretion and enhanced angiogenesis, emphasizing miR-16's role in angiogenesis regulation in high-risk MDS.

Regarding cytogenetic risk analysis, no statistically significant differences in miR-16 levels were found between good- and intermediate-risk groups.

2.4 Let-7a

Let-7a is a key regulator in HSC homeostasis through its interaction with the TGF-β and Wnt signaling pathways [88]. By suppressing the TGFBR1 and TGFBR2 receptors and the transcription factors SMAD2 and SMAD4, let-7a negatively regulates the TGF-β pathway, which is responsible for inhibiting proliferation and inducing apoptosis. Simultaneously, by reducing APC levels, let-7a activates the Wnt signaling pathway, stimulating HSC proliferation and self-renewal while preventing apoptosis [88].

The dysregulation of let-7a, observed in MDS, may disrupt this balance, leading to impaired hematopoiesis and disease progression. Suppression of TGF-β and aberrant activation of Wnt signaling due to let-7a dysregulation can contribute to apoptosis resistance and impaired differentiation, which are hallmarks of MDS.

Let-7a levels were significantly lower in MDS patients compared to healthy controls. The median level in MDS patients was 0.692, while in healthy controls, it was 1.669 ($p = 0.0002$). Similar results were reported by Zuo et al., who also found that let-7a levels were significantly lower in MDS

patients compared to controls ($p < 0.001$). ROC analysis confirmed the high diagnostic potential of let-7a, with an AUC of 0.8615 ($p = 0.0005$). The optimal cut-off value, set at 0.9775, provided a sensitivity of 71.79% and a specificity of 90.0%, highlighting let-7a's potential as a diagnostic biomarker.

Comparisons of let-7a levels across MDS subtypes and R-IPSS groups revealed statistically significant differences. Patients with multilineage dysplasia, ring sideroblasts, and del(5q) exhibited significantly lower let-7a levels than those with excess blasts ($p = 0.0250$). Similarly, patients with low-risk MDS (according to R-IPSS) had significantly lower let-7a levels compared to high-risk patients ($p = 0.0172$).

In contrast to our findings, Vasilatou et al. reported that let-7a levels were lower in CD34+ stem cells from the bone marrow of intermediate- and high-risk MDS patients (according to IPSS), whereas low-risk patients had higher levels [280]. They also found a strong negative correlation between let-7a expression and RAS protein levels in CD34+ cells. Reduced let-7a levels were associated with increased RAS protein expression, which in turn activates oncogenic signaling pathways and contributes to MDS progression and AML transformation.

These discrepancies between our study and Vasilatou et al. may be due to differences in the biological material analyzed—plasma in our study vs. bone marrow stem cells in theirs. Selective cellular mechanisms for microRNA release could lead to significant differences between extracellular and intracellular miRNA levels. As demonstrated in Pigati et al.'s study, cells selectively secrete specific miRNAs via exosomes

or other mechanisms, which may not reflect their intracellular levels [281]. This implies that in MDS, let-7a dysregulation in HSCs may result from direct pathogenic mechanisms, such as mutations or abnormal signaling. However, plasma let-7a levels might instead reflect a systemic response to the disease, including increased miRNA secretion from affected cells.

When analyzing cytogenetic risk, no statistically significant differences in let-7a levels were found between good- and intermediate-risk groups ($p = 0.8759$). These results suggest minimal clinical significance of let-7a in distinguishing cytogenetic risk groups. However, Vasilatou et al.'s analysis showed significantly lower let-7a expression in CD34+ bone marrow stem cells from intermediate- and high-cytogenetic risk patients compared to low-risk patients ($p = 0.035$) [280].

2.5 MicroRNA-451a

MiR-451a is a key regulator of erythropoiesis, whose expression is controlled by the transcription factor GATA-1 [99]. It is part of the miR-144/451 locus and is highly expressed in the erythroid lineage, playing a crucial role in maturation and survival of erythroid precursors. MiR-451a interacts with multiple target mRNAs, suppressing their expression and facilitating the late stages of erythropoiesis [112]. Experimental models demonstrate that miR-451a deficiency leads to severe erythroid maturation defects, highlighting its critical role in normal hematopoiesis [99].

