



# POPULATION-BASED APPROACHES USING THE SEER CANCER REGISTRY TO STUDY SURVIVAL OUTCOMES IN MYELOYDYSPLASTIC SYNDROMES AND CHRONIC MYELOMONOCYTIC LEUKEMIA

## Citation

Lesegretain, Arnaud. 2023. POPULATION-BASED APPROACHES USING THE SEER CANCER REGISTRY TO STUDY SURVIVAL OUTCOMES IN MYELOYDYSPLASTIC SYNDROMES AND CHRONIC MYELOMONOCYTIC LEUKEMIA. Master's thesis, Harvard Medical School.

## Link

<https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37375158>

## Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

<https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFIOWI2M2EwYTg>

## Accessibility

<https://accessibility.huit.harvard.edu/digital-accessibility-policy>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#)

**POPULATION-BASED APPROACHES USING THE SEER CANCER REGISTRY TO  
STUDY SURVIVAL OUTCOMES IN MYELOYDYSPLASTIC SYNDROMES AND  
CHRONIC MYELOMONOCYTIC LEUKEMIA.**

by

Arnaud Lesegetrain

A Dissertation Submitted to the Faculty of Harvard Medical School

In Partial Fulfillment of

The Requirements for the Degree of Master of Medical Sciences in Clinical Investigation

(MMSCI)

Harvard University

April 2023

Area of Concentration: Myeloid Malignancies

Project Advisors: Dr. Amir T. Fathi, Dr Naval Daver, Dr Finnian McCausland

I have reviewed this thesis. It represents work done by the author under my  
guidance/supervision.

Primary Mentor: Dr Amir T. Fathi

## TABLE OF CONTENT

|                                                               |    |
|---------------------------------------------------------------|----|
| THESIS COMMITTEE MEMBERS .....                                | 3  |
| ACKNOWLEDGMENTS.....                                          | 4  |
| OVERVIEW OF THESIS PAPERS.....                                | 5  |
| PAPER #1.....                                                 | 7  |
| SUPPLEMENTARY TABLES PAPER #1 .....                           | 10 |
| PAPER #2.....                                                 | 23 |
| SUPPLEMENTARY TABLES PAPER #2.....                            | 51 |
| PAPER SUMMARIES.....                                          | 57 |
| DISCUSSION, PERSPECTIVE, AND POSSIBLE FUTURE DIRECTIONS ..... | 59 |
| BIBLIOGRAPHY .....                                            | 62 |

## **THESIS COMMITTEE MEMBERS**

### **Primary Mentor**

Dr Amir T. Fathi, MD

Associate Professor, Medicine, Harvard Medical School

Director, Leukemia Program, Medicine, Massachusetts General Hospital

### **Content Expert**

Dr Naval Daver, MD

Associate Professor in the Department of Leukemia at MD Anderson Cancer Center

### **External Expert**

Pr Tapan Kadia, MD

Professor, Department of Leukemia, Division of Cancer Medicine, M.D. Anderson Cancer Center, Houston, TX

### **Program Representatives**

Ajay K. Singh, MBBS, FRCP, MBA

Senior Associate Dean for Postgraduate Medical Education, Director, Master in Medical Sciences in Clinical Investigation (MMSCI) Program Clinical Director of the Renal Division and Director, Dialysis Services and Associate Professor of Medicine at HMS

Dr Finnian McCausland, MBBCh, MMSc, FRCPI

Assistant Professor of Medicine, Faculty Director in Postgraduate Medical Education and Co-Director of the Master of Medical Sciences in Clinical Investigation program, Harvard Medical School. Associate physician and attending nephrologist at Brigham and Women's Hospital

Dr Martina McGrath, MBBCh, FRCPI

Associate Director of Harvard Medical School's Master of Medical Science in Clinical Investigation (MMSCI) program and co-directs the Translational Investigation track of the MMSCI program. Director of Academic Support for PGME Master's programs

## **ACKNOWLEDGMENTS.**

I wish to thank Dr Amir Fathi, my primary mentor, for his invaluable expertise, thoughtful guidance, and tireless patience throughout the program. I would also like to recognize all my co-authors, in particular Dr Andrew Brunner for his deep clinical insights and thorough reviews and Andrew King for his coaching about methods and biostatistics. Thanks to all of you, I have learned to think and write much more deeply and precisely about clinical and epidemiological questions. I have also been humbled by the complexity and power of statistical methods and their importance in the toolkit of a clinical researcher. I am also very grateful to Dr Naval Daver for his support as content expert during the program and for his encouragement to apply to the MMSCI program. Dr Antoine Yver was also instrumental in building my confidence and in giving me the courage to candidate for the program. I would also like to thank Simon King and Dr Ken Takeshita from Daiichi Sankyo for their unwavering support.

Furthermore, I would like to thank the MMSCI program leadership, in particular Dr Ajay Singh for his vision and passion to lead such a world-class curriculum and Dr Finnian McCausland for his guidance as program representative, his teaching and dry sense of humor. Nothing would be possible without the entire program team, in particular Molly Gallagher.

I also want to thank my parents for the values of hard work, curiosity, and humility they shared with me. Finally, I am forever grateful to my wife, Corine, and children, Solène, Tanguy and Garance who encouraged me from day 1 throughout the ups and downs of my research, including the many weekends spent at my desk and away from them.

## OVERVIEW OF THESIS PAPERS.

Racial disparities in cancer care and outcomes have been extensively documented among patients with solid tumor malignancies, such as cancers of the colon, breast, lung, and head and neck and in hematologic malignancies [1,2]. Black/African American patients have the highest death rate and shortest survival of any racial/ethnic group in the United States (US) for most cancers [3]. With the increase of overall cancer incidence and population growth among minorities in the US, health care disparities and access to care are increasingly important priorities [4,5].

In the first paper, we used the SEER cancer registry program to characterize and compare characteristics and survival outcomes between White and Black patients with myelodysplastic syndromes (MDS) [6]. We first found significant differences in epidemiological and disease characteristics between both cohorts. We then implemented various analytical methods to address challenges in data capture and diagnostic complexity in MDS in order to compare overall survival (OS) between White and Black patients. We reported an association between OS and race favoring Black patients which was unexpected. We further refined this overall association by the subgroup analyses we performed and reported a novel finding in the variation in both the direction and the magnitude of the association of race and OS between various MDS histology subsets. Our paper also shed light on the high proportion of the provisional diagnosis category of “MDS-not otherwise specified” in MDS cancer registry and its implication for disease surveillance [7].

Again, using the SEER registry, the second paper focused on another myeloid malignancy, chronic myelomonocytic leukemia (CMML) [8], formerly included in the

MDS category, a disease characterized by dysplasia, monocytosis, and a proliferative capacity, with a generally poor prognosis. The paper explored how survival outcomes have evolved over time with a focus on long-term survivors and on subgroups defined by sex, age and race.

**PAPER #1.**

Lesegretain A, Brunner A, King AJ, Laadem A, Fell G, Fathi AT. Comparison of demographics, disease characteristics, and outcomes between Black and White patients with myelodysplastic syndromes: A population-based study. *Leuk Res.* 2023;125:107006.  
doi:10.1016/j.leukres.2022.107006



## Research paper

# Comparison of demographics, disease characteristics, and outcomes between Black and White patients with myelodysplastic syndromes: A population-based study

Arnaud Lesegretain<sup>a,b</sup>, Andrew Brunner<sup>c</sup>, Andrew J. King<sup>d</sup>, Abderrahmane Laadem<sup>b</sup>, Geoffrey Fell<sup>c</sup>, Amir T. Fathi<sup>c,\*</sup>

<sup>a</sup> Harvard Medical School, 25 Shattuck St, Boston, MA 02115, USA

<sup>b</sup> Daiichi Sankyo, 211 Mt Airy Rd, Basking Ridge, NJ 07920, USA

<sup>c</sup> Massachusetts General Hospital, 55 Fruit St, Boston, MA 02114, USA

<sup>d</sup> Department of Health Care Policy, Harvard Medical School, 25 Shattuck St, Boston, MA 02115, USA

<sup>e</sup> Dana-Farber Cancer Institute, 450 Brookline Ave, Boston, MA 02215, USA

## ARTICLE INFO

## Keywords:

Myelodysplastic syndromes

Registry

Race

Survival analysis

MDS-U

MDS-NOS

## ABSTRACT

Racial disparities in cancer care and outcomes have been well documented in various malignancies, with Black patients having the highest death rate and shortest survival of any racial/ethnic group in the United States (US) for most cancers. However, there have been limited studies on racial/ethnic disparities in myelodysplastic syndromes (MDS). Our study characterized and compared differences in baseline demographics, clinical characteristics, socioeconomic factors, and overall survival (OS) between Black and White patients with MDS in the US. We used the Surveillance, Epidemiology, and End Results (SEER) Program and included 37,562 patients (Black, 8.1 %; White, 91.9 %) diagnosed between 2001 and 2013. We observed significant differences in baseline characteristics between cohorts. In a univariate analysis, Black race was associated with longer survival (hazard ratio [HR]: 0.83; 95 % confidence interval [CI], 0.79–0.86;  $p < 0.001$ ). The association between race and survival was attenuated but remained significant in various models to adjust for differences in baseline characteristics (HR in multivariable analysis, 0.92; 95 % CI, 0.87–0.96;  $p < 0.001$ ). Subgroup analysis by histology revealed differences in the association between race and OS. Refractory anemia (RA), RA with ring sideroblasts, and MDS—not otherwise specified, a category in SEER representing a poorly defined MDS subset for 52 % of cases in our study, favored Black patients. RA with excess blasts favored White patients. The overall finding that Black race is associated with better OS outcomes, when compared with White patients, needs to be interpreted with caution and nuanced by histology. Additional research to explore these associations is warranted.

## 1. Introduction

Myelodysplastic syndromes (MDS) are a heterogeneous group of clonal myeloid neoplasms characterized by morphologic dysplasia, ineffective hematopoiesis, progressive and refractory cytopenias, and risk of transformation to acute myeloid leukemia (AML) [1,2]. Clinical

outcomes are highly variable, with some lower-risk patients with MDS experiencing normal life expectancy with minimal cytopenias and those with higher-risk MDS having very poor outcomes owing to rapid bone marrow failure, infections, bleeding, and leukemic transformation [2].

Racial disparities in cancer care and outcomes have been well documented, particularly among patients with solid tumor

**Abbreviations:** AML, acute myeloid leukemia; CVD, cardiovascular disease; HR, hazard ratio; ICD-O3, International Classification of Diseases for Oncology version 3; IQR, interquartile range; MDS, myelodysplastic syndrome; MDS del5q, myelodysplastic syndrome associated with isolated del(5q); MDS-NOS, myelodysplastic syndrome—not otherwise specified; MDS-U, myelodysplastic syndrome-unclassifiable; OS, overall survival; RA, refractory anemia; RAEB, refractory anemia with excess blasts; RARS, refractory anemia with ring sideroblasts; RCMD, refractory cytopenia with multilineage dysplasia; RN, refractory neutropenia; RT, refractory thrombocytopenia; SEER, Surveillance, Epidemiology, and End Results; WHO, World Health Organization.

\* Correspondence to: Massachusetts General Hospital, 275 Cambridge St, PO Box 229, Boston, MA 02114, USA.

E-mail addresses: [alesegretain@hms.harvard.edu](mailto:alesegretain@hms.harvard.edu) (A. Lesegretain), [abrunner@mgh.harvard.edu](mailto:abrunner@mgh.harvard.edu) (A. Brunner), [King@hcp.med.harvard.edu](mailto:King@hcp.med.harvard.edu) (A.J. King), [alaadem@dsi.com](mailto:alaadem@dsi.com) (A. Laadem), [GeoffreyG\\_Fell@dfci.harvard.edu](mailto:GeoffreyG_Fell@dfci.harvard.edu) (G. Fell), [afathi@mgh.harvard.edu](mailto:afathi@mgh.harvard.edu) (A.T. Fathi).

<https://doi.org/10.1016/j.leukres.2022.107006>

Received 23 August 2022; Received in revised form 13 December 2022; Accepted 23 December 2022

Available online 26 December 2022

0145-2126/© 2022 Elsevier Ltd. All rights reserved.

malignancies, such as cancers of the colon, breast, lung, and head and neck and in hematologic malignancies, including AML [3–5]. Collectively, Black patients have the highest death rate and shortest survival of any racial/ethnic group in the United States (US) for most cancers [6]. With the increase of overall cancer incidence and population growth among minorities in the US, health care disparity and access to care are increasingly important priorities. As a result, prominent professional organizations have embraced ambitious programs to increase awareness and improve research on disparities and improve access to care [7,8].

In this context, the association between race and outcomes in MDS is understudied and not well characterized, both in terms of direction and magnitude. To our knowledge, three small retrospective single-center studies in the US have reported survival outcomes according to race with inconsistent results [9–11]. In a large retrospective cohort study from the Veterans administration population, no differences in survival were observed by race in a mostly male elderly population, although a trend in favor of Black patients was reported [12]. A study from the Surveillance, Epidemiology, and End Results (SEER) Program focusing on patients with MDS diagnosed between 2004 and 2016 compared disease characteristics and survival outcomes between patients who were non-Hispanic Black, Hispanic, and non-Hispanic White and found that the former group had longer overall survival (OS) than patients who were Hispanic and non-Hispanic White, even after adjusting for confounding variables [13].

Using the SEER cancer registry, we focused our study on the comparison of demographics, clinical characteristics, sociodemographic factors, and survival outcomes between Black and White patients with MDS. We expanded on prior work by looking at the consistency of the association between race and OS by MDS histology subset and at different time points to better understand the variations in outcomes between races. We also characterized and compared the most frequent causes of death between Black and White patients.

## 2. Methods

### 2.1. Data source

The SEER Program from the National Institutes of Health collects data on cancer incidence, survival, and mortality for 26.5 % of the US population (Black race, 25.5 %; White race, 23.6 %) [14,15] and is considered the gold standard for data quality among cancer registries in the US and globally. We used the SEER November 2020 submission to identify incident MDS cases reported to the SEER Program between 2001 and 2013 [16]. Cancer data are updated annually to capture vital status, survival time, and cause of death with a 90 % or greater completeness standard for malignant primary tumors [15].

### 2.2. Population

We identified patients with a microscopically confirmed MDS diagnosis reported to SEER between 2001 and 2013. We chose the starting date of 2001 as it represented the date when MDS was classified as a malignancy according to the World Health Organization (WHO) and became reportable to cancer registries. The data cutoff of December 2018 was chosen to allow a minimum 5 years of follow-up. We extracted cases based on International Classification of Diseases for Oncology version 3 (ICD-O-3) codes as reported to SEER as follows: refractory anemia (RA; 9980), refractory anemia with ring sideroblasts (RARS; 9982), refractory anemia with excess blasts (RAEB; 9983), refractory cytopenia with multilineage dysplasia (RCMD; 9985), myelodysplastic syndrome associated with isolated del(5q) (MDS del[5q]; 9986), myelodysplastic syndrome, unclassifiable (MDS-U; 9989), refractory neutropenia (RN; 9991), and refractory thrombocytopenia (RT; 9992). Of note, the ICD-O-3 category of MDS-U indicates that SEER had insufficient data to further subclassify, which is distinct from the WHO category of MDS-U. In this report, the SEER MDS-U subset will be referred to

as MDS—not otherwise specified (MDS-NOS) to describe the unspecified nature of MDS-U [17] and its distinction from the WHO diagnostic group [18].

Patients were included if they identified as Black or White race (based on SEER race recode) irrespective of Hispanic ethnicity, were  $\geq 20$  years of age at time of MDS diagnosis, and if MDS was the first cancer diagnosed in SEER. We excluded patients with therapy-related MDS and MDS in transformation, those with zero days of follow-up or zero days of survival, and those with diagnosis based on autopsy or death certificate only.

### 2.3. Measures

Other variables from SEER used to categorize patients at baseline were age as a categorical variable (20–44, 45–54, 55–64, 65–74, 75–84, or 85 + years), age as a continuous variable, sex (male or female), marital status (single, married, separated/divorced/widowed, or other/unknown), urban-rural continuum code representing metropolitan residential status (large metropolitan area defined as  $>1$  million inhabitants, small metropolitan area as inhabitants  $\leq 1$  million, or nonmetropolitan areas), median household income of the patient county of residence ( $< \$75,000/\text{year}$  or  $\geq \$75,000/\text{year}$ ), and year of diagnosis as a categorical variable (2001–2004, 2005–2013).

Other variables extracted were vital status recode (alive or dead) and cause of death recode for the various attributed causes of death that were then grouped according to common etiologies as described by Brunner et al. [19]. ICD-O-3 subtypes of MDS were grouped based on OS into lower risk (RA, RT, RN, or RARS), intermediate risk (RCMD, MDS del[5q], MDS-NOS), and higher-risk disease (RAEB) [20]. Owing to very small sample sizes for RN ( $n = 2$ ) and RT ( $n = 81$ ), these are not reported individually.

### 2.4. Definition of race

For our analyses, we modeled race in two different manners. First, we classified all patients with MDS as either White or Black race as per SEER original coding. We also defined three possible values for race based on a matching procedure between Black and White patients (see statistical analysis below).

### 2.5. Statistical analysis

To compare baseline characteristics between Black and White cohorts, we used the chi-square test for categorical covariates and the Wilcoxon rank-sum test for continuous covariates. Age-adjusted incidence was obtained directly from SEER dataset using the incidence-session and access via SEER\*Stat software on February 15, 2022 [21].

The Kaplan-Meier estimate was used to compare unadjusted survival estimates (model 1) from the time of diagnosis until death or censoring between the Black and White cohorts and to evaluate the proportional hazard of death between Black and White race and for all other baseline characteristic covariates (not reported) over time. The log-rank test was used to compare cohorts. All hazard ratios (HR) were calculated for Black patients compared with White patients (reference cohort).

Given the design of our study as a retrospectively cohort study, we anticipated that significant differences in presentation at diagnosis would be present between the Black and White cohorts. Therefore, we performed two analyses to reduce the impact of potential confounding factors due to imbalance in baseline covariates.

First, we performed multivariable analysis (model 2) controlling for all baseline covariates. To analytically address the question of missing data for 9.6 % of marital status values and to address the interpretability of the results due to the high proportion of MDS-NOS, we performed a data simulation via multiple imputation ( $N = 30$ , multiple imputation by predictive mean matching - see Appendix B). The first simulation only imputed missing marital status values and the second simulation

imputed both marital status and MDS-NOS histology. We used these imputed datasets to perform the multivariable analysis and for any analysis including these two variables. Second, we implemented a one-to-one matching between each Black patient and a corresponding White patient according to the following criteria: sex, age (increment of 5 years from 20 to 84 years, then  $\geq 85$  years), MDS histology, and urban-rural code. Given that the Black cohort represented a small percentage (8.1 %) of the study population, our goal was to match as many Black patients as possible to similar White patients. One-to-one matching was implemented based on the coarsened exact matching procedure that has been described previously [22] (see Appendix B in supplemental for details). We used this matching procedure to derive a new code for race with three possible values: one value for Black and White matched patients and one value for all other unmatched patients and used this new race recode in our third model (model 3) and in subgroup analyses.