In the context of MDS, erythropoiesis disorders are a major pathogenetic mechanism, leading to anemia and its associated clinical manifestations. Dysregulation of miR-451a may

contribute to ineffective erythropoiesis, a hallmark of MDS, by disrupting maturation and survival mechanisms of erythroid cells.

Consistent with these findings, miR-451a levels were significantly lower in MDS patients (median: 0.8215) compared to healthy controls (median: 2.132, $p < 0.0001$). This strong statistical significance highlights the diagnostic potential of miR-451a.

ROC analysis demonstrated a high discriminative value of miR-451a, with an AUC of 0.9175 ($p < 0.0001$). The optimal cut-off value was determined to be 1.161, ensuring a sensitivity of 70.0% and specificity of 100.0%, which is essential for minimizing false-positive results.

When comparing MDS subtypes (multilineage dysplasia, ring sideroblasts, and del(5q) vs. excess blasts), miR-451a levels did not show statistically significant differences ($p > 0.05$). However, in the R-IPSS-based stratification, low-risk patients had significantly lower miR-451a levels than high-risk patients ($p = 0.0449$), underscoring its clinical significance in risk stratification.

Comparing cytogenetic risk groups, patients with intermediate risk had significantly lower miR-451a levels than those with good risk ($p = 0.0481$). These results emphasize the potential role of miR-451a as a marker for distinguishing cytogenetic risk groups.

Similar to our findings, Merkerova et al. also reported lower miR-451a levels in MDS patients compared to healthy controls [233]. However, their study found even lower miR-451a levels in high-risk MDS (according to IPSS) compared to low-risk

MDS, whereas our study (using R-IPSS) observed the opposite trend. These discrepancies may arise from differences in risk classification scales (IPSS vs. R-IPSS) and the limited sample size.

Conversely, Zuo et al. reported higher miR-451a levels in MDS patients compared to healthy controls [225]. This discrepancy may be attributed to differences in analytical techniques, geographical and demographic variations, and population selection criteria.

Regarding cytogenetic risk, Zuo et al. found that patients with isolated del(7q)/-7 (poor cytogenetic risk) exhibited higher miR-451a levels, while those with isolated del(20q) (good cytogenetic risk) had lower levels [225]. In contrast, our study found that patients with good cytogenetic risk had higher miR-451a levels than those with intermediate risk. This divergence may result from differences in cohort composition—our study excluded poor-risk patients due to their low number, whereas Zuo et al. focused on specific cytogenetic abnormalities.

Despite these variations, all studies highlight the importance of miR-451a as a diagnostic marker in MDS. These findings further justify the need for larger cohort studies to validate these observations and fully explore miR-451a's potential in clinical practice.

3. Correlations Between microRNAs and Clinical Parameters

This study also analyzes the correlations between the five microRNAs (miR-22, miR-144, miR-16, let-7a, and miR-451a)

and their relationship with hematological and biochemical parameters in patients with MDS. The results show a statistically significant positive correlation between all the studied microRNAs. The strongest correlations were found between miR-16 and miR-451a, as well as between miR-16 and miR-144. Strong associations were also observed between miR-144 and miR-451a, and between miR-22 and miR-144. These results highlight the potential functional interconnection between microRNAs within key biological processes.

The strong correlations between miR-16, miR-144, and miR-451a are likely due to their key role in erythropoiesis. MiR-16 is involved in the regulation of ribosomal biogenesis and the proliferation of erythroid precursors [282]. MiR-144, on the other hand, regulates oxidative stress, influencing erythroid cell differentiation [283]. MiR-451a is a crucial regulator of GATA1—a key transcription factor for erythroid differentiation—further emphasizing the common mechanisms in which these microRNAs participate [99]. Their functional connectivity within erythropoiesis likely contributes to the observed strong correlations.

Additionally, miR-144 and miR-451a share a common locus on the long arm of chromosome 17 (17q11.2), which is essential for regulating erythroid homeostasis and terminal erythropoiesis [283]. The combined action of these two microRNAs ensures stability and efficiency in erythropoiesis, particularly under stress conditions. The established strong correlation between miR-144 and miR-451a may be explained by their shared regulatory mechanism and synergistic action on genes related to terminal erythroid cell differentiation. The

absence of these microRNAs leads to erythroid hyperplasia and ineffective erythropoiesis.

The strong correlation between miR-22 and miR-144 can be explained by their shared role in regulating the PTEN/PI3K/AKT signaling pathway. PTEN is a well-known tumor suppressor that regulates the PI3K/AKT signaling pathway by dephosphorylating PIP3 to PIP2 [284]. This leads to the inhibition of AKT activity, limiting cell proliferation and promoting apoptosis. PTEN dysfunction is associated with various myeloid neoplasms, including MDS and AML [285]. MiR-22 and miR-144 have been identified as regulators of PTEN, suppressing it through direct binding to its 3' end [156][286]. This reduces PTEN expression and leads to increased AKT activity. MiR-22 forms a regulatory loop in the PTEN/AKT pathway, modulating the activity of FoxO transcription factors, which are key to the cell cycle and apoptosis [156]. At the same time, miR-144 has been shown to directly inhibit PTEN, increasing tumor cell proliferation, migration, and invasion by activating the PI3K/AKT pathway [286]. Both microRNAs suppress PTEN expression, leading to increased AKT activity and stimulating cell growth and proliferation. This interaction is critical for cellular homeostasis and is associated with the development of myeloid neoplasms.

The results from the correlation analysis among the selected microRNAs emphasize the complex network of interactions between them, reflecting their shared functional roles in key biological processes. Future studies may confirm these mechanisms and expand their potential clinical applications.

A correlation analysis was also conducted between the selected five microRNAs and specific laboratory parameters

(hemoglobin, reticulocytes, leukocytes, neutrophils, platelets, LDH, ferritin, beta-2-microglobulin, and erythropoietin). These parameters represent key markers reflecting different aspects of the hematological and metabolic status of MDS patients. Including these markers in the correlation analysis with microRNAs aims to explore possible relationships between microRNA levels and clinical manifestations of the disease.

The results indicate a lack of significant correlations between microRNA levels and laboratory parameters in peripheral blood counts. This is consistent with existing studies. For example, Zuo et al. investigated the levels of miR-16 and let-7a in plasma from MDS patients and found that their levels did not significantly correlate with parameters such as hemoglobin, leukocytes, and platelets [232].

Similarly, Merkerova et al. found that the levels of miR-451a, miR-144, and miR-16 differ between MDS patients and healthy controls but lack clear associations with specific laboratory parameters [233]. Likewise, Ma et al. reported that although miR-22 is associated with high-risk forms of MDS, it does not show a direct relationship with routine laboratory parameters in peripheral blood [179]. These observations suggest that plasma microRNAs may reflect systemic or pathogenic processes that do not directly manifest in standard laboratory parameters of peripheral blood.

The similarities between the current results and the literature data highlight that plasma microRNAs may have potential as systemic biomarkers for MDS. However, their lack of direct correlation with laboratory parameters in peripheral blood underscores the need for additional functional studies to understand their mechanisms of action.

Despite the lack of significant correlations between the studied microRNAs and standard laboratory parameters in peripheral blood, our study expands the analysis by including parameters related to disease activity, erythroid, and iron metabolism. This more detailed assessment revealed significant correlations between certain microRNAs and key biochemical markers.

Lactate dehydrogenase (LDH) is an indicator of cellular metabolism and tissue damage, which is often elevated in MDS patients. In our study, we identified a moderate correlation between LDH levels and miR-22, miR-144, and let-7a.

MiR-22 plays a key role in apoptosis through its interaction with the p53 signaling pathway [153]. As described by Tsuchiya et al., miR-22 is a direct target of p53 and modulates cell fate by suppressing p21 expression—a key apoptosis inhibitor. This enhances p53-dependent apoptosis, directing cells toward programmed cell death. This mechanism is relevant in the context of MDS, where increased apoptosis and destruction of hematopoietic cells lead to elevated LDH. Therefore, the moderate positive correlation between miR-22 and LDH may reflect this relationship.

The correlation between let-7a, miR-144, and LDH observed in this study may be explained by their role in glycolysis regulation via GLUT transporters. Let-7a suppresses GLUT12 expression, leading to reduced glucose transport, lactate production, and glycolysis activity [287]. MiR-144 modulates glycolysis by suppressing GLUT1 expression, a major glucose transporter [288]. In cancer cells, this results in reduced glucose uptake and lactate production.