Based on the known association between age, sex, MDS histology, and socioeconomic factors with health [8] and survival, we also examined the robustness of the association between race and survival with baseline covariates (sex, age, MDS histology, marital status, urban-rural code, income group, and year of diagnosis period) via interaction testing. We reported the HR and  $p$  value from the Cox proportional model for the survival probability difference between Black and White cohorts for the MDS histology subset, as it was the main objective of our study and for any subgroups where we detected a statistically significant interaction.

We also investigated if the association of race and OS varied according to periods of diagnosis analyzed in two periods (2001–2004 and 2005–2013) or four periods (2001–2004, 2005–2006, 2007–2009, and 2010–2013).

We calculated descriptive statistics for the frequency of the attributed causes of death overall and between the Black and White cohorts.

All analyses were done using Stata 17.0 BE (StataCorp LLC; College Station, TX). Statistical significance for all two-sided tests was declared for  $p < 0.05$ .

This study was considered exempt by the Dana-Farber Harvard Cancer Center institutional review board.

### 3. Results

#### 3.1. Population and incidence rate

A total of 37,562 patients with MDS (Table 1) were identified with a median age of diagnosis of 76 years (interquartile range [IQR], 67–82). Of these, 81.1 % of patients were diagnosed after 65 years, 57.3 % were male, 8.1 % identified as Black, and 91.9 % identified as White. There was a greater percentage of female patients among the Black cohort (50.8 %) compared with the White cohort (42.0 %). Median age was younger among Black patients (71 years; IQR, 61–79) compared with White patients (76 years; IQR, 68–83). Overall, the age-adjusted MDS incidence rate between 2001 and 2013 for our cohort of adult patients

**Table 1**  
Patient demographics and disease characteristics at diagnosis.

|                                 | Overall |       | Black patients |      | White patients |      | $p$ value <sup>a</sup> |
|---------------------------------|---------|-------|----------------|------|----------------|------|------------------------|
|                                 | N       | %     | N              | %    | N              | %    |                        |
| Total                           | 37,562  | 100.0 | 3037           | 8.1  | 34,525         | 91.9 |                        |
| Sex                             |         |       |                |      |                |      | < 0.001                |
| Male                            | 21,515  | 57.3  | 1495           | 49.2 | 20,020         | 58.0 |                        |
| Female                          | 16,047  | 42.7  | 1542           | 50.8 | 14,505         | 42.0 |                        |
| Age group (years)               |         |       |                |      |                |      | < 0.001                |
| 20–44                           | 834     | 2.2   | 165            | 5.4  | 669            | 1.9  |                        |
| 45–54                           | 1711    | 4.6   | 258            | 8.5  | 1453           | 4.2  |                        |
| 55–64                           | 4559    | 12.1  | 555            | 18.3 | 4004           | 11.6 |                        |
| 65–74                           | 9622    | 25.6  | 849            | 28.0 | 8773           | 25.4 |                        |
| 75–84                           | 14,301  | 38.1  | 850            | 28.0 | 13,451         | 39.0 |                        |
| 85 +                            | 6535    | 17.4  | 360            | 11.9 | 6175           | 17.9 |                        |
| Marital status                  |         |       |                |      |                |      | < 0.001 <sup>b</sup>   |
| Never married                   | 3288    | 8.8   | 581            | 19.1 | 2707           | 7.8  |                        |
| Married                         | 19,536  | 52.0  | 1116           | 36.7 | 18,420         | 53.4 |                        |
| Separated/divorced/widowed      | 11,139  | 29.7  | 1064           | 35.0 | 10,075         | 29.2 |                        |
| Other/unknown                   | 3599    | 9.6   | 276            | 9.1  | 3323           | 9.6  |                        |
| Metropolitan residential status |         |       |                |      |                |      | < 0.001                |
| Large metropolitan area         | 22,836  | 60.8  | 2031           | 66.9 | 20,805         | 60.3 |                        |
| Small metropolitan area         | 10,078  | 26.8  | 694            | 22.9 | 9384           | 27.2 |                        |
| Nonmetropolitan area            | 4648    | 12.4  | 312            | 10.3 | 4336           | 12.6 |                        |
| Income bracket (US dollars)     |         |       |                |      |                |      | < 0.001                |
| < 75,000                        | 25,781  | 68.6  | 2453           | 80.8 | 23,328         | 67.6 |                        |
| $\geq 75,000$                   | 11,781  | 31.4  | 584            | 19.2 | 11,197         | 32.4 |                        |
| Histology                       |         |       |                |      |                |      | < 0.001                |
| RA                              | 4461    | 11.9  | 431            | 14.2 | 4030           | 11.7 |                        |
| RARS                            | 3669    | 9.8   | 268            | 8.8  | 3401           | 9.9  |                        |
| RAEB                            | 6077    | 16.2  | 442            | 14.6 | 5635           | 16.3 |                        |
| RCMD                            | 2693    | 7.2   | 198            | 6.5  | 2495           | 7.2  |                        |
| MDS del(5q)                     | 1148    | 3.1   | 61             | 2.0  | 1087           | 3.1  |                        |
| MDS-NOS                         | 19,514  | 52.0  | 1637           | 53.9 | 17,877         | 51.8 |                        |
| Prognostic risk group           |         |       |                |      |                |      | 0.016                  |
| Lower                           | 8130    | 21.6  | 699            | 23.0 | 7431           | 21.5 |                        |
| Intermediate                    | 23,355  | 62.2  | 1896           | 62.4 | 21,459         | 62.2 |                        |
| Higher                          | 6077    | 16.2  | 442            | 14.6 | 5635           | 16.3 |                        |
| Year of diagnosis               |         |       |                |      |                |      | 0.17                   |
| 2001–2004                       | 9554    | 25.4  | 804            | 26.5 | 8750           | 25.3 |                        |
| 2005–2013                       | 28,008  | 74.6  | 2233           | 73.5 | 25,775         | 74.7 |                        |

MDS del(5q), myelodysplastic syndrome associated with isolated del(5q); MDS-NOS, myelodysplastic syndrome—not otherwise specified; RA, refractory anemia; RAEB, refractory anemia with excess blasts; RARS, refractory anemia with ring sideroblasts; RCMD, refractory cytopenia with multilineage dysplasia; US, United States.

<sup>a</sup> Chi-squared test for categorical variables.

<sup>b</sup> Chi-squared test calculated from multiple imputed dataset.

with MDS was 6.7/100,000 and was statistically lower for Black patients (5.8/100,000) compared with White patients (7.0/100,000; [Supplemental Table S1](#)) for all MDS subsets except for RA. The lower age-adjusted MDS incidence for Black patients was observed during every year of the entire study period.

At diagnosis, 36.7 % of Black patients were married compared with 53.4 % of White patients ([Table 1](#)). More Black patients lived in large metropolitan areas (66.9 %) and in counties with lower median household incomes (<\$75,000/year [80.8 %]) compared with White patients (60.3 % and 67.6 %, respectively). In terms of histology, the percentages of Black patients with ICD-O-3 coding for RA, RAEB, and MDS-NOS were 14.2 %, 14.6 %, and 53.9 %, respectively, compared with 11.7 %, 16.3 %, 51.8 % for White patients. The proportion of Black patients diagnosed as RA and MDS-NOS were statistically higher compared with White patients (14.2 % vs 11.7 %;  $p < 0.001$  % and 53.9 % vs 51.8 %;  $p = 0.025$ , respectively). The proportion of Black patients with RAEB, the higher-risk disease category, was lower compared with White patients (14.6 % vs 16.3 %;  $p = 0.011$ ).

3.2. Constitution of the matched dataset

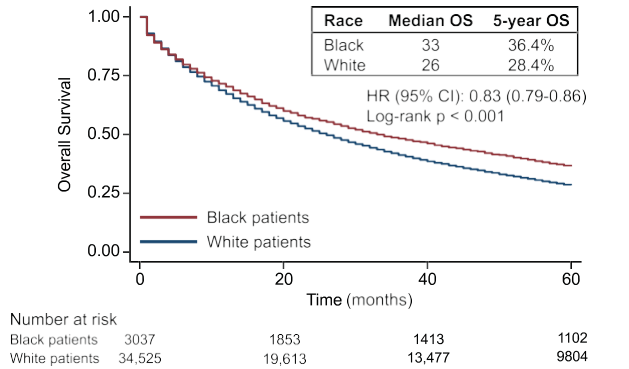
The matching procedure matched 99.6 % (3026/3037) of Black patients in a one-to-one manner with White patients. A total of 6052 patients constituted the matched set, with 4854 death events. A table of baseline characteristics after matching is in [Supplemental Table S2](#).

3.3. Univariate OS analysis

Median follow-up was 132 months, with 86.6 % of patients having died of any cause, split between Black and White patients at 79.0 % and 87.2 %, respectively ([Supplemental Table S3](#)). The univariate analysis of OS by race ([Table 2](#)) showed a statistically significant association between race and OS favoring Black patients after 3-year follow-up (HR [95 % confidence interval (CI)], 0.84 [0.80–0.89];  $p < 0.001$ ), after 5-year follow-up (0.83 [0.79–0.86];  $p < 0.001$ ; [Fig. 1](#)), and for the overall study period (0.79 [0.76–0.83];  $p < 0.001$ ). Median OS was 33 months (95 % CI, 31–37) for Black patients and 26 months (95 % CI, 26–27) for White patients ([Supplemental Table S4](#)). The 3-year and 5-year OS rates were 48.3 % and 36.4 %, respectively, for Black patients and 41.2 % and 28.4 % for White patients.

3.4. Analyses to adjust for confounding factors

The multivariable analysis of OS in the multiply imputed dataset (model 2; [Table 2](#)) showed an attenuated association between Black race and OS that remained statistically significant with a 3-year and 5-year HR of 0.92 (95 % CI, 0.87–0.97;  $p = 0.003$ ) and 0.92 (95 % CI,



**Fig. 1.** OS of patients with MDS diagnosed between 2001 and 2013 (n = 37,562). HR, hazard ratio; MDS, myelodysplastic syndrome; OS, overall survival.

0.87–0.96;  $p < 0.001$ ). In our third model, where we compared the survival between the matched Black and White patients while controlling for all other unmatched patients, as well as for income and marital status since they were not included in the matching procedure, the association between race and OS was not statistically significant at 3-year follow-up (HR, 0.94; 95 % CI, 0.87–1.00;  $p = 0.08$ ) but reached significance at 5-year follow-up (HR, 0.92; 95 % CI, 0.87–0.98;  $p = 0.01$ ).

3.5. Analysis of cohort effect

We did not detect a cohort effect as measured by interaction testing between race and OS with periods of diagnosis defined as either two periods (2001–2004, 2005–2013) or four periods (2001–2004, 2005–2006, 2007–2009, 2010–2013),  $\chi^2 = 0.8$  ( $p = 0.37$ ) and  $\chi^2 = 2.5$  ( $p = 0.06$ ), respectively.

3.6. Survival differences between geographic areas

We detected an interaction between the association of race and OS with metropolitan residential status. A subgroup analysis ([Table 3](#)) showed that the strongest and only statistically significant association was detected for patients living in nonmetropolitan areas (HR, 0.74; 95 % CI, 0.63–0.86;  $p < 0.001$ ).

3.7. Survival differences between specific MDS subtypes

When we analyzed how the association between race and survival varied by histology, we observed heterogeneity between MDS subgroups. We found no association between race and survival for RCMD

**Table 2**  
Association of race and OS for patients with MDS from Cox regression models (n = 37,562).

|                      |                   | Outcome: overall survival <sup>a</sup> |                           |                      |                     |                                 |                      |
|----------------------|-------------------|----------------------------------------|---------------------------|----------------------|---------------------|---------------------------------|----------------------|
|                      |                   | Three-year follow-up                   |                           |                      | Five-year follow-up |                                 |                      |
|                      | Black vs White    | Events                                 | HR <sup>b</sup> (95 % CI) | p value <sup>b</sup> | Events              | HR <sup>b</sup> (95 % CI)       | p value <sup>b</sup> |
| Model 1 <sup>c</sup> | Race per SEER     | 21,805                                 | 0.84 (0.80–0.89)          | < 0.001              | 26,578              | 0.83 (0.79–0.86)                | < 0.001              |
| Model 2 <sup>d</sup> | Race per SEER     | 21,805                                 | 0.92 (0.87–0.97)          | 0.003                | 26,578              | 0.92 <sup>b,c</sup> (0.87–0.96) | < 0.001              |
| Model 3 <sup>f</sup> | Race per matching | 3171                                   | 0.94 (0.87–1.00)          | 0.08                 | 3917                | 0.92 (0.87–0.98)                | 0.01                 |

HR, hazard ratio; CI, confidence interval; MDS, myelodysplastic syndrome; SEER, Surveillance, Epidemiology, and End Results.

<sup>a</sup> Univariate HR (95 %CI) for the entire study period is: 0.79 (0.76–0.83;  $p < 0.001$ ).

<sup>b</sup> p values are from Cox proportional hazard regression model. A hazard ratio < 1 (> 1) corresponds to a lower (higher) value for Black compared with White patients.

<sup>c</sup> In model 1, the functional form of race is coded Black vs White, as defined by SEER. Model is unadjusted and runs in observed dataset.

<sup>d</sup> In model 2, the functional form of race is coded Black vs White, as defined by SEER. Model is adjusted for all baseline covariates (sex, age group, histology, marital status, metropolitan residential status, income group, year of diagnosis period). Model runs in imputed dataset for both marital status and MDS–NOS.

<sup>e</sup> HR with age as continuous variable, 0.93; 95 % CI, 0.88–0.97;  $p = 0.002$ . HR with marital status variable only imputed, 0.92; 95 % CI, 0.88–0.96;  $p = 0.001$ .

<sup>f</sup> In model 3, the functional form of race is coded as Black-matched vs White-matched, while controlling for all other remaining unmatched patients and income and marital status covariates. HR for unmatched patients is 1.16 (95 % CI, 1.10–1.21);  $p < 0.001$ ). Count of events are for matched patients. Model runs in imputed dataset.

**Table 3**

Cox regression-based associations of Black versus White race and overall survival by SEER-defined metropolitan residential area subgroups (n = 37,562).

| Metropolitan Residential Area subgroup <sup>b</sup> | Overall Survival <sup>a</sup> |                           |                      |
|-----------------------------------------------------|-------------------------------|---------------------------|----------------------|
|                                                     | Frequency <sup>c</sup>        | HR <sup>d</sup> (95 % CI) | p value <sup>d</sup> |
| Large metropolitan area                             | 66.9 %                        | 0.94 (0.88–1.01)          | 0.10                 |
| Small metropolitan area                             | 22.8 %                        | 0.94 (0.85–1.04)          | 0.26                 |
| Nonmetropolitan area                                | 10.2 %                        | 0.74 (0.63–0.86)          | < 0.001              |
| $\chi^2_2$                                          |                               | 9.59                      | 0.01                 |

CI, confidence interval; HR, hazard ratio; PH, proportional hazard; SEER, Surveillance, Epidemiology, and End Results;  $\chi^2_{2df}$ : Chi-squared test, degree of freedom.<sup>a</sup> Functional form of race is coded as Black-matched versus White-matched, while controlling for all other remaining unmatched patients and income and marital status covariates. HR shown are for matched patients. Model run in observed dataset. 5-year follow-up.<sup>b</sup> In a preliminary model, a statistically significant interaction between race and overall survival with metropolitan residential status was detected ( $\chi^2_2 = 10.4$ ;  $p = 0.005$ ).<sup>c</sup> The frequency percentages are for matched patients (n = 6052).<sup>d</sup> p values are from Cox PH regression model. A HR < 1 (>1) corresponds to a lower (higher) value for Black compared with White MDS patients.

and MDS del(5q). Using observed data, RA, RARS, and MDS-NOS showed a statistically significant association favoring Black patients with the following HRs: 0.64 (95 % CI, 0.56–0.74;  $p < 0.001$ ), 0.69 (95 % CI, 0.58–0.82;  $p < 0.001$ ), and 0.84 (95 % CI, 0.78–0.91;  $p < 0.001$ ). RAEB favored White patients in both models with HR of 1.25 (95 % CI, 1.13–1.38;  $p < 0.001$ ) in the analysis using the imputed data. Significant differences for RA, RARS, and RAEB remained in the imputed data set (Table 4).

### 3.8. Frequency of attributed cause of death

We evaluated how deaths were attributed within the SEER MDS population (Supplemental Table S3). Death attributed to MDS/leukemia represented the most frequent cause of death overall (46.6 %) and was higher for White patients (47.2 %) compared with Black patients (39.3 %). Death attributed to cardiovascular disease (CVD) was the second most common overall (18.0 %) and was more common among Black versus White patients (19.6 % vs 17.8 %). The third most frequent cause of death was that due to other malignancies (10.9 %) and was 12.4 % for Black patients and 10.8 % for White patients.

## 4. Discussion

We conducted a large retrospective analysis of the SEER cancer registry to characterize and compare baseline demographics, clinical characteristics, sociodemographic factors, and survival outcomes between Black and White patients with MDS in the US. The univariate OS

analysis showed that Black race was associated with longer survival for patients with MDS during the study period (HR, 0.83; 95 % CI, 0.79–0.86;  $p < 0.001$ ; Fig. 1). This was unexpected given the absence to our knowledge of biological rationale for difference in outcome between race for patients with MDS and given the existing literature about racial disparities in cancer care and outcomes between Black and White patients that consistently highlights poorer outcomes for Black patients. In the multivariable analysis to adjust for differences in baseline characteristics between both cohorts, results were attenuated (5-year OS HR, 0.92; 95 % CI, 0.87–0.96;  $p < 0.001$ ) although still significant and results were also significant in the adjusted analysis using the matched patients. Nonetheless, these results need to be interpreted with caution given the lack of comprehensive clinical, cytogenetic, prognostic, molecular, therapy annotation in SEER that prevented us from including these variables in our analysis. Although previous studies have reported inconclusive results about the association between race and OS outcomes in MDS [9–11], our study builds and expands on prior work reporting reported a trend toward an association between race and survival outcomes in favor of Black patients [10–13]. The key finding of our study is that we found meaningful variation in both the direction and the magnitude of the association of race and OS between various MDS histology. We observed the strongest association favoring Black patients in the lower-risk (RA, RARS) subsets and in the MDS-NOS group, a category of patients with an MDS diagnosis but for whom there were insufficient data to be better classified in SEER. When looking at the higher-risk RAEB group, the results favored White patients, which is consistent with results previously reported in AML [3] and with the

**Table 4**

Cox regression-based associations of Black versus White race and overall survival by SEER-defined MDS histology subgroups (n = 37,562).

| MDS subgroups <sup>c</sup> | Overall Survival           |                           |                      |                                |                           |                      |
|----------------------------|----------------------------|---------------------------|----------------------|--------------------------------|---------------------------|----------------------|
|                            | Observed data <sup>a</sup> |                           |                      | Imputed dataset <sup>a,b</sup> |                           |                      |
|                            | Frequency <sup>d</sup>     | HR <sup>e</sup> (95 % CI) | p value <sup>e</sup> | Frequency <sup>d</sup> (SE)    | HR <sup>e</sup> (95 % CI) | p value <sup>e</sup> |
| RA                         | 14.2 %                     | 0.64 (0.56–0.74)          | < 0.001              | 28.8 % (0.8)                   | 0.75 (0.67–0.83)          | < 0.001              |
| RARS                       | 8.8 %                      | 0.69 (0.58–0.82)          | < 0.001              | 19.2 % (0.6)                   | 0.80 (0.71–0.92)          | 0.001                |
| RAEB                       | 14.6 %                     | 2.12 (1.90–2.37)          | < 0.001              | 32.7 % (0.9)                   | 1.25 (1.13–1.38)          | < 0.001              |
| RCMD                       | 6.4 %                      | 0.94 (0.77–1.43)          | 0.57                 | 14.4 % (0.6)                   | 0.88 (0.76–1.03)          | 0.11                 |
| MDS del(5q)                | 1.9 %                      | 1.05 (0.77–1.43)          | 0.76                 | 4.9 % (0.4)                    | 0.97 (0.74–1.25)          | 0.79                 |
| MDS-NOS                    | 54.1 %                     | 0.84 (0.78–0.91)          | < 0.001              |                                |                           |                      |
| $\chi^2_5$                 |                            | 308.9                     | < 0.001              | $\chi^2_4$                     | 15.9                      | < 0.001              |

HR, hazard ratio; CI, confidence interval; MDS, myelodysplastic syndrome; MDS del(5q), myelodysplastic syndrome associated with isolated del(5q); MDS-NOS, myelodysplastic syndrome—not otherwise specified; OS, overall survival; PH, proportional hazard; RA, refractory anemia; RAEB, refractory anemia with excess blasts; RARS, refractory anemia with ring sideroblasts; RCMD, refractory cytopenia with multilineage dysplasia; SEER, Surveillance, Epidemiology, and End Results;  $\chi^2_{2df}$ : chi-squared test, degree of freedom.<sup>a</sup> Functional form of race is coded as Black-matched versus White-matched, while controlling for all other remaining unmatched patients and income and marital status covariates. HR shown are for matched patients. 5-year follow-up.<sup>b</sup> Multiply imputed dataset where missing marital status data and the poorly defined MDS-NOS subset were simulated.<sup>c</sup> In a preliminary model, a statistically significant interaction between race and OS with MDS histology was detected ( $\chi^2_5 = 11.3$ ;  $p = 0.045$ ).