Ferritin is a key marker of iron overload in the body and plays an important role in MDS pathogenesis. Elevated ferritin levels are often associated with transfusion dependence, ineffective erythropoiesis, and increased iron accumulation in tissues. Studies highlight that ferritin levels above 1000 ng/mL are an independent prognostic factor for lower overall survival in MDS patients, regardless of the number of transfusions [289]. Additionally, iron overload worsens bone marrow function by increasing oxidative stress and genomic instability. These effects are particularly harmful to erythroid precursors, which are vulnerable to damage from reactive oxygen species (ROS) [290]. Furthermore, iron overload can disrupt the bone marrow microenvironment and suppress hematopoiesis, leading to ineffective erythropoiesis and worsening anemia.

In this study, moderate positive correlations were found between ferritin levels and three of the studied microRNAs: miR-144, miR-16, and miR-451a. These microRNAs have a proven role in regulating erythroid differentiation, influencing the maturation and functionality of erythroid precursors. This shared association with erythropoiesis may explain their correlation with ferritin, which is directly linked to iron metabolism and erythropoiesis in MDS. The correlations between ferritin and microRNAs highlight their potential role as molecular markers for assessing iron overload and related damage.

In our study, a moderate positive correlation was found between miR-22 and erythropoietin levels. This relationship can be explained by miR-22's role as a modulator of erythropoietin-stimulated pathways that promote erythroid cell maturation and differentiation [219].

To better understand the relationships between microRNAs and laboratory parameters, an additional analysis was conducted using simple linear regression. The results suggest that such predictive models cannot be reliably constructed, indicating the complexity of microRNA interactions and their indirect impact on laboratory markers.

A correlation analysis between microRNA levels and bone marrow blast percentage revealed a statistically significant moderate positive correlation between let-7a levels and blast percentage. This suggests that let-7a could serve as a biomarker for disease progression in MDS, though further studies are needed to confirm its predictive value.

4. Predictive Value of microRNAs in MDS

4.1 Diagnostic Potential of microRNAs

This study conducted an in-depth analysis of the predictive diagnostic value of five microRNAs in differentiating MDS patients from healthy controls. The research is based on simple and multivariate logistic regression analysis, as well as the LASSO regression method, to identify the strongest biomarkers and evaluate their combined effectiveness.

The initial results from the simple logistic regression analysis show that four of the microRNAs with established differences between the groups—miR-16, miR-144, miR-451a, and let-7a—are statistically significant predictors. The highest predictive value was demonstrated by miR-451a, which explains 42% of the variation in the model (Pseudo-R^2 = 0.4193) and exhibits exceptional discriminatory ability. MiR-16 and miR-144 also show moderate predictive value (Pseudo-R^2

= 0.2428 and 0.2552, respectively), while let-7a has the lowest value among the four studied microRNAs (Pseudo-R^2 = 0.2048).

Following the simple logistic regression analysis to assess the diagnostic value of individual microRNAs, a LASSO analysis was performed to identify the strongest predictors and evaluate how they interact within the model. The LASSO analysis helps address multicollinearity, which is common among highly correlated microRNAs, while eliminating less significant variables. This provides a more precise, interpretable, and stable diagnostic model that accounts for the complex interactions between different microRNAs. The LASSO analysis identified miR-22 and miR-451a as the most significant biomarkers. These two microRNAs were included in a simplified model that demonstrated excellent discriminatory ability (AUC = 0.994) and high accuracy (95.92%). These results confirm the key role of miR-22 and miR-451a as reliable predictors for distinguishing MDS patients from healthy controls.

Additionally, a multivariate logistic regression analysis based on miR-22 and miR-451a confirmed their combined predictive effect, with minimal multicollinearity (VIF ≈ 1.18) and significantly improved model performance (AUC = 0.995, AIC = 13.73). These results demonstrate that the combined use of these two microRNAs can provide a reliable tool for MDS diagnosis while ensuring model stability and accuracy.

Although miR-22 did not show statistically significant differences in levels between MDS patients and healthy controls, it was included in the LASSO analysis due to its potential contribution to improving the diagnostic model. This

decision is justified by the principles of LASSO regression, which aim to identify the most significant predictors by minimizing model complexity and eliminating variables with negligible contributions.