<sup>d</sup> The frequency percentages are for matched patients (n = 6052).

<sup>e</sup> *p* values are from Cox PH regression model. A HR < 1 (>1) corresponds to a lower (higher) value for Black compared with White MDS patients

disease continuum between high-risk MDS and AML. For RARS, which is a relatively well-defined subgroup (associated with and enriched for *SF3B1* mutations [23]), results favored Black patients. We believe that our study is the first to perform a more in-depth analysis of how the association between survival and race varies by MDS histology, as reported to SEER.

The findings in our study should be interpreted in the context of the detail of case capture in a larger database such as SEER. MDS is a highly heterogeneous clinical, morphologic, and molecular entity [24]. The diagnosis of MDS can be challenging, particularly for cases of lower-grade MDS (without excess blasts) [1]. MDS diagnosis often requires distinguishing from many alternative neoplastic, pre-MDS and non-neoplastic disorders that can mimic MDS [25–28]. This is important for disease surveillance, as registries depend on accurate source data for case inclusion. Inter-observer variability has been observed in all subtypes and especially in the low-grade forms of MDS [29]. In a large study from MD Anderson Cancer Center, a 12 % disagreement rate was observed between diagnosis at referral and the one performed at the tertiary center, with most cases being reclassified as higher-risk disease [30].

In addition, the relatively high frequency of patients with disease classified as MDS-NOS (52.0 %) could indicate that a large proportion of cases remained with an initial provisional diagnosis. A higher proportion of Black patients was diagnosed as MDS-NOS in SEER, which may introduce bias in diagnosis and in our analyses. The percentage of MDS-NOS has been relatively stable over time in SEER, ranging from 48.1 % to 54.6 % over the study period. However, we observed large variability in the proportion of MDS-NOS in the various regional cancer registries included in SEER, ranging from 31.8 % in Seattle to 60.1 % in New Jersey (Supplemental Table S5). Furthermore, the SEER designation of MDS-NOS and the WHO definition for MDS-U differ and do not fully overlap. Per SEER, MDS-NOS corresponds to “myelodysplastic syndrome, NOS, a generic disease description. NOS histology is only the provisional diagnosis; the physician will run further diagnostic procedures and look for various clinical presentations to identify a more specific disease” [17]. A study from the Mayo Clinic on 1000 primary MDS cases from 1989 to 2014 reported the percentage of MDS-U declining from 28 % to 4 % over that time [31], reflecting advances in diagnostic criteria and accuracy. Maassen et al. [32] also reported that MDS-U represented 6 % of patients in the Dusseldorf MDS Registry per 2008 WHO. Thus, it is difficult to assess the true proportion of MDS-U per the WHO definition in our SEER dataset.

Given the challenging diagnostic that may have impacted the MDS-NOS population in SEER and the limitations associated with the SEER dataset, it is not possible to pinpoint one overarching reason why survival outcomes favored Black patients. One possibility is that the entity of benign ethnic neutropenia, a condition seen among Black patients, may be inadvertently included in the MDS-NOS population given the criteria of subjective interpretation for inclusion. A subsequent impact of such an inclusion would not be surprising, given that benign ethnic neutropenia is associated with low absolute neutrophil count but no increased risk of infections [33–36]. Regarding RA, several groups have reported lower hemoglobin concentrations in the Black population compared with White; one can speculate if this could be a source of bias resulting in the classification as RA of Black patients with laboratory values below reference ranges but otherwise favorable prognosis [37, 38].

We also observed significant difference in demographics, socioeconomic factors, and clinical characteristics between Black and White patients (Table 1). Regarding baseline demographics, the lower MDS age-adjusted incidence among Black patients (except for RA) has also been observed in other hematologic diseases such as AML, chronic lymphocytic leukemia, chronic myeloid leukemia, and chronic myelomonocytic leukemia [14]. To our knowledge, the cause of these phenomena is unclear. In regard to the higher proportion of males in the White cohort, a review by Karantanos et al. [39] summarized evidence

supporting sex-related differences in the genomic landscape of patients with myeloid neoplasms, indicating that men have higher incidence of mutations in high-risk genes such as *ZRSR2*, *U2AF1*, *SRSF2*, and *ASXL1*. Others reported differential response and outcomes based on sex [40]. Such variability could impact our results, although the methods used in our study for controlling differences in baseline characteristics should have accounted for this.

The higher proportion of MDS cases diagnosed in the Black cohort in the ill-defined category of MDS-NOS and in MDS-RA may have contributed to our finding. However, even we attempted to remove the impact of MDS-NOS via multiple imputation, we observed significant statistical association with three remaining MDS subsets (Table 4). Therefore, the observation of longer survival among Black patients with MDS in SEER may theoretically reflect earlier detection of truly malignant cancers, the inclusion of conditions with limited clinical impact due to bias in case ascertainment and diagnosis, or a poorly understood difference in biology between Black and White patients with MDS. Until US registries can address the potential impact of misclassification of subtypes in MDS, it will be challenging to fully address this point. Furthermore, the evolving diagnostic criteria of MDS and AML, such as the recent change in the European Leukemia Net 2022 AML classification [41] introducing a new category of MDS/AML for cases of 10 %–19 % blasts may impact the interpretability of our results and the comparison between studies, especially for HR MDS.

In a prior study, Brunner et al. [19] reported the relative importance of MDS/leukemia-related and non-cancer-related causes of death for patients with MDS, such as CVD. They reported important findings demonstrating that CVD as the attributed cause of death increased as a function of increased survival time and the overall elevated CVD risk compared with the general population [19]. Our study confirms the high frequency of CVD as the second most frequently attributed cause of death overall and for Black and White patients (19.6 % and 17.8 %, respectively; Supplemental Table S3). Therefore, even though Black patients with MDS were observed to have an overall favorable OS outcome compared with White patients with MDS, the higher frequency of CVD-related deaths in the Black population [42] supports the need for continued screening and appropriate management of this increased risk.

In summary, in this large analysis of MDS using data reported to SEER, we report an association between race and OS favoring Black patients. Given the limitation of the source data, these results need to be interpreted with caution. We show potential significant variation, in terms of MDS subtypes and association with outcomes, with RA, RARS, and MDS-NOS favoring Black patients and RAEB favoring White patients. These underline the need to minimize and better understand the contribution of the poorly defined MDS-NOS subset of MDS in the SEER cancer registry. We also confirm the high frequency of CVD as an attributed cause of death in SEER and its continued importance in the clinical management of patients with MDS.

## Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Editorial assistance was provided by The Lockwood Group (Stamford, CT, US) and was supported with private funding by the first author.

## CRedit authorship contribution statement

**Arnaud Lesegretain:** Conceptualization, Methodology, Data curation, Formal analysis, Software, Visualization, Writing – original draft, Writing – review & editing. **Andrew Brunner:** Conceptualization, Writing – review & editing. **Andrew J. King:** Supervision, Conceptualization, Methodology, Writing – review & editing. **Abderrahmane Laadem:** Visualization, Writing – review & editing. **Geoffrey Fell:** Methodology, Conceptualization. **Amir T. Fathi:** Conceptualization, Writing – review & editing, Visualization, Supervision.

## Previous congress disclosures

Lesegretain A, Laadem A, Fell G, Fathi AT. Comparison of demographics, disease characteristics, and outcomes between African American patients and White patients with myelodysplastic syndrome: A population-based study, *J. Clin. Oncol.* 40(16\_suppl) (2022) 7051.

Lesegretain A, Laadem A, Fell G, Fathi AT. Comparison of demographics, disease characteristics and outcomes between black and white patients with myelodysplastic syndrome: a population-based study. Presented at the Society of Hematologic Oncology Annual Meeting in Houston, TX, USA; September 28–October 1, 2022.

## Authors' disclosure of potential conflicts of interest

No potential conflicts of interest relevant to this article were reported.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.leukres.2022.107006.

## References

- [1] A.J. Siddon, R.P. Hasserjian, How I diagnose low-grade myelodysplastic syndromes, *Am. J. Clin. Pathol.* 154 (1) (2020) 5–14.
- [2] A.M. Zeidan, R.M. Shallis, R. Wang, A. Davidoff, X. Ma, Epidemiology of myelodysplastic syndromes: why characterizing the beast is a prerequisite to taming it, *Blood Rev.* 34 (2019) 1–15.
- [3] B. Bhatnagar, J. Kohlschmidt, K. Mro'zek, Q. Zhao, J.L. Fisher, D. Nicolet, C. J. Walker, A.S. Mims, C. Oakes, B. Giacomelli, S. Orwick, I. Boateng, J.S. Blachly, S. E. Maharry, A.J. Carroll, B.L. Powell, J.E. Kolitz, R.M. Stone, J.C. Byrd, E. D. Paskett, A. de la Chapelle, R. Garzon, A.K. Eisfeld, Poor survival and differential impact of genetic features of Black patients with acute myeloid leukemia, *Cancer Disco* 11 (3) (2021) 626–637.
- [4] K. Kirtane, S.J. Lee, Racial and ethnic disparities in hematologic malignancies, *Blood* 130 (15) (2017) 1699–1705.
- [5] L.A. Siminoff, L. Ross, Access and equity to cancer care in the USA: a review and assessment, *Postgrad. Med. J.* 81 (961) (2005) 674–679.
- [6] American Cancer Society, Cancer Facts & Figures for African Americans 2019–2021, 2019. (<https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/cancer-facts-and-figures-for-african-americans/cancer-facts-and-figures-for-african-americans-2019–2021.pdf>). (Accessed August 8, 2022).
- [7] American Society of Hematology, Message from the ASH President on Diversity, Equity, and Inclusion., 2020. (<https://www.hematology.org/newsroom/press-releases/2020/message-from-the-ash-president-on-diversity-equity-and-inclusion/>). (Accessed September 2, 2020).
- [8] E. Goss, A.M. Lopez, C.L. Brown, D.S. Wollins, O.W. Brawley, D. Raghavan, American society of clinical oncology policy statement: disparities in cancer care, *J. Clin. Oncol.* 27 (17) (2009) 2881–2885.
- [9] A. Nazha, K. Al-Issa, B. Przychodzen, N. Abuhadra, C. Hirsch, J.P. Maciejewski, M. A. Sekeres, Differences in genomic patterns and clinical outcomes between African-American and white patients with myelodysplastic syndromes, *Blood Cancer J.* 7 (9) (2017), e602.
- [10] A. Sridharan, R. Jain, M.A. Bachhuber, Y. Yu, K. Ramesh, K. Gundabolu, E. W. Friedman, A.K. Verma, Epidemiologic study of myelodysplastic syndromes in a multiethnic, inner city cohort, *Exp. Hematol. Oncol.* 3 (2014) 22.
- [11] D.P. Zandberg, F. Hendrick, E. Vannorsdall, J. Bierenbaum, M.L. Tidwell, Y. Ning, X.F. Zhao, A.J. Davidoff, M.R. Baer, Tertiary center referral patterns for patients with myelodysplastic syndrome are indicative of age and race disparities: a single-institution experience, *Leuk. Lymphoma* 54 (2) (2013) 304–309.
- [12] R.S. Komrokji, G.M. Matacia-Murphy, N.H. Al Ali, M.S. Beg, M.M. Safa, D. E. Rollison, A.F. List, Outcome of patients with myelodysplastic syndromes in the Veterans Administration population, *Leuk. Res.* 34 (1) (2010) 59–62.
- [13] S.Y. Goksu, M. Ozer, B.B. Goksu, R. Wang, J. Khatib, P.A. Patel, M. Vusirikala, S. Cole, A. Seyhanli, R.H. Collins, S. Chung, A.M. Zeidan, Y.F. Madanat, The impact of race and ethnicity on outcomes of patients with myelodysplastic syndromes: a population-based analysis, *Leuk. Lymphoma* 63 (7) (2022) 1651–1659.
- [14] Surveillance, Epidemiology, and End Results (SEER) Program ([www.seer.cancer.gov](http://www.seer.cancer.gov)) SEER\*Stat Database, Number of Persons by Race and Hispanic Ethnicity for SEER Participants (2020 Census Data). (<https://seer.cancer.gov/registries/data.html>). (Accessed August 2, 2022).
- [15] M.A. Duggan, W.F. Anderson, S. Altekruse, L. Penberthy, M.E. Sherman, The surveillance, epidemiology, and end results (SEER) program and pathology: toward strengthening the critical relationship, *Am. J. Surg. Pathol.* 40 (12) (2016) e94–e102.
- [16] Surveillance, Epidemiology, and End Results (SEER) Program ([www.seer.cancer.gov](http://www.seer.cancer.gov)) SEER\*Stat Database, SEER Research Plus Data, 18 Registries, Nov 2020 Sub (2000–2018) - Linked To County Attributes - Time Dependent (1990–2018) Income/Rurality, 1969–2019 Counties. (Accessed August 2, 2022).
- [17] D. Ruhl, M. Adamo, L. Dickie, S. Negoita, Hematopoietic and Lymphoid Neoplasms Coding Manual, National Cancer Institute, Bethesda, MD, 2021.
- [18] D.A. Arber, A. Orazi, R. Hasserjian, J. Thiele, M.J. Borowitz, M.M. Le Beau, C. D. Bloomfield, M. Cazzola, J.W. Vardiman, The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia, *Blood* 127 (20) (2016) 2391–2405.
- [19] A.M. Brunner, T.M. Blonquist, G.S. Hobbs, P.C. Amrein, D.S. Neuberg, D. P. Steensma, G.A. Abel, A.T. Fathi, Risk and timing of cardiovascular death among patients with myelodysplastic syndromes, *Blood Adv.* 1 (23) (2017) 2032–2040.
- [20] L. Malcovati, M.G. Della Porta, C. Pascutto, R. Invernizzi, M. Boni, E. Travaglino, F. Passamonti, L. Arcaini, M. Maffioli, P. Bernasconi, M. Lazzarino, M. Cazzola, Prognostic factors and life expectancy in myelodysplastic syndromes classified according to WHO criteria: a basis for clinical decision making, *J. Clin. Oncol.* 23 (30) (2005) 7594–7603.
- [21] Surveillance Research Program, National Cancer Institute SEER\*Stat software ([seer.cancer.gov/seerstat](http://seer.cancer.gov/seerstat)) version 8.3.9.2. (Accessed February 15, 2022).
- [22] M. Blackwell, S. Iacus, G. King, G. Porro, Cem: coarsened exact matching in Stata, *Stata J.* 9 (4) (2009) 524–546.
- [23] L. Malcovati, M. Karimi, E. Papaemmanuil, I. Ambaglio, M. Ja'rdenstein, M. Jansson, C. Elena, A. Galli, G. Walldin, M.G. Della Porta, K. Raaschou-Jensen, E. Travaglino, K. Kallenbach, D. Pietra, V. Ljungström, S. Conte, E. Boveri, R. Invernizzi, R. Rosenquist, P.J. Campbell, M. Cazzola, E. Hellström, Lindberg, SF3B1 mutation identifies a distinct subset of myelodysplastic syndrome with ring sideroblasts, *Blood* 126 (2) (2015) 233–241.
- [24] O.K. Weinberg, R.P. Hasserjian, The current approach to the diagnosis of myelodysplastic syndromes, *Semin. Hematol.* 56 (1) (2019) 15–21.
- [25] D.P. Steensma, Dysplasia has a differential diagnosis: distinguishing genuine myelodysplastic syndromes (MDS) from mimics, imitators, copycats and impostors, *Curr. Hematol. Malig. Rep.* 7 (4) (2012) 310–320.
- [26] T.N. Tanaka, R. Bejar, MDS overlap disorders and diagnostic boundaries, *Blood* 133 (10) (2019) 1086–1095.
- [27] P. Valent, ICUS, IDUS, CHIP and CCUS: diagnostic criteria, separation from MDS and clinical implications, *Pathobiology* 86 (1) (2019) 30–38.
- [28] P. Valent, B.J. Bain, J.M. Bennett, F. Wimalaz, W.R. Sperr, G. Mufti, H.P. Horny, Idiopathic cytopenia of undetermined significance (ICUS) and idiopathic dysplasia of uncertain significance (IDUS), and their distinction from low risk MDS, *Leuk. Res.* 36 (1) (2012) 1–5.
- [29] P. Font, J. Loscertales, C. Benavente, A. Bermejo, M. Callejas, L. Garcia-Alonso, A. Garcia-Marcilla, S. Gil, M. Lopez-Rubio, E. Martin, C. Mun'oz, P. Ricard, S. Soto, P. Balsalobre, A. Villegas, Inter-observer variance with the diagnosis of myelodysplastic syndromes (MDS) following the 2008 WHO classification, *Ann. Hematol.* 92 (1) (2013) 19–24.
- [30] K. Naqvi, E. Jabbour, C. Bueso-Ramos, S. Pierce, G. Borthakur, Z. Estrov, F. Ravandi, S. Faderl, H. Kantarjian, G. Garcia-Manero, Implications of discrepancy in morphologic diagnosis of myelodysplastic syndrome between referral and tertiary care centers, *Blood* 118 (17) (2011) 4690–4693.
- [31] N. Gangat, M.M. Patnaik, K. Begna, A. Al-Kali, M.R. Litzow, R.P. Ketterling, C. A. Hanson, A.D. Pardanani, A. Tefferi, Survival trends in primary myelodysplastic syndromes: a comparative analysis of 1000 patients by year of diagnosis and treatment, *Blood Cancer J.* 6 (2016), e414.
- [32] A. Maassen, C. Strupp, A. Giagounidis, A. Kuendgen, K. Nachtigam, B. Hildebrandt, N. Gattermann, C. Aul, R. Haas, U. Germing, Validation and proposals for a refinement of the WHO 2008 classification of myelodysplastic syndromes without excess of blasts, *Leuk. Res.* 37 (1) (2013) 64–70.
- [33] B.J. Bain, Ethnic and sex differences in the total and differential white cell count and platelet count, *J. Clin. Pathol.* 49 (8) (1996) 664–666.
- [34] L.E. Merz, M. Achebe, When non-Whiteness becomes a condition, *Blood* 137 (1) (2021) 13–15.
- [35] D. Reich, M.A. Nalls, W.H. Kao, E.L. Akyzbekova, A. Tandon, N. Patterson, J. Mullikin, W.C. Hsueh, C.Y. Cheng, J. Coresh, E. Boerwinkle, M. Li, A. Waliszewska, J. Neubauer, R. Li, T.S. Leal, L. Ekinwe, J.C. Files, C.L. Hardy, J. M. Zmuda, H.A. Taylor, E. Ziv, T.B. Harris, J.G. Wilson, Reduced neutrophil count in people of African descent is due to a regulatory variant in the Duffy antigen receptor for chemokines gene, *PLoS Genet* 5 (1) (2009), e1000360.
- [36] S.L. Van Driest, N.S. Abul-Husn, J.T. Glessner, L. Bastarache, S. Nirenberg, J. S. Schildcrout, M.S. Eswarappa, G.M. Belbin, C.M. Shaffer, F. Mentch, J. Connolly, M. Shi, C.M. Stein, D.M. Roden, H. Hakonarson, N.J. Cox, S.C. Borinstein, J. D. Mosley, Association between a common, benign genotype and unnecessary bone marrow biopsies among African American patients, *JAMA Intern. Med* 181 (8) (2021) 1100–1105.
- [37] E. Beutler, C. West, Hematologic differences between African-Americans and whites: the roles of iron deficiency and  $\alpha$ -thalassemia on hemoglobin levels and mean corpuscular volume, *Blood* 106 (2) (2005) 740–745, <https://doi.org/10.1182/blood-2005-02-0713>.
- [38] G.S. Perry, T. Byers, R. Yip, S. Margen, Iron nutrition does not account for the hemoglobin differences between blacks and whites, *J. Nutr.* 122 (7) (1992) 1417–1424, <https://doi.org/10.1093/jn/122.7.1417>.
- [39] T. Karantanos, T. Jain, A.R. Moliterno, R.J. Jones, A.E. DeZern, Sex-related differences in chronic myeloid neoplasms: from the clinical observation to the underlying biology, *Int. J. Mol. Sci.* 22 (5) (2021) 2595, <https://doi.org/10.3390/ijms22052595>.
- [40] A.E. DeZern, A.M. Zeidan, J. Barnard, W. Hand, N. Al Ali, F. Brown, C. Zimmerman, G.J. Roboz, G. Garcia-Manero, D.P. Steensma, R.S. Komrokji, M. A. Sekeres, M.D.S. Evans, Clinical research consortium, differential response to