MiR-22 is associated with the regulation of epigenetic mechanisms and hematopoietic differentiation, with elevated levels linked to disease progression and high-risk MDS. This suggests that even without significant differences in its levels between patients and healthy controls, miR-22 may interact synergistically with other microRNAs (such as miR-451a) to enhance the model's discriminatory ability.

The LASSO analysis accounts for this complex contribution by considering not only the individual effect of each microRNA but also its role in the context of other predictors. Including miR-22 in the model reduces the AIC value and increases accuracy, indicating that it improves the model's overall performance. This is likely due to its ability to complement the information provided by other microRNAs, such as miR-451a, and to capture additional aspects of MDS pathophysiology that are not fully reflected by other markers.

Therefore, the results of the LASSO analysis demonstrate that miR-22, despite its initial lack of diagnostic significance, plays a crucial role in improving the model, highlighting the importance of a combined approach in biomarker analysis. This underscores the potential of integrative methods such as LASSO to identify hidden interconnections that might otherwise remain unnoticed.

The current results emphasize the significance of the selected microRNAs as potential diagnostic biomarkers for MDS, with

miR-451a standing out as the leading marker, while miR-22 further enhances the model's diagnostic value.

In addition to the diagnostic analysis, this study also evaluated the predictive value of the selected microRNAs in relation to risk stratification according to R-IPSS and MDS subtypes. For this purpose, LASSO analysis was again applied, identifying let-7a as the strongest predictor for risk stratification and subtype differentiation. This microRNA shows a significant association with belonging to the high-risk group according to R-IPSS, as well as with the group of patients with a high percentage of blasts.

The results of the simple logistic regression analysis further confirm the importance of let-7a. It shows that elevated levels of let-7a are associated with increased odds of belonging to the high-risk group. Although the model's discriminatory ability (AUC ≈ 0.71–0.74) remains moderate, these results highlight the potential of let-7a as a biomarker for risk stratification.

For distinguishing MDS subtypes, the LASSO analysis again identified let-7a as the only significant predictor. Higher levels of let-7a are associated with belonging to subtypes with a high percentage of blasts. The simple logistic regression analysis confirms these results, with the model demonstrating good discriminatory ability (AUC = 0.7219).

To our knowledge, no published studies have directly used simple and multivariate logistic regression analysis to assess the diagnostic value of the selected microRNAs in MDS patients. While previous studies, such as those by Zuo et al. [232] and Merkerova et al. [233], have used ROC analyses to evaluate the discriminatory ability of microRNAs, our approach

is distinguished by the integration of logistic regression to calculate regression coefficients (β), odds ratios (OR), and the coefficient of determination (Pseudo-R^2). This allows not only for the assessment of diagnostic accuracy but also for determining the relative contribution of each microRNA to the model.

Our results highlight the advantages of this approach by identifying miR-22 and miR-451a as key biomarkers with high discriminatory ability and model stability. This suggests that using logistic regression can provide a deeper understanding of the role of microRNAs in MDS diagnostics, complementing existing data. The present study underscores the potential of miR-22 and miR-451a as non-invasive biomarkers for MDS diagnosis. Since these microRNAs were studied in plasma samples, this offers an alternative to invasive diagnostic procedures such as bone marrow biopsy and aspiration, which are currently necessary for diagnosis. However, while microRNAs cannot replace existing diagnostic algorithms, they can significantly aid them, especially in cases where bone marrow morphological characteristics are subtle or inconclusive. In such cases, the presence of additional diagnostic tools, such as molecular-genetic and cytogenetic analyses, in combination with microRNAs, may play a key role in confirming the diagnosis.

The limitations of this study include the relatively small sample size and the lack of external validation of the results in independent patient cohorts. Future studies with larger populations and long-term follow-up will be necessary to confirm the applicability of these microRNAs in clinical practice.

4.2 Prognostic value for overall survival

This study also conducted an analysis aimed at evaluating the prognostic value of five microRNAs concerning the survival of MDS patients using a Cox proportional hazards analysis. The results show that all examined microRNAs have hazard ratios (HR) greater than 1, suggesting a trend toward an increased risk of events with higher expression levels. However, the confidence intervals for HR include 1 for all microRNAs, indicating a lack of statistical significance. Among the studied microRNAs, the lowest AIC value was observed for the model with let-7a (AIC = 103.8), which may suggest some potential for this microRNA as a prognostic marker.

Following the univariate analysis, a combined model for miR-144 and miR-16 was developed to assess their joint prognostic value for survival in MDS patients. These two microRNAs were selected based on their individual results from the Cox proportional hazards analysis, which indicated a trend toward an increased risk of events with higher expression levels (HR > 1). Additionally, both microRNAs demonstrated lower AIC values compared to the null model, suggesting that they may contribute to prognostic value. By combining these two microRNAs, the analysis aimed to explore their independence and potential synergistic effect on the prognostic model. Although both microRNAs showed a trend toward increased risk in the individual analysis, their combined effect was not statistically significant.

Despite these limitations, the observed trends for miR-144, miR-16, and let-7a suggest that these microRNAs may have prognostic value in larger cohorts or when used in combination with other clinical and molecular markers. Further studies with

larger sample sizes and multivariate models are needed to confirm the potential role of these microRNAs as prognostic biomarkers in MDS patients. The observed trends toward increased risk with miR-144, miR-16, and let-7a are consistent with literature findings highlighting their potential role in MDS prognosis.

A study by Ma et al. analyzed samples from 45 MDS patients, divided into low- and high-risk groups according to R-IPSS, along with 20 healthy controls. A Kaplan-Meier analysis revealed that elevated miR-22 levels were associated with lower overall survival [179]. The authors explained these results by the role of miR-22 in inhibiting TET2, leading to hypermethylation and impaired differentiation of hematopoietic cells. Our results also show a trend toward increased risk with high miR-22 levels, which aligns with the reported findings.

Zuo et al. studied plasma levels of miR-16 and let-7a in 50 MDS patients and 76 healthy controls [232]. They found that lower levels of both microRNAs were associated with decreased overall survival and progression-free survival ($p < 0.001$). In another study involving 72 patients with normal cytogenetics, higher levels of let-7a and miR-16, as part of a prognostic model including seven microRNAs, predicted worse survival outcomes [225]. The differences in findings are likely due to variations in study populations and methodologies. In our study, a trend toward increased risk with high levels of let-7a and miR-16 was observed, which is consistent with the results from Zuo et al. [225].

MiR-144 and miR-451a were also included in the combined model in the study by Zuo et al. [225]. Specifically for miR-451a, a multivariate analysis showed that low levels were

an independent predictor of shorter progression-free survival (HR = 0.072, p = 0.006). This aligns with the findings of Merkerova et al., who also identified miR-451a as a prognostic marker [233]. In their study, low plasma levels of miR-451a were associated with poor prognosis, with a Kaplan-Meier analysis demonstrating significantly shorter progression-free survival in patients with low miR-451a levels. For miR-144, the study by Zuo et al. found that higher levels were linked to worse overall survival [225]. Our study also observed a trend toward increased risk with high levels of miR-144 and miR-451a, but without statistical significance.

The results of our study indicate trends toward an increased risk of events with elevated levels of the studied microRNAs, although statistical significance was not achieved. Data from the literature support our observations, emphasizing the potential of miR-22, miR-144, miR-16, let-7a, and miR-451a as prognostic biomarkers in MDS. Future studies with larger cohorts are necessary to validate these findings and clarify the role of microRNAs as prognostic biomarkers in MDS.

5. Conclusion

In summary, this study expands the understanding of the role of microRNAs as diagnostic and prognostic biomarkers in MDS. Through an in-depth analysis of five microRNAs—miR-22, miR-144, miR-16, let-7a, and miR-451a—significant differences in their expression levels between MDS patients and healthy controls were identified, highlighting their diagnostic potential. Among the studied microRNAs, miR-451a demonstrated the highest discriminatory power, while the other microRNAs also showed significant associations with clinical features of the disease.

Beyond their diagnostic potential, the analyses revealed specific correlations between microRNAs and laboratory parameters, as well as their interactions with biological processes related to MDS pathogenesis. However, the lack of significant correlations with certain clinical parameters suggests that these molecules play a more indirect role in the disease's pathophysiology. Although the prognostic value of microRNAs was not definitively established, observed trends suggest their potential for inclusion in prognostic models.