hypomethylating agents based on sex: a report on behalf of the MDS clinical research consortium (MDS CRC), *Leuk. Lymphoma* 58 (6) (2017) 1325–1331, <https://doi.org/10.1080/10428194.2016.1246726>.

- [41] H. Döhner, A.H. Wei, F.R. Appelbaum, C. Craddock, C.D. DiNardo, H. Dombret, B. L. Ebert, P. Fenaux, L.A. Godley, R.P. Hasserjian, R.A. Larson, R.L. Levine, Y. Miyazaki, D. Niederwieser, G.J. Ossenkoppele, C. Roßlrig, J. Sierra, E.M. Stein, M.S. Tallman, H.F. Tien, J. Wang, A. Wierzbowska, B. Löwenberg, Diagnosis and management of AML in adults: 2022 ELN recommendations from an international expert panel, *Blood* 2022 Jul 7: blood.2022016867. doi: [10.1182/blood.2022.016867](https://doi.org/10.1182/blood.2022.016867) . Epub ahead of print.
- [42] M. Van Dyke, S. Greer, E. Odom, L. Schieb, A. Vaughan, M. Kramer, M. Casper, Heart disease death rates among blacks and whites aged ≥/35 years - United States, 1968-2015, *MMWR Surveill. Summ.* 67 (5) (2018) 1–11.

## **SUPPLEMENTARY TABLES PAPER #1**

**Appendix A. Supplementary data****Supplemental Table S1**

Age-adjusted incidence of MDS in adult patients ( $\geq 20$  years of age) in SEER between 2001 and 2013.

| Rate (95% CI)                                    | Black patients               | White patients             |
|--------------------------------------------------|------------------------------|----------------------------|
| Overall                                          | 5.8 <sup>a,b</sup> (5.6–6.0) | 7.0 <sup>a</sup> (6.9–7.0) |
| Refractory anemia                                | 0.8 (0.7–0.8)                | 0.8 (0.7–0.8)              |
| Refractory anemia with ring sideroblasts         | 0.4 <sup>b</sup> (0.4–0.5)   | 0.6 (0.6–0.6)              |
| Refractory anemia with excess blasts             | 0.7 <sup>b</sup> (0.6–0.7)   | 0.9 (0.9–1.0)              |
| Refractory cytopenia with multilineage dysplasia | 0.3 <sup>b</sup> (0.3–0.3)   | 0.4 (0.4–0.4)              |
| MDS del(5q)                                      | 0.1 <sup>b</sup> (0.1–0.1)   | 0.2 (0.2–0.2)              |
| MDS-NOS                                          | 3.5 <sup>b</sup> (3.3–3.6)   | 4.1 (4.0–4.1)              |

Rates are per 100,000 and age-adjusted to the 2000 United States standard population (19 age groups - Census P25-1130) standard; 95% CIs (Tiwari mod) are presented.

CI, confidence interval; MDS, myelodysplastic syndrome; MDS del(5q), myelodysplastic syndrome associated with isolated del(5q); MDS-NOS, myelodysplastic syndrome—not otherwise specified; SEER, Surveillance, Epidemiology, and End Results.

<sup>a</sup> MDS age-adjusted incidence for 2001-2013 without age or race restriction is 4.8/100,000.

<sup>b</sup> Rate ratio (not shown) for Black patients is statistically different from White patients as calculated using SEER\*STAT.

## Supplemental Table S2

Patient demographics and disease characteristics: matched and unmatched dataset.

|                                 | Matched |       |       |      |       |      | <i>p</i> <sup>a</sup> value | Unmatched |      |                             |
|---------------------------------|---------|-------|-------|------|-------|------|-----------------------------|-----------|------|-----------------------------|
|                                 | Total   |       | Black |      | White |      |                             | Total     |      |                             |
|                                 | N       | %     | N     | %    | N     | %    |                             | N         | %    | <i>p</i> <sup>b</sup> value |
| Total                           | 6052    | 100.0 | 3026  | 50.0 | 3026  | 50.0 |                             | 31,510    |      |                             |
| Sex                             |         |       |       |      |       |      | 1.00                        |           |      | <0.001                      |
| Male                            | 2974    | 49.1  | 1487  | 49.1 | 1487  | 49.1 |                             | 18,541    | 58.8 |                             |
| Female                          | 3078    | 50.9  | 1539  | 50.9 | 1539  | 50.9 |                             | 12,969    | 41.2 |                             |
| Age group (years)               |         |       |       |      |       |      | 1.00                        |           |      | <0.001                      |
| 20-24                           | 18      | 0.3   | 9     | 0.3  | 9     | 0.3  |                             | 41        | 0.1  |                             |
| 25-29                           | 30      | 0.5   | 15    | 0.5  | 15    | 0.5  |                             | 50        | 0.2  |                             |
| 30-34                           | 54      | 0.9   | 27    | 0.9  | 27    | 0.9  |                             | 83        | 0.3  |                             |
| 35-39                           | 70      | 1.2   | 35    | 1.2  | 35    | 1.2  |                             | 126       | 0.4  |                             |
| 40-44                           | 138     | 2.3   | 69    | 2.3  | 69    | 2.3  |                             | 224       | 0.7  |                             |
| 45-49                           | 214     | 3.5   | 107   | 3.5  | 107   | 3.5  |                             | 441       | 1.4  |                             |
| 50-54                           | 300     | 5.0   | 150   | 5.0  | 150   | 5.0  |                             | 756       | 2.4  |                             |
| 55-69                           | 508     | 8.4   | 254   | 8.4  | 254   | 8.4  |                             | 1332      | 4.2  |                             |
| 60-64                           | 602     | 9.9   | 301   | 9.9  | 301   | 9.9  |                             | 2117      | 6.7  |                             |
| 65-69                           | 802     | 13.3  | 401   | 13.3 | 401   | 13.3 |                             | 3298      | 10.5 |                             |
| 70-74                           | 896     | 14.8  | 448   | 14.8 | 448   | 14.8 |                             | 4626      | 14.7 |                             |
| 75-79                           | 908     | 15.0  | 454   | 15.0 | 454   | 15.0 |                             | 6184      | 19.6 |                             |
| 80-84                           | 792     | 13.1  | 396   | 13.1 | 396   | 13.1 |                             | 6417      | 20.4 |                             |
| 85+                             | 720     | 11.9  | 360   | 11.9 | 360   | 11.9 |                             | 5815      | 18.5 |                             |
| Marital status                  |         |       |       |      |       |      | <0.001                      |           |      | <0.001                      |
| Never married                   | 874     | 14.4  | 574   | 19.0 | 300   | 9.9  |                             | 2414      | 7.7  |                             |
| Married                         | 2712    | 44.8  | 1114  | 36.8 | 1598  | 52.8 |                             | 16,824    | 53.4 |                             |
| Separated/divorced/<br>widowed  | 1888    | 31.2  | 1064  | 35.2 | 824   | 27.2 |                             | 9251      | 29.4 |                             |
| Other/unknown                   | 578     | 9.6   | 274   | 9.1  | 304   | 10.0 |                             | 3021      | 9.6  |                             |
| Metropolitan residential status |         |       |       |      |       |      | 1.00                        |           |      | <0.001                      |
| Large metro area                | 4050    | 66.9  | 2025  | 66.9 | 2025  | 66.9 |                             | 18,786    | 59.6 |                             |
| Small metro area                | 1382    | 22.8  | 691   | 22.8 | 691   | 22.8 |                             | 8696      | 27.6 |                             |
| Nonmetropolitan area            | 620     | 10.2  | 310   | 10.2 | 310   | 10.2 |                             | 4028      | 12.8 |                             |
| Income bracket (US dollars)     |         |       |       |      |       |      | <0.001                      |           |      | <0.001                      |
| <75,000                         | 4383    | 72.4  | 2446  | 80.8 | 1937  | 64.0 |                             | 21,398    | 67.9 |                             |
| ≥75,000                         | 1669    | 27.6  | 580   | 19.2 | 1089  | 36.0 |                             | 10,112    | 32.1 |                             |
| Histology                       |         |       |       |      |       |      | 1.00                        |           |      | <0.001                      |
| RA                              | 860     | 14.2  | 430   | 14.2 | 430   | 14.2 |                             | 3601      | 11.4 |                             |
| RARS                            | 532     | 8.8   | 266   | 8.8  | 266   | 8.8  |                             | 3137      | 10.0 |                             |
| RAEB                            | 882     | 14.6  | 441   | 14.6 | 441   | 14.6 |                             | 5195      | 16.5 |                             |
| RCMD                            | 388     | 6.4   | 194   | 6.4  | 194   | 6.4  |                             | 2305      | 7.3  |                             |
| del(5q)                         | 118     | 1.9   | 59    | 1.9  | 59    | 1.9  |                             | 1030      | 3.3  |                             |
| MDS-NOS                         | 3272    | 54.1  | 1636  | 54.1 | 1636  | 54.1 |                             | 16,242    | 51.5 |                             |
| Year of diagnosis               |         |       |       |      |       |      | 0.23                        |           |      | 0.41                        |
| 2001–2004                       | 1565    | 25.9  | 803   | 26.5 | 762   | 25.2 |                             | 7989      | 25.4 |                             |
| 2005–2013                       | 4487    | 74.1  | 2223  | 73.5 | 2264  | 74.8 |                             | 23,521    | 74.6 |                             |

<sup>a</sup> p-value is from the Chi-Squared test for categorical variables comparing Black and White matched patients<sup>b</sup> p-value is from the Chi-Squared test for categorical variables comparing all matched patients with all unmatched patients

MDS del(5q), myelodysplastic syndrome associated with isolated del(5q); MDS-NOS, myelodysplastic syndrome—not otherwise specified; RA, refractory anemia; RAEB, refractory anemia with excess blasts; RARS, refractory anemia with ring sideroblasts; RCMD, refractory cytopenia with multilineage dysplasia.

**Supplemental Table S3**

Vital status and frequency of attributed cause of death.

|                              | All patients with MDS |      | Black |      | White  |      | <i>p</i> value <sup>a</sup> |
|------------------------------|-----------------------|------|-------|------|--------|------|-----------------------------|
|                              | N                     | %    | N     | %    | N      | %    |                             |
| Vital status at study cutoff |                       |      |       |      |        |      | < 0.001                     |
| Alive                        | 5049                  | 13.4 | 638   | 21.0 | 4411   | 12.8 |                             |
| Dead                         | 32,513                | 86.6 | 2399  | 79.0 | 30,114 | 87.2 |                             |
| COD <sup>b</sup>             |                       |      |       |      |        |      | < 0.001                     |
| MDS/Leukemia                 | 15,154                | 46.6 | 944   | 39.3 | 14,210 | 47.2 |                             |
| CVD                          | 5841                  | 18.0 | 470   | 19.6 | 5371   | 17.8 |                             |
| Other malignancies           | 3543                  | 10.9 | 297   | 12.4 | 3246   | 10.8 |                             |
| Other COD                    | 3121                  | 9.6  | 270   | 11.3 | 2851   | 9.5  |                             |
| Infectious CODs              | 1316                  | 4.0  | 133   | 5.5  | 1183   | 3.9  |                             |
| COPD                         | 996                   | 3.1  | 49    | 2.0  | 947    | 3.1  |                             |
| Renal Disease                | 551                   | 1.7  | 88    | 3.7  | 463    | 1.5  |                             |
| Accident/Suicide             | 549                   | 1.7  | 19    | 0.8  | 530    | 1.8  |                             |
| Diabetes mellitus            | 524                   | 1.6  | 68    | 2.8  | 456    | 1.5  |                             |
| Dementia                     | 339                   | 1.0  | 21    | 0.9  | 318    | 1.1  |                             |
| Cirrhosis                    | 314                   | 1.0  | 12    | 0.5  | 302    | 1.0  |                             |
| Unknown                      | 265                   | 0.8  | 28    | 1.2  | 237    | 0.8  |                             |

COD, cause of death; COPD, chronic obstructive pulmonary disease; CVD, cardiovascular disease; MDS, myelodysplastic syndrome.

<sup>a</sup> *p* value is from the Chi-squared test.

<sup>b</sup> Percentages represent frequency for patients with an event.

**Supplemental Table S4**

OS in SEER by histology for patients (≥20 years of age) with MDS diagnosed 2001–2013.

|                                    | Overall          | Black patients   | White patients   | OS-based risk grouping |
|------------------------------------|------------------|------------------|------------------|------------------------|
| Median OS (95% CI), months         | 27 (26–27)       | 33 (31–37)       | 26 (26–27)       |                        |
| Alive at 3 years, % (95% CI)       | 41.8 (41.3–42.3) | 48.3 (46.5–50.1) | 41.2 (40.7–41.8) |                        |
| Alive at 5 years, % 95% CI)        | 29.0 (28.6–29.5) | 36.4 (34.7–38.1) | 28.4 (27.9–28.9) |                        |
| mOS by histology, (95% CI), months |                  |                  |                  |                        |
| RA                                 | 43 (41–45)       | 57 (50–67)       | 41 (38–44)       | Lower risk             |
| RARS                               | 51 (49–53)       | 51 (45–59)       | 51 (49–54)       | Lower risk             |
| RCMD                               | 27 (25–28)       | 34 (23–46)       | 26 (25–28)       | Intermediate risk      |
| MDS del(5q)                        | 29 (26–31)       | 25 (15–33)       | 29 (27–32)       | Intermediate risk      |
| MDS-NOS                            | 28 (28–29)       | 38 (34–42)       | 28 (27–29)       | Intermediate risk      |
| RAEB                               | 11 (10–11)       | 11 (9–14)        | 11 (10–11)       | Higher risk            |

CI, confidence interval; MDS del(5q), myelodysplastic syndrome associated with isolated del(5q); MDS-NOS, myelodysplastic syndrome–not otherwise specified; OS, overall survival; RA, refractory anemia; RAEB, refractory anemia with excess blasts; RARS, refractory anemia with ring sideroblasts; RCMD, refractory cytopenia with multilineage dysplasia; SEER, Surveillance, Epidemiology, and End Results.

**Supplemental Table S5**

Information on MDS-NOS subset in each regional cancer registry.

| Registry     | Overall       |             | MDS-NOS subgroup  |                   |               |
|--------------|---------------|-------------|-------------------|-------------------|---------------|
|              | MDS cases     | MDS-NOS, %  | White patients, % | Black patients, % | Difference, % |
| California   | 13,239        | 55.8        | 55.8              | 56.0              | 0.2           |
| Connecticut  | 1912          | 58.5        | 58.7              | 55.3              | –3.4          |
| Georgia      | 3476          | 58.1        | 57.4              | 61.5              | 4.1           |
| Hawaii       | 158           | 41.8        | 42.3              | 0.0               | –42.3         |
| Iowa         | 2128          | 45.2        | 45.3              | 27.8              | –17.5         |
| Kentucky     | 2004          | 50.0        | 49.9              | 51.5              | 1.6           |
| Louisiana    | 2224          | 40.1        | 40.5              | 38.4              | –2.1          |
| Detroit      | 2836          | 44.1        | 43.6              | 46.3              | 2.7           |
| New Jersey   | 5428          | 60.1        | 59.6              | 65.3              | 5.7           |
| New Mexico   | 697           | 50.6        | 50.8              | 37.5              | –13.3         |
| Utah         | 824           | 44.5        | 44.6              | 25.0              | –19.6         |
| Seattle      | 2636          | 31.8        | 31.7              | 37.7              | 6.0           |
| <b>Total</b> | <b>37,562</b> | <b>52.0</b> | <b>51.8</b>       | <b>53.9</b>       | <b>2.1</b>    |

MDS, myelodysplastic syndrome; MDS-NOS, myelodysplastic syndrome–not otherwise specified.

## Appendix B. Supplementary methods and results

### *Matching procedure*

The overall imbalance between baseline covariate is given by the L1 statistic as a comprehensive measure of global imbalance with values of L1 varying between 0 and 1. Perfect global balance (up to coarsening) is indicated by  $L1 = 0$ , and larger values indicate larger imbalance between the groups, with a maximum of  $L1 = 1$ , which indicates complete separation.

Out of the 504 possible cells, 455 were populated based on actual values associated with variables used for matching and 334 of these cells included the matched dataset. The matching procedure reduced imbalance between covariates between both groups as measured by the L1 statistics after matching = 0 (compared with a L1 value = 0.235 before matching).

In summary, there were 6052 patients matched (3026 Black and 3026 White) and 31,510 unmatched patients (11 Black and 31,499 White).

The *Cem* command was used in Stata to implement the matching procedure.

### *Multiple imputation procedure*

Model 2 in Table 2 reports the results from the multiple imputation of both the 3599 other/unknown values for marital status and the 19,514 cases associated with myelodysplastic syndrome—not otherwise specified from the observed dataset; these data were imputed (30 imputations) by predictive mean matching in exactly one procedure that controlled for race, sex, age group, income group, residential status group, year of diagnosis periods. The multinomial logistic function, *mlogit* command in Stata, was used to impute the missing values and *mi* command was used to run the multiple imputation procedure.

## PAPER #2.

| Leukemia and Lymphoma                                                                                      |                                                                                                                                                                                                                                                   |
|------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                           |                                                                                                                                                                                                                                                   |
| <b>Temporal trend in survival for Chronic Myelomonocytic Leukemia in the US: A population-based study.</b> |                                                                                                                                                                                                                                                   |
| Journal:                                                                                                   | <i>Leukemia and Lymphoma</i>                                                                                                                                                                                                                      |
| Manuscript ID                                                                                              | Draft                                                                                                                                                                                                                                             |
| Manuscript Type:                                                                                           | Original Article                                                                                                                                                                                                                                  |
| Date Submitted by the Author:                                                                              | n/a                                                                                                                                                                                                                                               |
| Complete List of Authors:                                                                                  | Lesegretain, Arnaud; Harvard Medical School; Daiichi Sankyo Inc<br>Brunner, Andrew; Massachusetts General Hospital, Center for Leukemia<br>Benzohra, Aziz; Daiichi Sankyo Inc<br>Fathi, Amir; Massachusetts General Hospital, Medicine - Oncology |
| Keywords:                                                                                                  | CMML, SEER, Survival, Registry, Chronic Myelomonocytic Leukemia                                                                                                                                                                                   |
|                         |                                                                                                                                                                                                                                                   |

Submitted March 16<sup>th</sup> 2023

Following input from the Thesis committee, additional details were added compared to the submitted version of the original manuscript.

## **ABSTRACT**

Outcomes for Chronic Myelomonocytic Leukemia are insufficiently characterized at the population level. We analyzed epidemiological trends for patients between 2001 and 2017, focusing on age, sex, race and long-term survivors. Using the Surveillance, Epidemiology, and End Results Program, we studied 3929 patients, in 4 time-period (tp) cohorts, based on year of diagnosis [2001-2004 (tp1); 2005-2009 (tp2); 2010-2013 (tp3); 2014-2017 (tp4)]. Stable incidence overall, male predominance and higher incidence for White versus Black and “Other” races were noted. Three-year Relative Survival (RS) increased from 27.9% to 36.9% between tp1 and tp4. The most pronounced increase occurred between tp1 and tp2. All subgroups generally experienced RS improvements over time, except notably Black patients. Improvements for patients aged 85+ (3-year RS 8.4% to 23.6% between tp1 and tp4) and increases in long-term survivors (5-year OS from 13.2% to 22.3%) were observed. Additional study is warranted to explore these associations, particularly for Black patients.

*Keywords:* CMML, Chronic Myelomonocytic Leukemia, Survival, SEER, registry, outcomes

## **Introduction**

CMML is a myeloid malignancy of the elderly (median age at diagnosis 71-73 years) with a variable but typically poor prognosis. The median OS estimates is 12–37 months, with approximately 15–25% transforming to AML [1-3]. It was classified as a form of myelodysplastic syndrome (MDS) in the French-American-British classification system in 1982, but was subsequently reclassified as a mixed myelodysplastic/myeloproliferative disorder by the World Health Organization (WHO) in 2001, and subsequently updated in 2016 [4].

CMML has traditionally been studied epidemiologically and therapeutically as a subset of MDS, although in recent years, it has gained significant interest as an individual entity [1-8]. The

management of CMML and MDS patients often share the use of erythropoiesis stimulating agents, granulocyte colony-stimulating factor, transfusion support, hypomethylating agents (HMA), and allogeneic hematopoietic stem cell transplantation as the only curative option [5-7]. CMML is distinguished clinically by persistent (>3 months) monocytosis ( $\geq 1 \times 10^9/L$ ; monocytes  $\geq 10\%$  of white blood cell count), splenomegaly, a characteristic mutational profile [1,2]. WHO 2016 classified CMML into “proliferative” and “dysplastic” sub-types with further subclassification in three categories based on peripheral and bone marrow blast % [4].

The lack of CMML-specific clinical trials makes it challenging to understand the natural course of the disease and the impact of treatment used in clinical practice. Allogeneic stem cell transplantation is the only potential curative option for higher-risk patients, but advanced age at diagnosis often limits the proportion of patients eligible [1]. Hypomethylating agents (HMA) are commonly used for the treatment of CMML patients based on pivotal trials for MDS [10-14]. Previous small and retrospective studies in CMML hinted at an improvement in survival outcomes associated with the availability of HMAs [15-26]. In a SEER-medicare study on 1368 CMML patients (age  $\geq 66$  years) diagnosed between 2001 and 2011, Zeidan et al. reported an improvement in overall survival for a matched cohort of HMA treated patients between pre and post HMA approval periods (hazard ratio, 0.72; 95% confidence interval [CI], 0.58-0.91;  $p=.005$ ) [26]. However, it is unclear if these improvements have continued in recent time and if various subgroups have benefited equally. There is also very little data about long-term survivors given the short follow-up of clinical trials. Using the SEER NIH cancer registry [19], our study aims at including the most recent longitudinal data to describe at the population level the temporal trends in survival outcome in the US for CMML patients diagnosed between 2001 and 2017. Our study had three primary objectives: to describe how survival outcomes evolved over time, including in a post-HMA period; to determine if the overall trend in survival differs by key

subgroups; and lastly, to characterize how the group of patients considered long-term survivors has evolved over time.

## **2. Methods**

### **2.1. Data sources and population**

We used the SEER Program from the National Institutes of Health that collects data on cancer incidence, survival, and mortality for 26.5% of the population in the US [27-28]. It is an authoritative source of information on cancer incidence and survival in the United States. Cancer data are updated annually to capture vital status, survival time, and cause of death with a 90% or greater completeness standard for cancers [28]. Expected survival was obtained from life tables produced by the Center for Disease Control [34] and available until 2017 by age, sex and race coded as following: White, Black, Other. We assumed same expected survival for 2018 and 2019 as in 2017 to keep consistent with our classification of race.

We only included patients with a microscopically confirmed CMML diagnosis reported to SEER between 2001 and 2017 [29]. We chose 2001 as a starting date as it represented the date when CMML was classified as a malignancy according to the World Health Organization (WHO) and became reportable to cancer registries. The data cutoff was December 2019. Cases were extracted based on International Classification of Diseases for Oncology version 3 (ICD-O3) code 9945 (chronic myelomonocytic leukemia; CMML).

Patients were included if they were  $\geq 20$  years of age at time of CMML diagnosis, and if CMML was the first cancer diagnosed in SEER. We excluded patients with zero days of follow-up or zero days of survival, and those with diagnosis based on autopsy or death certificate only.

### **2.2. Measures**

Cases were included in 4 time-period cohorts based on year of diagnosis defined as a categorical variable (cohort 1: 2001–2004, cohort 2: 2005–2009, cohort 3: 2010–2013, cohort 4:

2014-2017). Other variables from SEER used to categorize patients at baseline were age as a categorical variable (20–64, 65–74, 75–84, or 85+ years), age as a continuous variable, sex (male, female), race (Black, White, Other).

### 2.3. Statistical analysis

To compare baseline characteristics between diagnosis cohorts, we used the Chi-squared test for categorical covariates and the Wilcoxon rank-sum test for continuous covariates. We reported age-adjusted incidence rate to reduce confounding due to change in the mix of the age structure of our studied population over time. It is calculated as a weighted average of the age-specific incidence rates, where the weights are the proportions of persons in the corresponding age groups of a standard population defined as the US population in 2000. Age-adjusted disease incidence was obtained directly from SEER dataset using the incidence-session and access via SEER\*Stat software on November 21st, 2022 [30].

We used net survival (a measure that provides estimates of survival of the disease in the absence of other causes) in a relative survival (RS) framework as the primary method used to analyze the survival outcomes of CMML patients (for the entire population and by subgroups defined by sex, age groups and race) and how it evolved over time [31-33]. The major advantage of RS is that it does not require information on specific causes of death which have inherent inaccuracies in cancer registries. RS is used in population health research to compare the survival of a cohort of cancer patients to the survival of the general population. It is calculated as the ratio of observed survival (all-cause deaths) for a group of cancer patients divided by the expected survival of a corresponding group from the general population based on age, sex and race. The *strs* command in Stata was used to compute relative survival [35].

The analytical plan to calculate and model Relative Survival Estimate is as follows:

First, a datafile is extracted from SEER which contains individual patient data with sex, race, year of cancer diagnosis, age at diagnosis, number of months of survival and alive status at the study cut-off date. Information of survival in month is transformed in years.

Second, a grouped file is created by the strs Stata command to derive RS point estimates (and 95% confidence intervals) for each stratum corresponding to the relevant predictor variable (sex, age group, race). Because RS computation matches survival information from individual patients to annual lifetables, observed survival is calculated with the actuarial method and assumes that half of the population is at risk at the midpoint for each age-year. This assumes that death and censoring are evenly distributed throughout the interval under consideration, an assumption that is reasonable over a long study period. Cumulative expected survival was calculated with the Ederer II method, considered the preferred method among the three various estimators of relative survival (Ederer I, Ederer II, Hakulinen) [33].

An example of selected fields from the grouped survival file used to derive 3-year RS by sex is provided below:

| Diag | Sex | start | end | n   | cp     | cp_e2  | cr_e2  | se_cr_e2 | lo_cr_e2 | hi_cr_e2 |
|------|-----|-------|-----|-----|--------|--------|--------|----------|----------|----------|
| 1    | 1   | 0     | 3   | 475 | 0.2363 | 0.8618 | 0.2742 | 0.0226   | 0.231    | 0.3195   |
| 1    | 2   | 0     | 3   | 318 | 0.2516 | 0.8785 | 0.2864 | 0.0277   | 0.2337   | 0.3419   |
| 2    | 1   | 0     | 3   | 700 | 0.2817 | 0.8803 | 0.32   | 0.0193   | 0.2826   | 0.3583   |
| 2    | 2   | 0     | 3   | 419 | 0.3182 | 0.893  | 0.3563 | 0.0255   | 0.3069   | 0.4066   |
| 3    | 1   | 0     | 3   | 570 | 0.3234 | 0.8864 | 0.3648 | 0.0221   | 0.3218   | 0.4084   |
| 3    | 2   | 0     | 3   | 342 | 0.3411 | 0.9002 | 0.379  | 0.0285   | 0.3236   | 0.4349   |
| 4    | 1   | 0     | 3   | 684 | 0.3179 | 0.8918 | 0.3564 | 0.0205   | 0.3166   | 0.3969   |
| 4    | 2   | 0     | 3   | 421 | 0.3503 | 0.9006 | 0.3889 | 0.0267   | 0.3369   | 0.4412   |

n is number of alive patients at beginning of interval; cp represents cumulated observed survival; cp\_e2 represents cumulative expected survival and cr\_e2 gives the cumulative relative survival. For example, for male (sex=1) CMMML patients diagnosed during the first period of diagnosis (diag=1), the 3-year RS is calculated as:  $23.63 / 86.18 = 27.42$  with a corresponding 95% CI: 0.231-0.3195.

Lastly, we modeled excess deaths in order to characterize differences in RS. Regression modeling for relative survival has been extensively described in the literature with various

approaches producing similar estimates [36]. A brief summary of our chosen approach is provided below.

When analyzing RS, the entity of interest is excess death due to cancer and this is the excess mortality that is modelled. The underlying model for RS assumes an additive model where the total hazard is the sum of known baseline hazard (estimated from the general population mortality rates) and the excess hazard associated with the CMML. The excess hazards are assumed to be constant within pre-specified bands of follow-up. This assumption of piecewise constant hazards implies a Poisson process for the number of deaths in each interval.

Therefore, we have decided to compare differences in RS by calculating and comparing Excess Hazard Ratio (EHR) of deaths between groups of patients using a generalized linear model (GLM) approach, with a Poisson assumption for the number of observed deaths and death as an outcome. The GLM models compares observed death to expected deaths and returns a EHR as a measure of relative excess risk. We compared EHRs for each time-period cohort to the one immediately before and compared period 4 with period 1, using the chi-square test statistic with 1 and 3 degrees of freedom, respectively. Analytically, the grouped survival file generated by the `strs` command is used to calculate EHR with a sample size equal to the number of cross-classified demographic groupings of expected survival computation; using the example cited previously of the 3-year RS by sex, the sample size of the file used to model EHR has 8 records (number of periods of diagnosis (4) x number of sex (2)). The GLM model runs with a user-defined link function previously described [35,36].

We also performed age-standardized analysis using period 4 as our reference in order facilitate comparison of RS between different periods where the age group mix may have fluctuated. In detail, we used the proportion of age group by sex and age group by race of period 4 as standard for all 3 previous periods to calculate age-standardized RS point estimates.

As a supplementary measure, we also performed overall survival analysis using the Kaplan-Meier method and compared survival using the log-rank test. OS (in months) was calculated from the time of diagnosis until death or censoring.

We characterized long-term survivors with 5-year RS % and 5-year OS %.

All analyses were done using Stata 17.0 BE (StataCorp LLC; College Station, TX). Statistical significance for all two-sided tests was declared for  $p < 0.05$ .

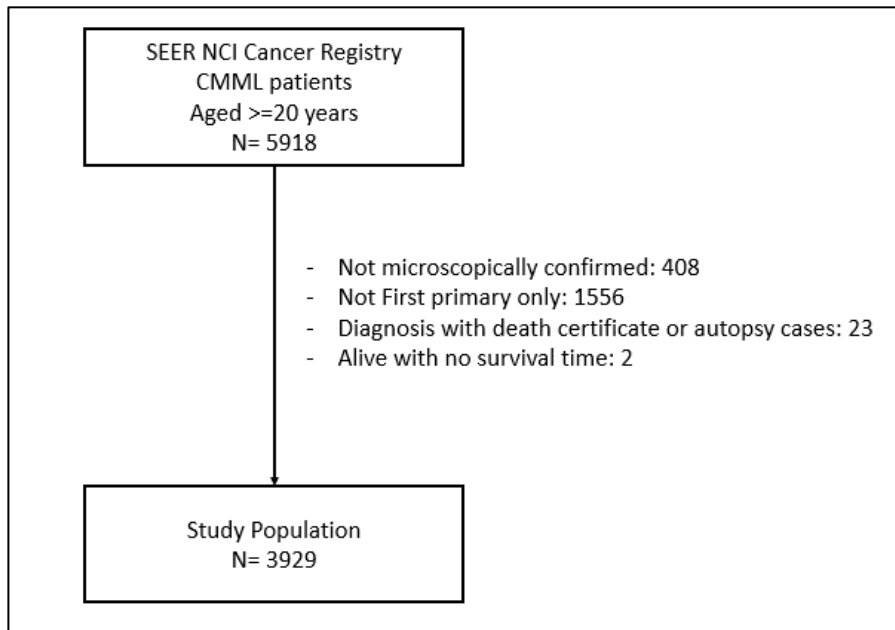
This study was considered exempt by the Dana-Farber Harvard Cancer Center institutional review board.

### **3. Results**

#### **3.1. Population and incidence rate**

A total of 3929 patients for a total of 9,345 person-years at-risk (Table 1) were identified after applying inclusion (CMML diagnosis aged 20 years or older) and exclusion criteria (Figure 1), the latter of which included “not first primary only” “diagnosis, “diagnosis microscopically not confirmed”, cases with diagnosis based only on death certificate or autopsy cases” and cases “alive with no survival time”. Supplemental Table S6 provides further details about the study population and the excluded population.

#### **Figure 1**



Median age at diagnosis was 75 years (interquartile range [IQR], 67–82) and 80.6% of all cases were diagnosed at age >65 years. There was a male predominance with 61.8% of males in our cohort. Approximately 86.0%, 6.2% and 7.8% of the population was classified as White, Black and Other in SEER, respectively. Over the 4 periods of diagnosis, there was no significant difference in the composition of each cohort based on median age, age group and sex.

**Table 1**

Patient demographics at diagnosis - CMML patient diagnosed 2001-2017 - SEER (n=3929)

|                         | <b>Total</b> |          | <b>2001-2004</b> |          | <b>2005-2009</b> |          | <b>2010-2013</b> |          | <b>2014-2017</b> |          | <b>p<sup>a</sup><br/>value</b> |
|-------------------------|--------------|----------|------------------|----------|------------------|----------|------------------|----------|------------------|----------|--------------------------------|
|                         | <b>N</b>     | <b>%</b> | <b>N</b>         | <b>%</b> | <b>N</b>         | <b>%</b> | <b>N</b>         | <b>%</b> | <b>N</b>         | <b>%</b> |                                |
| <b>N</b>                | 3,929        |          | 793              | 20.2%    | 1,119            | 28.5%    | 912              | 23.2%    | 1,105            | 28.1%    |                                |
| <b>Sex</b>              |              |          |                  |          |                  |          |                  |          |                  |          | 0.64                           |
| Male                    | 2,429        | 61.8%    | 475              | 59.9%    | 700              | 62.6%    | 570              | 62.5%    | 684              | 61.9%    |                                |
| Female                  | 1,500        | 38.2%    | 318              | 40.1%    | 419              | 37.4%    | 342              | 37.5%    | 421              | 38.1%    |                                |
| <b>Median age - IQR</b> | 75 (67-82)   |          | 75 (68-82)       |          | 75 (66-82)       |          | 74 (66-82)       |          | 74 (67-82)       |          | 0.14                           |
| <b>Age Group</b>        |              |          |                  |          |                  |          |                  |          |                  |          | 0.05                           |
| 20-64                   | 759          | 19.3%    | 129              | 16.3%    | 239              | 21.4%    | 186              | 20.4%    | 205              | 18.6%    |                                |
| 65-74                   | 1,196        | 30.4%    | 233              | 29.4%    | 320              | 28.6%    | 277              | 30.4%    | 366              | 33.1%    |                                |
| 75-84                   | 1,372        | 34.9%    | 302              | 38.1%    | 398              | 35.6%    | 313              | 34.3%    | 359              | 32.5%    |                                |
| 85+                     | 602          | 15.3%    | 129              | 16.3%    | 162              | 14.5%    | 136              | 14.9%    | 175              | 15.8%    |                                |
| <b>Race</b>             |              |          |                  |          |                  |          |                  |          |                  |          | 0.03                           |
| White                   | 3,378        | 86.0%    | 705              | 88.9%    | 958              | 85.6%    | 787              | 86.3%    | 928              | 84.0%    |                                |
| Black                   | 243          | 6.2%     | 46               | 5.8%     | 65               | 5.8%     | 60               | 6.6%     | 72               | 6.5%     |                                |
| Other <sup>b</sup>      | 308          | 7.8%     | 42               | 5.3%     | 96               | 8.6%     | 65               | 7.1%     | 105              | 9.5%     |                                |

CMML, chronic myelomonocytic leukemia; IQR: interquartile range; SEER, Surveillance, Epidemiology, and End Results  
a: p-value is from the Chi-Squared test for categorical variables and Wilcoxon Rank-Sum test for continuous variable b:  
Includes: n=291 Asian/Pacific-Islander, American Indians/Alaska Natives; n=17: unknown

Overall, the age adjusted CMML incidence rate (IR) between 2001 and 2017 for our cohort of adult patients with CMML was 0.4/100,000 and stable during the study period (Table 2 and Supplemental Table S1). The IR was more than double in male (0.6/100,000) compared to female (0.3/100,000) and was statistically higher for White patients (0.4/100,000) compared with Black and “Other” race patients (0.3/100,000). Age-adjusted incidence nearly doubled for patients diagnosed between 75-84 years (2.5/100,000) compared to patients aged 65-74 years (1.3/100,000).