The study results indicate that microRNAs can be used as non-invasive molecular biomarkers that complement traditional diagnostic methods and aid in risk stratification for MDS patients. The use of multivariable approaches, such as LASSO analysis, enabled the identification of miR-22 and miR-451a as key elements in a diagnostic model with high accuracy and stability. The limitations of this study, including sample size and the lack of external validation, underscore the need for future research on larger cohorts to confirm these findings and uncover the full potential of microRNAs in clinical practice.

The conclusions of this study reinforce the significance of microRNAs as biomarkers for MDS, offering new opportunities for improving diagnosis, prognosis, and the individualization of therapeutic strategies for this heterogeneous disease.

Key Findings:

1. MDS patients have significantly lower levels of miR-451a, miR-144, miR-16, and let-7a compared to healthy controls.

2. Increased levels of let-7a are associated with a higher blast percentage in the bone marrow and higher risk according to R-IPSS.
3. Lower levels of miR-144 and miR-451a are associated with intermediate cytogenetic risk in MDS patients.
4. Levels of miR-22, miR-144, and let-7a show a moderate positive correlation with LDH, reflecting their role in cellular metabolism and apoptosis.
5. High levels of miR-144, miR-16, and miR-451a correlate with elevated ferritin levels, suggesting their involvement in iron metabolism.
6. MiR-22 and miR-451a were identified as the strongest biomarkers in the diagnostic model.

VI. Contributions:

- For the first time in Bulgaria, a study has been conducted on plasma microRNAs (miR-16, miR-144, miR-22, miR-451a, and let-7a) in patients with myelodysplastic syndrome (MDS), enhancing traditional diagnostic methods and offering new perspectives in disease assessment.
- For the first time, the relationship between the levels of specific microRNAs (miR-16, miR-144, miR-451a, etc.) and key biochemical parameters (ferritin, erythropoietin, LDH) reflecting erythropoiesis and iron metabolism in MDS has been analyzed. These correlations highlight the potential of microRNAs to provide additional insights into pathological processes in the bone marrow.
- An integrated statistical approach (simple and multivariate logistic regression, LASSO analysis, ROC analysis) has been applied for the first time, allowing for the assessment of both the individual and combined predictive value of the studied microRNAs for MDS diagnosis and risk stratification.
- It has been confirmed that the levels of miR-16, miR-144, miR-451a, and let-7a are significantly lower in patients with MDS compared to healthy controls, aligning with international data on their role as potential diagnostic markers.
- It has been confirmed that let-7a correlates with blast percentage and R-IPSS risk, supporting literature data on its role in MDS progression and its potential role in risk stratification.

- A diagnostic model has been developed based on LASSO regression and the combination of multiple microRNAs (miR-22 and miR-451a), demonstrating exceptionally high sensitivity and specificity in distinguishing MDS patients from healthy controls.
- The potential non-invasiveness of plasma microRNAs highlights their role as an additional tool to standard diagnostic methods (morphological assessment, cytogenetics, molecular tests), especially in cases with challenging histopathological evaluation.

VII. Scientific publications related to the dissertation:

1. Атанасова С, Мичева И. Нарушения в метилирането при миелодиспластичен синдром – терапевтични възможности. MEDINFO. 2022;4

2. Атанасова С., Мичева И. Епигенетични механизми при миелодиспластичен синдром. Медицински преглед. 2022

3. Micheva, I. D., & Atanasova, S. A. MicroRNA dysregulation in myelodysplastic syndromes: implications for diagnosis, prognosis, and therapeutic response. *Frontiers in oncology*, *14*, 2024

VIII. Acknowledgments

This dissertation is the result of years of effort, which would not have been possible without the support, guidance, and assistance of many individuals to whom I am deeply grateful:

- I sincerely express my gratitude to my scientific supervisor, Prof. Ilina Micheva, MD, for her patience, professionalism, and invaluable guidance throughout the research process.
- Special thanks to Assoc. Prof. Trifon Chervenkov, MD and Maria Teneva, whose dedication to the preservation and processing of the samples made this study possible.
- I am also grateful to the entire team of the Clinic of Hematology for providing the necessary conditions and support throughout the research.
- A heartfelt thank you to all the patients and healthy volunteers who participated in this study. Your cooperation was invaluable and contributed significantly to scientific progress in this field.
- I extend my deepest gratitude to my family for their unconditional support, understanding, and patience throughout this journey. Your encouragement and belief in me have been a constant source of motivation.