**Table 2**Age-adjusted CMML incidence - patient diagnosed 2001-2017 - SEER (n=3991<sup>a</sup>)

|                          | Entire period<br>2001-2017 |                         |                           |
|--------------------------|----------------------------|-------------------------|---------------------------|
|                          | N                          | IR <sup>b</sup> (95%CI) | IRR <sup>c</sup> (95% CI) |
| <b>Total population</b>  | <b>3991</b>                | 0.4 (0.4-0.4)           | n/a                       |
| <b>Sex</b>               |                            |                         |                           |
| Male                     | 2462                       | 0.6 (0.6-0.6)           | Ref.                      |
| Female                   | 1529                       | 0.3 (0.3-0.3)           | 0.47** (0.44-0.50)        |
| <b>Age Group - years</b> |                            |                         |                           |
| 20-64                    | 768                        | 0.1 (0.1-0.1)           | Ref.                      |
| 65-74                    | 1205                       | 1.3 (1.3-1.4)           | 16.3** (14.9-17.9)        |
| 75-84                    | 1402                       | 2.5 (2.4-2.7)           | 31.1** (28.5-34.1)        |
| 85+                      | 616                        | 2.7 (2.5-3.0)           | 33.6** (30.2-37.4)        |
| <b>Race</b>              |                            |                         |                           |
| White                    | 3431                       | 0.4 (0.4-0.4)           | Ref.                      |
| Black                    | 246                        | 0.3 (0.3-0.3)           | 0.70** (0.61-0.80)        |
| Other                    | 314                        | 0.3 (0.3-0.4)           | 0.76** (0.67-0.85)        |

CMML, chronic myelomonocytic leukemia; IR: Incidence Rate; IRR:

Incidence Rate Ratio; CI, Confidence Interval

a: the incidence session used to obtain IRR from SEER included 62 more patients than the population used for the rest of the study

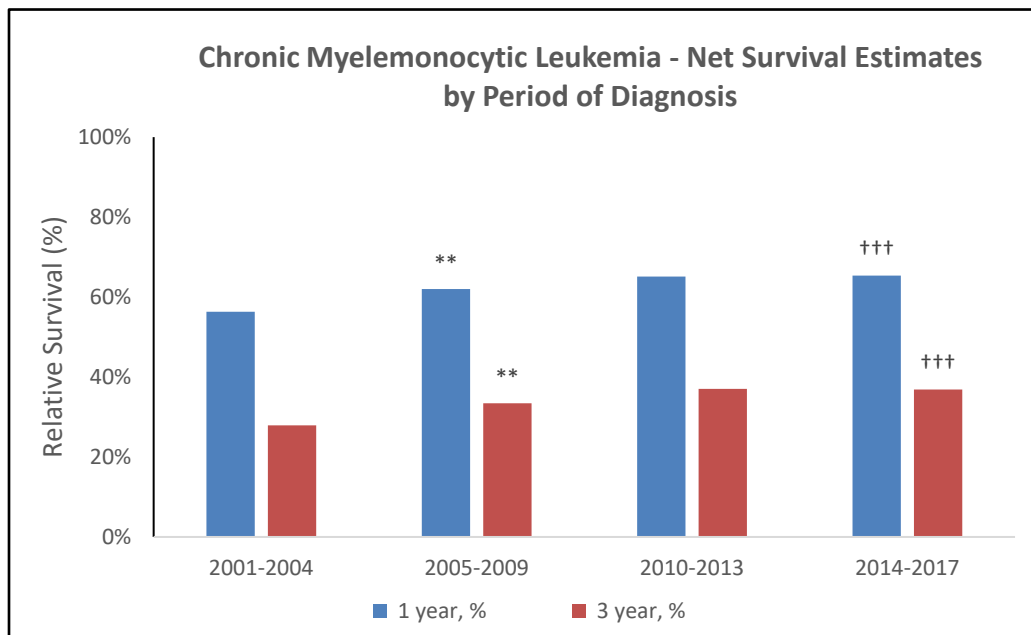
b: Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard; Confidence intervals (Tiwari mod) are 95% for rates

c: Test of difference: \*: p&lt;0.05; \*\*: p&lt;.001

### 3.2. Relative Survival and Overall Survival - entire CMML cohort.

Figure 2 and Supplemental Table S2 show the 1- and 3- relative survival for the entire cohort of CMML patients for the 4 pre-specified periods of diagnosis. In the most recent period analyzed, relative survival (RS) was 65.4%, 36.9% and 24.5% at 1, 3 and 5 years respectively. We report statistical differences across the 4 periods and observed an increase in 1-, 3- and 5-year RS over time. The strongest improvement was observed between 2001-2004 and 2005-2009 with 1-, 3-, and 5-year RS all showing statistically significant improvements. There was no statistically significant change in RS after period 2 except for the 5-year RS between cohort 2 and cohort 3. An analysis of 1-, 3- and 5-year RS by year is available in

**Figure 2**



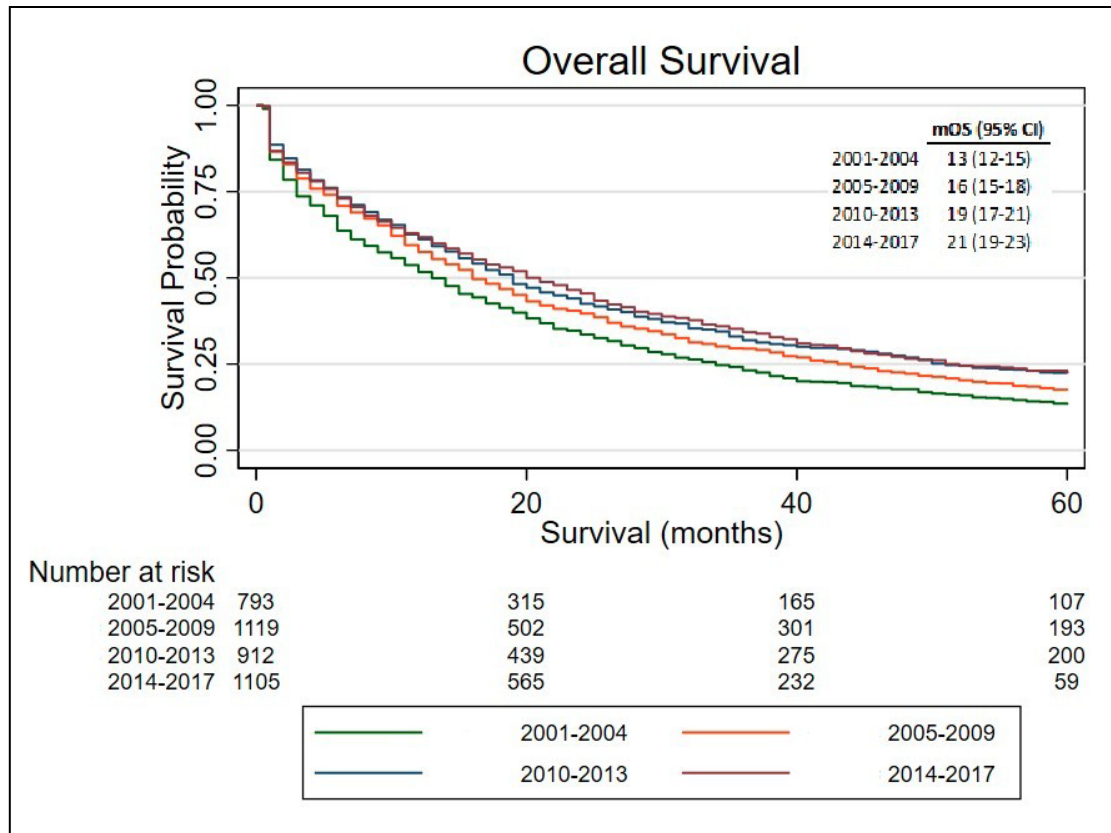
Statistically significant difference between Excess Hazard Ratio with previous cohort: \*:<.05; \*\*:<.01; \*\*\*:<.001

Statistically significant difference between Excess Hazard Ratio between the 1st and 4th cohorts: †:<.05; ††:<.01; †††:<.001

Relative Survival trend for CMML patients 2001-2017- SEER (n=3929), compared to a non-diseased US population by historical time-based period cohorts, as described in Section 2.3 of the Methods

As clinicians may be more familiar with measures associated with overall survival, we also report median OS and Kaplan-Meier estimates for each study-period cohort. mOS was 13, 16, 19 and 21 months in the 4 time-periods respectively, reflecting in absolute terms the improvement in survival to cancer and other non-cancer related factors. Detailed mOS by subgroups are available in Supplemental Table S5. Univariate cox-regression analysis using period 1 (2001-2004) as reference resulted in the following hazard ratio for the risk of all-cause death: 0.85 (95% CI:0.76-0.94), p=.002, 0.76 (95% CI:0.68-0.85), p<.001; 0.74 (95% CI:0.67-0.83), p<.001 for period 2, 3 and 4, respectively.

**Figure 3**



5-year Overall Survival by period of diagnosis for CMML patients 2001-2017- SEER (n=3929)  
 mOS: median Overall Survival

### 3.3. 3-year Relative Survival by subgroups

Table 3 shows the 3-year relative survival for various cohorts defined by sex, age groups and race. Overall, RS was significantly different across the 4 periods of diagnosis and increased over time for males and females, all age groups, White and “Other” races. The Black race was not associated with an increase in RS overall although an increase in RS was observed between period 1 and period 2. When we compared the last cohort with the first cohort, RS increased for all subgroups, except for Black patients. 3-year RS point estimates for White was above that for Black patients for three out of the four time periods but our model did not detect a difference in Excess Hazard Ratio between Black and White patients, possibly due the small sample size of the Black cohort.

Not surprisingly, RS depended on age, and in period 1, RS was 37.4% and 8.4% for 20-64 and 85+ years age groups, respectively. The largest improvement in unit % occurred in the 85+ years age group with RS increasing from 8.4% for cohort 1 to 23.6% for cohort 4. All other age groups also saw an improvement in RS between period 1 and period 4.

**Table 3**

Relative Survival trend by various - CMML cohorts 2001-2017 - SEER (n=3929)

|                    | 3-year Relative Survival <sup>a,b,c</sup> % (95% CI) |                                |                                |                                 | Difference <sup>d</sup> |
|--------------------|------------------------------------------------------|--------------------------------|--------------------------------|---------------------------------|-------------------------|
|                    | 2001-2004                                            | 2005-2009                      | 2010-2013                      | 2014-2017                       |                         |
| <b>Overall</b>     | 27.9 (24.5-31.4)                                     | 33.4 <sup>**</sup> (30.4-36.4) | 37.0 (33.6-40.5)               | 36.9 <sup>+++</sup> (33.7-40.1) | $\chi^2_3=37.6^{***}$   |
| <b>Sex</b>         |                                                      |                                |                                |                                 |                         |
| Male               | 27.4 (23.1-32.0)                                     | 32.0 <sup>*</sup> (28.3-35.8)  | 36.5 <sup>*</sup> (32.2-40.8)  | 35.6 <sup>+++</sup> (31.7-39.7) | $\chi^2_3=23.3^{***}$   |
| Female             | 28.6 (23.4-34.2)                                     | 35.6 <sup>*</sup> (30.7-40.7)  | 37.9 (32.4-43.5)               | 38.9 <sup>+++</sup> (33.7-44.1) | $\chi^2_3=15.2^{**}$    |
| <b>Age groups</b>  |                                                      |                                |                                |                                 |                         |
| 20-64              | 37.4 (29.0-45.8)                                     | 37.5 (31.2-43.7)               | 49.6 <sup>**</sup> (42.1-56.7) | 44.7 <sup>†</sup> (37.4-51.8)   | $\chi^2_3=12.0^{**}$    |
| 65-74              | 27.2 (21.4-33.4)                                     | 37.5 <sup>**</sup> (32.0-43.1) | 38.0 (32.0-44.0)               | 35.4 <sup>††</sup> (30.2-40.7)  | $\chi^2_3=13.7^{**}$    |
| 75-84              | 31.1 (25.4-37.2)                                     | 33.1 (28.0-38.4)               | 31.8 (26.2-37.7)               | 39.5 <sup>+++</sup> (33.7-45.3) | $\chi^2_3=10.9^*$       |
| 85+                | 8.4 (3.9-15.3)                                       | 16.4 <sup>**</sup> (10.4-23.8) | 26.0 (17.9-35.3)               | 23.6 <sup>+++</sup> (16.6-31.7) | $\chi^2_3=24.1^{***}$   |
| <b>Race</b>        |                                                      |                                |                                |                                 |                         |
| White <sup>c</sup> | 28.3 (24.7-32.0)                                     | 32.5 <sup>*</sup> (29.3-35.8)  | 38.6 <sup>**</sup> (34.9-42.4) | 36.9 <sup>+++</sup> (33.5-40.4) | $\chi^2_3=33.4^{***}$   |
| Black <sup>c</sup> | 24.6 (12.5-39.6)                                     | 37.9 <sup>*</sup> (25.3-51.0)  | 31.6 (19.7-44.7)               | 29.1 (18.3-41.0)                | $\chi^2_3=6.5$          |
| Other <sup>c</sup> | 25.1 (12.7-39.9)                                     | 38.4 <sup>*</sup> (28.3-48.7)  | 23.2 <sup>*</sup> (13.5-34.6)  | 42.0 <sup>+++</sup> (31.6-52.3) | $\chi^2_3=12.5^*$       |

CMML, chronic myelomonocytic leukemia; SEER, Surveillance, Epidemiology, and End Results

a: actuarial method used for observed survival. Ederer II method used for cumulative expected survival.

b: statistically significant difference between Excess Hazard Ratio with previous cohort: <sup>\*</sup>:<.05; <sup>\*\*</sup>:<.01; <sup>+++</sup>:<.001

c: statistically significant difference between Excess Hazard Ratio between the 1st and 4th cohorts: <sup>†</sup>:<.05; <sup>††</sup>:<.01; <sup>+++</sup>:<.001

d: the str command was run for each row of the table and then we ran a regression on the 4 observations to get a 3 degrees of freedom chi-square test. <sup>\*</sup>:<.05; <sup>\*\*</sup>:<.01; <sup>+++</sup>:<.001

An age-standardized analysis (Supplemental – Table S4) using period 4 age distribution was consistent with the age-unadjusted analysis.

### 3.3. 3-year Relative Survival stratified by age

Given that RS varied by age, we conducted a stratified analysis of RS by age group between the 2001-2004 and 2014-2017 cohorts. Table 4 shows RS within each age group for the entire population and by sex (Table 4). Race was excluded from this analysis due to very

small sample size for Black and “Other” subgroups (n<10 in many cells). RS improved between the 2 time periods overall except for female aged 20-64.

**Table 4**

3-year Relative Survival stratified by age group - CMML patients diagnosed 2001-2017- SEER (n=3929)

| 3-year Relative Survival <sup>a,b</sup> % (95% CI) |                     |                                  |                     |                                    |
|----------------------------------------------------|---------------------|----------------------------------|---------------------|------------------------------------|
|                                                    | 20-64               |                                  | 65-74               |                                    |
|                                                    | 2001-2004           | 2014-2017                        | 2001-2004           | 2014-2017                          |
| <b>Overall</b>                                     | 37.4<br>(29.0-45.8) | 44.7 <sup>†</sup><br>(37.4-51.8) | 27.2<br>(21.4-33.4) | 35.4 <sup>††</sup><br>(30.2-40.7)  |
| <b>Sex</b>                                         |                     |                                  |                     |                                    |
| Male                                               | 30.8<br>(20.8-41.5) | 42.7 <sup>†</sup><br>(33.7-51.4) | 27.9<br>(20.9-35.3) | 34.5 <sup>†</sup><br>(28.2-40.9)   |
| Female                                             | 46.9<br>(32.8-59.9) | 48.5 <sup>†</sup><br>(35.8-60.2) | 25.7<br>(15.6-37.2) | 37.5 <sup>†</sup><br>(28.1-46.9)   |
| <b>Race</b>                                        |                     |                                  |                     |                                    |
| White                                              | 36.3<br>(27.5-45.2) | 46.3 <sup>†</sup><br>(38.0-54.2) | 27.6<br>(21.2-34.4) | 35.1 <sup>†</sup><br>(29.3-40.9)   |
| Black                                              | 44.1<br>(10.1-75.6) | 40.0<br>(18.6-61.1)              | 25.5<br>(9.3-46.4)  | 30.9<br>(14.2-50.0)                |
| Other                                              | 50.5<br>(11.2-81.1) | 37.3<br>(16.7-58.3)              | 25.3<br>(7.0-49.5)  | 42.1<br>(24.7-58.8)                |
|                                                    | 75-84               |                                  | 85+                 |                                    |
|                                                    | 2001-2004           | 2014-2017                        | 2001-2004           | 2014-2017                          |
| <b>Overall</b>                                     | 31.1<br>(25.4-37.2) | 39.5 <sup>†</sup><br>(33.7-45.3) | 8.4<br>(3.9-15.3)   | 23.6 <sup>†††</sup><br>(16.6-31.7) |
| <b>Sex</b>                                         |                     |                                  |                     |                                    |
| Male                                               | 31.6<br>(23.9-39.8) | 37.5 <sup>†</sup><br>(30.1-45.0) | 4.9<br>(0.9-15.1)   | 20.4 <sup>†</sup><br>(11.2-32.1)   |
| Female                                             | 30.5 (22.1-39.6)    | 42.4 <sup>†</sup> (33.2-51.6)    | 11.0 (4.5-21.2)     | 26.4 <sup>††</sup> (16.7-37.8)     |
| <b>Race</b>                                        |                     |                                  |                     |                                    |
| White                                              | 32.1<br>(26.0-38.4) | 39.7 <sup>†</sup><br>(33.4-46.0) | 8.4<br>(3.7-15.8)   | 23.4 <sup>†††</sup><br>(16.0-32.0) |
| Black                                              | 25.4<br>(2.5-65.4)  | 23.8<br>(5.8-50.3)               | n/a                 | n/a                                |
| Other                                              | 16.0<br>(2.6-41.0)  | 44.5 <sup>†</sup><br>(26.4-62.0) | 17.8<br>(8.9-57.8)  | 42.7<br>(13.1-73.7)                |

CMML, chronic myelomonocytic leukemia; SEER, Surveillance, Epidemiology, and End Results

a: actuarial method used for observed survival. Ederer II method used for cumulative expected survival. The *strs* command was run for each row of the table for each age group and then we ran a regression on the 2 observations to get a 1 degree of freedom chi-square test

b: statistically significant difference of the Excess Hazard Ratio across the two periods: <sup>†</sup>:<.05; <sup>††</sup>:<.01; <sup>†††</sup>:<.001

### 3.4. Long-term survivors

5-year RS % and 5-year OS % are showed in Table 5 and Supplemental S5 (Appendix) respectively. For the entire CMML population, there was a statistically difference in RS and OS across the 4 cohorts, reflecting an increase in the percentage of patients surviving their cancer

and alive 5 years after diagnosis. Males and females both saw an increase in long-term survival, measured by RS or OS. All age groups also saw an increase in 5-year RS and 5-year OS across the 4 periods. In terms of race, our results show a statistically significant difference across the 4 periods for White and “Other” Races. For the Black subgroup of patients, we did not detect a statistically significant change in RS across the 4 periods. The 5-year OS% for the Black cohort reach the 20% threshold in period 4 while the White cohort exceeded this mark earlier during period 3.

**Table 5**

Long-term survivors - CMML patients diagnosed 2001-2017- SEER (n=3929)

|                    | 5-year Relative Survival <sup>a,b,c</sup> % (95% CI) |                               |                               |                               | Difference <sup>d</sup> |
|--------------------|------------------------------------------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------|
|                    | 2001-2004                                            | 2005-2009                     | 2010-2013                     | 2014-2017                     |                         |
| <b>Total</b>       | 17.2 (14.3-20.3)                                     | 21.6*** (19.0-24.4)           | 27.0** (23.8-30.3)            | 24.5*** (21.5-27.6)           | $\chi^2_3=41.3^{***}$   |
| <b>Sex</b>         |                                                      |                               |                               |                               |                         |
| Male               | 13.7 (10.4-17.5)                                     | 20.9*** (16.6-24.4)           | 26.9** (22.9-31.2)            | 24.3*** (20.6-28.3)           | $\chi^2_3=36.1^{***}$   |
| Female             | 22.2 (17.2-27.6)                                     | 22.6 (18.3-27.3)              | 27.2 (22.1-32.6)              | 24.8 <sup>†</sup> (20.0-29.9) | $\chi^2_3=8.2^*$        |
| <b>Age group</b>   |                                                      |                               |                               |                               |                         |
| 20-64              | 25.9 (18.5-33.9)                                     | 28.7 (22.9-34.7)              | 40.9 <sup>†</sup> (33.6-48.2) | 34.0 (26.9-41.4)              | $\chi^2_3=15.0^{**}$    |
| 65-74              | 14.5 (10.0-19.8)                                     | 24.5*** (19.6-29.7)           | 27.8 (22.3-33.6)              | 21.7*** (17.1-26.7)           | $\chi^2_3=20.3^{***}$   |
| 75-84              | 18.5 (13.7-24.1)                                     | 18.0 (13.8-22.7)              | 20.3 (15.4-25.8)              | 25.3*** (20.0-31.2)           | $\chi^2_3=9.7^*$        |
| 85+                | 6.4 (2.4-13.4)                                       | 7.9 <sup>†</sup> (3.7-14.5)   | 15.0 <sup>†</sup> (8.4-23.9)  | 15.7*** (9.5-23.8)            | $\chi^2_3=22.1^{***}$   |
| <b>Race</b>        |                                                      |                               |                               |                               |                         |
| White              | 17.2 (14.1-20.5)                                     | 21.0** (18.2-24.1)            | 28.9** (25.3-32.6)            | 24.0*** (20.8-27.4)           | $\chi^2_3=40.0^{***}$   |
| Black              | 18.7 (7.9-33.9)                                      | 22.0 (11.9-34.6)              | 18.0 (8.9-30.2)               | 18.6 (9.6-30.4)               | $\chi^2_3=4.1$          |
| Other <sup>b</sup> | 15.3 (5.9-30.1)                                      | 26.0 <sup>†</sup> (17.2-36.0) | 13.9 <sup>†</sup> (6.5-24.2)  | 33.1*** (22.7-43.9)           | $\chi^2_3=13.6^{**}$    |

CMML, chronic myelomonocytic leukemia

a: actuarial method used for observed survival. Ederer II method used for cumulative expected survival

b: statistically significant difference between Excess Hazard Ratio with previous cohort: \*:<.05; \*\*:<.01; \*\*\*:<.001

c: statistically significant difference between Excess Hazard Ratio between the 1st and 4th cohorts: <sup>†</sup>:<.05; <sup>††</sup>:<.01; <sup>†††</sup>:<.001

d: overall test of difference between Excess Hazard Ratio between the 4 cohorts: \*:<.05; \*\*:<.01; \*\*\*:<.001

## 4. Discussion

In this study, we report population-based epidemiologic data for CMML patients diagnosed in the US and reported to SEER between 2001 and 2017. We compared patient demographics, incidence and survival outcomes time-trends between 4 time-period cohorts and analyzed trends by sex, age group and race subgroups and lastly characterized long-term of survivors.

The first key finding from our study is the non-linear improvement in outcome over time for CMML in the US although relative survival (RS) and overall survival (OS) both improved significantly across the 4 time-periods analyzed. The most pronounced increase in RS occurred between the 2001-2004 and 2005-2009 time-periods with statistically significant improvement in 1-, 3- and 5-year RS of 56.3% vs 62.0%, 27.9% vs 33.4% and 17.2% vs 17.2% respectively. Period 2 corresponds to the timeframe when the FDA approved the hypomethylating agents azacitidine and decitabine for MDS and CMML, and this data supports the notion that the use of HMAs may have been a factor associated with improvement in outcomes for CMML patients.

It may be that the approval of HMAs acted as a natural experiment in our study, allowing us to also observe a plateauing trend in RS for time-periods 3 and 4, following the integration of HMAs into clinical practice. This possible contribution of HMA treatment to outcomes is consistent with the findings previously reported by Zeidan et al in their SEER-Medicare study [26]. They analyzed 1378 elderly CMML patients ( $\geq 66$  year) diagnosed between 2001 and 2011 and used propensity score to match patients pre and post HMA approvals. The median OS times were 11 and 17 months, respectively, translating into a significant improvement in OS and a 28% reduction in the risk of death in the matched-HMA cohort for this elderly patient population. Although our choice to categorize years of diagnosis in 4 categories was somewhat arbitrary and could impact our findings in terms of temporal trend for RS, the 1-, 3- and 5- RS

graphical analysis using year as continuous variable (Supplemental – Figures F1) support a similar observation.

The second key finding from our study is that the oldest group of patients ( $\geq 85+$  years) appear to have seen the largest improvement in RS, with a 3-year (95% CI) RS of 23.6% (16.6-31.7) compared to 8.4% (3.9-15.3),  $p < .001$  for 2001-2004 and 2014-2017 respectively. As patients in these age categories are not candidate for transplantation or intensive chemotherapy, we can possibly infer from our study that both HMAs and supportive care played a role in this improvement. Our study is not able to tease out the relative contribution of specific treatment strategies but the study from Zeidan et al. in elderly patients (Medicare enrollees in the United States) reported that only 18% of patients received HMAs in this cohort of patients aged  $> 66$  years between 2001 and 2011 [26] and that supportive care did not appear to explain improvement in OS. We also observed that the group of Black patients appear to have benefited less from the increase in RS and OS seen overall. The overall result of inferior outcomes associated with Black race are consistent with findings reported by Kirtane and Lee in their review of racial and ethnic disparities in hematological malignancies [37], where Black patients had inferior 5-year RS for Non-Hodgkin Lymphoma, Myeloma, Acute Myeloid Leukemia, Hodgkin Lymphoma and Acute Lymphoblastic Leukemia. To our knowledge, there is no CMML specific data that establishes a biological difference between races and outcomes. One could infer that the inferior outcomes associated with Black race may result to some extent from disparities in access to care (clinical trials, new therapy, transplantation, medical coverage), seen and described in other cancer types and hematological malignancies [37-40]. In AML and ALL, Black patients have worse survival outcomes and decreased likelihood of receiving definitive therapy with hematopoietic cell transplant (HCT) [41-43]. In another study between 2000-2010, authors found that patients diagnosed with ALL, AML, and MDS from areas with higher poverty rates were less likely to receive an unrelated donor allogeneic hematopoietic

transplantation [44]. In a study of 4659 AML patients from the Florida cancer registry, Byrne et al also reported that being African American, living in areas with higher poverty level and covered only by Medicaid were all associated with poor outcomes [45]. Alternatively, two studies from the SEER Program on patients with MDS found that Black patients had longer overall survival (OS) compared to White patients [46,47]. In our study, a significantly higher proportion of Black CMML patients were living in counties with median income  $\leq$ \$75,000/year compared to White and “Other” races, 80.7% vs 70.9 and 48.4% respectively. However, the Black population in our study was younger (60.9% vs 48.8% of patients  $\leq$ 74 years) with a higher proportion of female compared to White patients (45.3% vs 38.1%), both factors with favorable prognosis.

Our third notable observation was an increase in long-term survivors characterized by 5-year RS % and 5-year OS% for the entire population and most subgroups. To our knowledge, our study is the first to characterize in detail these encouraging, and highly relevant long-term efficacy data given the limited number of clinical trials dedicated to CMML and their limitation in terms of follow-up to allow a proper assessment of the “tail” at the end of survival curves.

The low incidence rate observed confirms CMML is a rare disease. Diagnosis patterns appeared stable during the study period for cases reported to SEER as microscopically confirmed. We confirm the male predominance of CMML and report a higher incidence in White patients compared to Black and “Other” race patients. This higher incidence in White compared to Black patients has been reported in other hematological malignancies such as AML, chronic lymphocytic leukemia, chronic myeloid leukemia, and chronic myelomonocytic leukemia [25] but to our knowledge, an explanation for these phenomena is not yet available.

Our study included more than 3900 patients, over a period of 17 years, indirectly sheds light on the impact of therapeutic progress on CMML, a disease associated with a poor prognosis and without specific or approved treatment options tailored to its distinct clinical and molecular background.

Our study has several limitations. The lack of therapeutic, clinical, cytogenetics, molecular, and prognosis score data in SEER did not allow the inclusion of these important variables in our analysis. Life tables according to our coding of race were not available in 2018 and 2019, and we used the same estimates as for 2017, hence the precision of the RS estimates may be slightly affected. Although HMAs have likely played an important role in the overall improvement in outcomes for CMML patients, other factors such as progress in supportive care, in the ease and outcome of transplant and the availability of new targeted agents [8] are also expected to have had a favorable impact on outcomes.

In summary, we have characterized a clinically meaningful and statistically significant improvement in survival outcomes over time for CMML cases diagnosed between 2001 to 2017 in the US, as reported to SEER. The most significant improvement was associated with the time-periods closely following the availability and increased integration of hypomethylating agents. This suggests that the increased incorporation of azacitidine and decitabine into the treatment of CMML may have had a significant impact on outcomes in the US. We also report a clinically important increase in long-term survivors and a notable improvement of survival for patients  $\geq 85+$  years of age, two important markers of therapeutic progress.

## **Acknowledgments**

### **Authors' contributions**

Arnaud Lesegretain: conceptualization, data acquisition & analysis, methodology, writing (drafting, review & editing), visualization

Andrew Brunner: conceptualization, writing (review and editing)

Aziz Benzohra: writing (review and editing)

Amir T. Fathi: conceptualization, writing (review & editing), visualization, supervision

## **Funding**

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## **Authors' disclosure of potential conflicts of interest**

Arnaud Lesegretain is an employee of Daiichi-Sankyo Inc. and has stock and stock options of Daiichi Sankyo. Aziz Benzohra was an employee of Daiichi-Sankyo during the research and writing of this manuscript and has since left the company. Andrew Brunner reports having received consulting fees from Acceleron, Agios, BMS/Celgene, Gilead, Keros Therapeutics, Novartis, Taiho, Takeda.

Amir Fathi reports having received research funding for clinical trials from Abbvie, BMS, Servier and consulting fees from Abbvie, Agios, Amgen, Astellas, Blueprint, BMS, Celgene, Daiichi Sankyo, EnClear, Forma, Genentech, Kite, Kura, Immunogen, Ipsen, Mablytics, Morphosys, Novartis, Orum, PureTech, Rige, Servier, Takeda, Pfizer, Trilium over the last 36 months.

## **References**

- [1] Itzykson R, Duchmann M, Lucas N, Solary E. CMML: Clinical and molecular aspects. *Int J Hematol.* 2017 Jun;105(6):711-719.
- [2] Patnaik MM, Tefferi A. Chronic Myelomonocytic leukemia: 2020 update on diagnosis, risk stratification and management. *Am J Hematol.* 2020 Jan;95(1):97-115
- [3] Geissler K, Jäger E, Barna A, Gurbisz M, Marschon R, Graf T, Graf E, Borjan B, Jilch R, Geissler C, Hoermann G, Esterbauer H, Schwarzingier I, Nösslinger T, Pfeilstöcker M, Tüchler H, Reisner R, Sliwa T, Keil F, Bettelheim P, Machherndl-Spandl S, Doleschal B, Zach O, Weltermann A, Heibl S, Thaler J, Zebisch A, Sill H, Stauder R, Webersinke G, Petzer A, Kusec R, Ulsperger E, Schneeweiss B, Berger J, Öhler L, Germing U, Sperr WR, Knöbl P, Jäger U, Valent P. The Austrian biodatabase for chronic myelomonocytic

- leukemia (ABCMML) : A representative and useful real-life data source for further biomedical research. *Wien Klin Wochenschr.* 2019 Sep;131(17-18):410-418.
- [4] Arber, D.A.; Orazi, A.; Hasserjian, R.; Thiele, J.; Borowitz, M.J.; Le Beau, M.M.; Bloomfield, C.D.; Cazzola, M.; Vardiman, J.W. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood* 2016, 127, 2391–2405.
- [5] Greenberg PL, Stone RM, Al-Kali A, Barta SK, Bejar R, Bennett JM, Carraway H, De Castro CM, Deeg HJ, DeZern AE, Fathi AT, Frankfurt O, Gaensler K, Garcia-Manero G, Griffiths EA, Head D, Horsfall R, Johnson RA, Juckett M, Klimek VM, Komrokji R, Kujawski LA, Maness LJ, O'Donnell MR, Pollyea DA, Shami PJ, Stein BL, Walker AR, Westervelt P, Zeidan A, Shead DA, Smith C. Myelodysplastic Syndromes, Version 2.2017, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw.* 2017 Jan;15(1):60-87.
- [6] Solary E, Itzykson R. How I treat chronic myelomonocytic leukemia. *Blood.* 2017 Jul 13;130(2):126-136.
- [7] Liapis K, Kotsianidis I. Approaching First-Line Treatment in Patients With Advanced CMML: Hypomethylating Agents or Cytotoxic Treatment? *Front Oncol.* 2021 Dec 13;11:801524.
- [8] Bewersdorf JP, Zeidan AM. Risk-Adapted, Individualized Treatment Strategies of Myelodysplastic Syndromes (MDS) and Chronic Myelomonocytic Leukemia (CMML). *Cancers (Basel).* 2021 Mar 31;13(7):1610.
- [9] Short NJ, Kantarjian H. Hypomethylating agents for the treatment of myelodysplastic syndromes and acute myeloid leukemia: Past discoveries and future directions. *Am J Hematol.* 2022 Dec;97(12):1616-1626.
- [10] Kantarjian H, Issa JP, Rosenfeld CS, Bennett JM, Albitar M, DiPersio J, Klimek V, Slack J, de Castro C, Ravandi F, Helmer R 3rd, Shen L, Nimer SD, Leavitt R, Raza A, Saba H. Decitabine improves patient outcomes in myelodysplastic syndromes: results of a phase III randomized study. *Cancer.* 2006 Apr 15;106(8):1794-803.

- [11] Silverman LR, Demakos EP, Peterson BL, Kornblith AB, Holland JC, Odchimar-Reissig R, Stone RM, Nelson D, Powell BL, DeCastro CM, Ellerton J, Larson RA, Schiffer CA, Holland JF. Randomized controlled trial of azacitidine in patients with the myelodysplastic syndrome: a study of the cancer and leukemia group B. *J Clin Oncol*. 2002 May 15;20(10):2429-40.
- [12] Fenaux P, Mufti GJ, Hellstrom-Lindberg E, Santini V, Finelli C, Giagounidis A, Schoch R, Gattermann N, Sanz G, List A, Gore SD, Seymour JF, Bennett JM, Byrd J, Backstrom J, Zimmerman L, McKenzie D, Beach C, Silverman LR; International Vidaza High-Risk MDS Survival Study Group. Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: a randomised, open-label, phase III study. *Lancet Oncol*. 2009 Mar;10(3):223-32.
- [13] Azacitidine FDA Package Insert:  
[https://www.accessdata.fda.gov/drugsatfda\\_docs/label/2022/050974s034lbl.pdf](https://www.accessdata.fda.gov/drugsatfda_docs/label/2022/050974s034lbl.pdf)
- [14] Decitabine FDA Package Insert:  
[https://www.accessdata.fda.gov/drugsatfda\\_docs/label/2018/021790s021lbl.pdf](https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/021790s021lbl.pdf)
- [15] Braun T, Itzykson R, Renneville A, de Renzis B, Dreyfus F, Laribi K, Bouabdallah K, Vey N, Toma A, Recher C, Royer B, Joly B, Vekhoff A, Lafon I, Sanhes L, Meurice G, Or  ar C, Preudhomme C, Gardin C, Ades L, Fontenay M, Fenaux P, Droin N, Solary E; Groupe Francophone des My  lodysplasies. Molecular predictors of response to decitabine in advanced chronic myelomonocytic leukemia: a phase 2 trial. *Blood*. 2011 Oct 6;118(14):3824-31.
- [16] Wijermans PW, R  ter B, Baer MR, Slack JL, Saba HI, L  bbert M. Efficacy of decitabine in the treatment of patients with chronic myelomonocytic leukemia (CMML). *Leuk Res*. 2008 Apr;32(4):587-91.

- [17] Fianchi L, Criscuolo M, Breccia M, Maurillo L, Salvi F, Musto P, Mansueto G, Gaidano G, Finelli C, Aloe-Spiriti A, Santini V, Greco M, Hohaus S, Leone G, Voso MT. High rate of remissions in chronic myelomonocytic leukemia treated with 5-azacytidine: results of an Italian retrospective study. *Leuk Lymphoma*. 2013 Mar;54(3):658-61.
- [18] Thorpe M, Montalvão A, Pierdomenico F, Moita F, Almeida A. Treatment of chronic myelomonocytic leukemia with 5-Azacitidine: a case series and literature review. *Leuk Res*. 2012 Aug;36(8):1071-3.
- [19] Adès L, Sekeres MA, Wolffromm A, Teichman ML, Tiu RV, Itzykson R, Maciejewski JP, Dreyfus F, List AF, Fenaux P, Komrokji RS. Predictive factors of response and survival among chronic myelomonocytic leukemia patients treated with azacitidine. *Leuk Res*. 2013 Jun;37(6):609-13. doi: 10.1016/j.leukres.2013.01.004.
- [20] Costa R, Abdulhaq H, Haq B, Shadduck RK, Latsko J, Zenati M, Atem FD, Rossetti JM, Sahovic EA, Lister J. Activity of azacitidine in chronic myelomonocytic leukemia. *Cancer*. 2011 Jun 15;117(12):2690-6.
- [21] Wong E, Seymour JF, Kenealy M, Westerman D, Herbert K, Dickinson M. Treatment of chronic myelomonocytic leukemia with azacitidine. *Leuk Lymphoma*. 2013 Apr;54(4):878-80.
- [22] Drummond MW, Pocock C, Boissinot M, Mills J, Brown J, Cauchy P, Cross NC, Hartley S, Kell J, Szubert A, Cockerill PN, Bowen DT. A multi-centre phase 2 study of azacitidine in chronic myelomonocytic leukaemia. *Leukemia*. 2014 Jul;28(7):1570-2. doi: 10.1038/leu.2014.85. Epub 2014 Feb 26. PMID: 24569776.
- [23] Tantravahi SK, Szankasi P, Khorashad JS, Dao KH, Kovacsovics T, Kelley TW, Deininger MW. A phase II study of the efficacy, safety, and determinants of response to 5azacitidine (Vidaza®) in patients with chronic myelomonocytic leukemia. *Leuk Lymphoma*. 2016 Oct;57(10):2441-4.
- [24] Santini V, Allione B, Zini G, Gioia D, Lunghi M, Poloni A, Cilloni D, Sanna A, Masiera E,

- Ceccarelli M, Abdel-Wahab O, Terenzi A, Angelucci E, Finelli C, Onida F, Pelizzari A, Ferrero D, Saglio G, Figueroa M, Levis A. A phase II, multicentre trial of decitabine in higher-risk chronic myelomonocytic leukemia. *Leukemia*. 2018 Feb;32(2):413-418.
- [25] Pleyer L, Germing U, Sperr WR, Linkesch W, Burgstaller S, Stauder R, Girschikofsky M, Schreder M, Pfeilstocker M, Lang A, Sliwa T, Geissler D, Schlick K, Placher-Sorko G, Theiler G, Thaler J, Mitrovic M, Neureiter D, Valent P, Greil R. Azacitidine in CMML: matched-pair analyses of daily-life patients reveal modest effects on clinical course and survival. *Leuk Res*. 2014 Apr;38(4):475-83.
- [26] Zeidan AM, Hu X, Long JB, Wang R, Ma X, Podoltsev NA, Huntington SF, Gore SD, Davidoff AJ. Hypomethylating agent therapy use and survival in older patients with chronic myelomonocytic leukemia in the United States: A large population-based study. *Cancer*. 2017 Oct 1;123(19):3754-3762.
- [27] Surveillance, Epidemiology, and End Results (SEER) Program ([www.seer.cancer.gov](http://www.seer.cancer.gov))
- [28] M.A. Duggan, W.F. Anderson, S. Altekruse, L. Penberthy, M.E. Sherman, The Surveillance, Epidemiology, and End Results (SEER) program and pathology: toward strengthening the critical relationship, *Am. J. Surg. Pathol.* 40(12) (2016) e94–e102.
- [29] Products - Life Tables - Homepage ([cdc.gov](http://cdc.gov))
- [30] Surveillance, Epidemiology, and End Results (SEER) Program ([www.seer.cancer.gov](http://www.seer.cancer.gov))  
SEER\*Stat Database: Incidence - SEER Research Plus Data, 17 Registries, Nov 2021 Sub (2000-2019) - Linked To County Attributes - Time Dependent (1990-2019) Income/Rurality, 1969-2020 Counties, National Cancer Institute, DCCPS, Surveillance Research Program, released April 2022, based on the November 2021 submission. Accessed November 21st 2022.
- [31] Surveillance Research Program, National Cancer Institute SEER\*Stat software – ([www.seer.cancer.gov/seerstat](http://www.seer.cancer.gov/seerstat)) version 8.4.0.1. Accessed November 21st 2022.

- [32] Ederer F, Axtell LM, Cutler SJ. The relative survival rate: a statistical methodology. *Natl Cancer Inst Monogr.* 1961 Sep;6:101-21.
- [33] Henson DE, Ries LA. The relative survival rate. *Cancer.* 1995 Nov 15;76(10):1687-8.
- [34] Ederer F, Heise H. Instructions to IBM 650 programmers in processing survival computations. Methodological note no. 10. Bethesda, MD: End Results Evaluation Section, National Cancer Institute; 1959
- [35] Dickman, P. W., & Coviello, E. (2015). Estimating and Modeling Relative Survival. *The Stata Journal*, 15(1), 186–215.
- [36] Dickman PW, Sloggett A, Hills M, Hakulinen T. Regression models for relative survival. *Stat Med.* 2004;23(1):51-64.
- [37] K. Kirtane, S.J. Lee, Racial and ethnic disparities in hematologic malignancies, *Blood* 130(15) (2017) 1699–1705.
- [38] American Cancer Society, *Cancer Facts & Figures for African Americans 2019–2021*, 2019. <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-andstatistics/cancer-facts-and-figures-for-african-americans/cancer-facts-and-figures-forafrican-americans-2019-2021.pdf>. (Accessed November 21st, 2022).
- [39] L.A. Siminoff, L. Ross, Access and equity to cancer care in the USA: a review and assessment, *Postgrad. Med. J.* 81(961) (2005) 674–679.
- [40] American Cancer Society, *Cancer Facts & Figures for African Americans 2019–2021*, 2019. <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-andstatistics/cancer-facts-and-figures-for-african-americans/cancer-facts-and-figures-forafrican-americans-2019-2021.pdf>. (Accessed November 21st, 2022).
- [41] B. Bhatnagar, J. Kohlschmidt, K. Mrózek, Q. Zhao, J.L. Fisher, D. Nicolet, C.J. Walker, A.S. Mims, C. Oakes, B. Giacomelli, S. Orwick, I. Boateng, J.S. Blachly, S.E. Maharry, A.J.

- Carroll, B.L. Powell, J.E. Kolitz, R.M. Stone, J.C. Byrd, E.D. Paskett, A. de la Chapelle, R. Garzon, A.K. Eisfeld, Poor survival and differential impact of genetic features of Black patients with acute myeloid leukemia, *Cancer Discov.* 11(3) (2021) 626–637.
- [42] Bhatt VR, Shostrom V, Gundabolu K, Armitage JO. Utilization of initial chemotherapy for newly diagnosed acute myeloid leukemia in the United States. *Blood Adv.* 2018 Jun 12;2(11):1277-1282
- [43] Pulte D, Redaniel MT, Jansen L, Brenner H, Jeffreys M. Recent trends in survival of adult patients with acute leukemia: overall improvements, but persistent and partly increasing disparity in survival of patients from minority groups. *Haematologica.* 2013 Feb;98(2):222-9
- [44] Paulson K, Brazauskas R, Khera N, He N, Majhail N, Akpek G, Aljurf M, Buchbinder D, Burns L, Beattie S, Freytes C, Garcia A, Gajewski J, Hahn T, Knight J, LeMaistre C, Lazarus H, Szwajcer D, Seftel M, Wirk B, Wood W, Saber W. Inferior Access to Allogeneic Transplant in Disadvantaged Populations: A Center for International Blood and Marrow Transplant Research Analysis. *Biol Blood Marrow Transplant.* 2019 Oct;25(10):2086-2090
- [45] Byrne MM, Halman LJ, Koniaris LG, Cassileth PA, Rosenblatt JD, Cheung MC. Effects of poverty and race on outcomes in acute myeloid leukemia. *Am J Clin Oncol.* 2011 Jun;34(3):297-304.
- [46] S.Y. Goksu, M. Ozer, B.B. Goksu, R. Wang, J. Khatib, P.A. Patel, M. Vusirikala, S. Cole, A. Seyhanli, R.H. Collins, S. Chung, A.M. Zeidan, Y.F. Madanat, The impact of race and ethnicity on outcomes of patients with myelodysplastic syndromes: a population-based analysis, *Leuk. Lymphoma* 63(7) (2022) 1651–1659.
- [47] Lesegretain A, Brunner A, King AJ, Laadem A, Fell G, Fathi AT. Comparison of demographics, disease characteristics, and outcomes between Black and White patients

with myelodysplastic syndromes: A population-based study. Leuk Res. 2022 Dec  
26;125:107006

## SUPPLEMENTARY TABLES PAPER #2

### Appendix A. Supplementary data

#### Supplemental Table S1

Age-adjusted incidence for adult CMML patients ( $\geq 20$  years) by period of diagnosis between 2001-2017 - SEER (n=3991)

|                  | 2001-2004        |                    | 2005-2009        |                     | 2010-2013        |                     | 2014-2017        |                       |
|------------------|------------------|--------------------|------------------|---------------------|------------------|---------------------|------------------|-----------------------|
|                  | IR <sup>a</sup>  | IRR <sup>b</sup>   | IR <sup>a</sup>  | IRR <sup>b</sup>    | IR <sup>a</sup>  | IRR <sup>b</sup>    | IR <sup>a</sup>  | IRR <sup>b</sup>      |
|                  | (95%CI)          | (95% CI)           | (95%CI)          | (95% CI)            | (95%CI)          | (95% CI)            | (95%CI)          | (95% CI)              |
| <b>Total</b>     | 0.4<br>(0.4-0.4) | Ref.               | 0.4<br>(0.4-0.4) | 1.05<br>(0.96-1.15) | 0.4<br>(0.4-0.4) | 0.97<br>(0.88-1.07) | 0.4<br>(0.4-0.4) | 1.05<br>(0.96-1.16)   |
| <b>Sex</b>       |                  |                    |                  |                     |                  |                     |                  |                       |
| Male             | 0.6<br>(0.5-0.6) | Ref.               | 0.6<br>(0.6-0.7) | Ref.                | 0.6<br>(0.5-0.6) | Ref.                | 0.6<br>(0.5-0.6) | Ref.                  |
| Female           | 0.3<br>(0.2-0.3) | 0.46** (0.40-0.53) | 0.3<br>(0.2-0.3) | 0.45** (0.39-0.50)  | 0.3<br>(0.2-0.3) | 0.47** (0.41-0.53)  | 0.3<br>(0.3-0.3) | 0.50**<br>(0.43-0.56) |
| <b>Age Group</b> |                  |                    |                  |                     |                  |                     |                  |                       |
| 20-64            | 0.1<br>(0.1-0.1) | Ref.               | 0.1<br>(0.1-0.1) | Ref.                | 0.1<br>(0.1-0.1) | Ref.                | 0.1<br>(0.1-0.1) | Ref.                  |
| 65-74            | 1.3<br>(1.2-1.5) | 18.6** (15.0-23.3) | 1.4<br>(1.2-1.5) | 15.1** (12.7-18.0)  | 1.2<br>(1.1-1.4) | 15.7** (12.9-19.0)  | 1.4<br>(1.2-1.5) | 16.7** (14.1-19.9)    |
| 75-84            | 2.4<br>(2.2-2.7) | 34.3** (27.8-42.4) | 2.6<br>(2.3-2.8) | 28.5** (24.2-33.6)  | 2.5<br>(2.2-2.8) | 31.1** (25.8-37.5)  | 2.7<br>(2.4-2.9) | 32.6**<br>(27.3-38.9) |
| 85+              | 3.1<br>(2.6-3.6) | 43.0** (33.5-55.3) | 2.7<br>(2.3-3.1) | 29.4** (24.0-36.1)  | 2.4<br>(2.1-2.9) | 30.8** (24.5-38.5)  | 2.9<br>(2.5-3.3) | 35.3**<br>(28.7-43.3) |
| <b>Race</b>      |                  |                    |                  |                     |                  |                     |                  |                       |
| White            | 0.4<br>(0.4-0.5) | Ref.               | 0.4<br>(0.4-0.5) | Ref.                | 0.4<br>(0.4-0.5) | Ref.                | 0.4<br>(0.4-0.5) | Ref.                  |
| Black            | 0.3<br>(0.2-0.4) | 0.72* (0.52-0.97)  | 0.3<br>(0.2-0.4) | 0.71** (0.54-0.92)  | 0.3<br>(0.2-0.4) | 0.71* (0.53-0.93)   | 0.3<br>(0.2-0.4) | 0.65**<br>(0.50-0.84) |
| Other            | 0.3<br>(0.3-0.4) | 0.66* (0.47-0.90)  | 0.4<br>(0.3-0.5) | 0.87 (0.70-1.08)    | 0.3<br>(0.2-0.3) | 0.63** (0.48-0.82)  | 0.4<br>(0.3-0.4) | 0.81*<br>(0.65-0.99)  |

CMML, chronic myelomonocytic leukemia; IR: Incidence Rate; IRR: Incidence Rate Ratio; CI, Confidence Interval; SEER, Surveillance, Epidemiology, and End Results

a: Rates are per 100,000 and age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130) standard; Confidence intervals (Tiwari mod) are 95% for rates. b: Test of difference; \*:  $p < 0.05$ ; \*\*:  $p < .001$

#### Supplemental Table S2

Relative Survival trend CMML patients diagnosed 2001-2017- SEER (n=3929)

|                    | Relative Survival <sup>a,b,c</sup> |                    |                  |                     | Difference <sup>d</sup> |
|--------------------|------------------------------------|--------------------|------------------|---------------------|-------------------------|
|                    | 2001-2004                          | 2005-2009          | 2010-2013        | 2014-2017           |                         |
| 1-year, % (95% CI) | 56.3 (52.6-59.9)                   | 62.0** (59.0-65.0) | 65.1 (61.7-68.3) | 65.4*** (62.4-68.3) | $\chi^2_3=22.8^{***}$   |
| 3-year, % (95% CI) | 27.9 (24.5-31.4)                   | 33.4** (30.4-36.4) | 37.0 (33.6-40.5) | 36.9*** (33.7-40.1) | $\chi^2_3=37.6^{***}$   |

CMML, chronic myelomonocytic leukemia;  $\chi^2_{df}$ , Chi-Squared test/Degree of freedom; SEER, Surveillance, Epidemiology, and End Results.  
a: actuarial method used for observed survival. Ederer II method used for cumulative expected survival  
b: statistically significant difference between Excess Hazard Ratio with previous cohort: \*: <.05; \*\*: <.01; \*\*\*: <.001  
c: statistically significant difference between Excess Hazard Ratio the 1st and 4th cohorts: †: <.05; ††: <.01; †††: <.001  
d: overall test of difference between Excess Hazard Ratio between the 4 cohorts: \*: <.05; \*\*: <.01; \*\*\*: <.001

### Supplemental Table S3

3-year Relative Survival with/without age-standardization CMML patients 2001-2017- SEER (n=3929).

|                | Unadjusted <sup>a,b</sup> % (95% CI) |                                 | Age-Standardized <sup>a,b,c</sup> % (95% CI) |                                 |
|----------------|--------------------------------------|---------------------------------|----------------------------------------------|---------------------------------|
|                | 2001-2004                            | 2014-2017                       | 2001-2004                                    | 2014-2017                       |
| <b>Overall</b> | 27.9 (24.5-31.4)                     | 36.9 <sup>†††</sup> (33.7-40.1) | 27.4 (24.1-30.8)                             | 36.6 <sup>†††</sup> (33.4-39.7) |
| <b>Sex</b>     |                                      |                                 |                                              |                                 |
| Male           | 27.4 (23.1-32.0)                     | 35.6 <sup>†††</sup> (31.7-39.7) | 26.9 (22.7-31.2)                             | 35.3 <sup>†††</sup> (31.4-39.3) |
| Female         | 28.6 (23.3-34.1)                     | 38.9 <sup>†††</sup> (33.7-44.1) | 27.7 (22.6-32.9)                             | 38.6 <sup>†††</sup> (33.4-43.7) |
| <b>Race</b>    |                                      |                                 |                                              |                                 |
| White          | 28.3 (24.7-32.0)                     | 36.9 <sup>†††</sup> (33.5-40.4) | 27.4 (24.0-31.0)                             | 36.6 <sup>†††</sup> (33.2-40.0) |
| Black          | 24.6 (12.5-39.6)                     | 29.1 (18.3-41.0)                | 28.1 (14.1-43.8)                             | 28.6 (18.2-39.8)                |
| Other          | 25.1 (12.7-39.9)                     | 42.0 <sup>†</sup> (31.6-52.3)   | 26.1 (13.7-40.4)                             | 42.1 <sup>†</sup> (31.5-52.3)   |

CMML, chronic myelomonocytic leukemia; SEER, Surveillance, Epidemiology, and End Results.

<sup>a</sup> actuarial method used for observed survival. Ederer II method used for cumulative expected survival <sup>b</sup> statistically significant difference between Excess Hazard Ratio for the 1st and 4th time-periods: †: <.05; ††: <.01; †††: <.001 <sup>c</sup> age-standardized using 2014-2017 age group proportion as standard weight

### Supplemental Table S4

Overall survival for CMML patients diagnosed 2001-2017- SEER (n=3929).

|                   | median Overall Survival, months (95% CI) |            |            |            |
|-------------------|------------------------------------------|------------|------------|------------|
|                   | 2001-2004                                | 2005-2009  | 2010-2013  | 2014-2017  |
| <b>Overall</b>    | 13 (12-15)                               | 16 (15-18) | 19 (17-21) | 21 (19-23) |
| <b>Sex</b>        |                                          |            |            |            |
| Male              | 13 (11-15)                               | 16 (15-18) | 19 (17-23) | 20 (17-23) |
| Female            | 14 (11-17)                               | 17 (13-20) | 18 (15-21) | 21 (17-26) |
| <b>Age groups</b> |                                          |            |            |            |
| 20-64             | 19 (12-27)                               | 19 (15-25) | 32 (22-49) | 29 (21-43) |
| 65-74             | 15 (12-19)                               | 20 (17-26) | 24 (19-29) | 22 (19-25) |
| 75-84             | 15 (12-18)                               | 16 (12-19) | 16 (12-19) | 23 (18-26) |
| 85+               | 6 (3-7)                                  | 8 (6-11)   | 11 (8-16)  | 9 (6-14)   |
| <b>Race</b>       |                                          |            |            |            |
| White             | 14 (12-15)                               | 16 (14-18) | 19 (17-22) | 21 (18-23) |
| Black             | 5 (2-15)                                 | 14 (9-22)  | 19 (13-26) | 14 (6-24)  |

|       |           |            |           |            |
|-------|-----------|------------|-----------|------------|
| Other | 14 (6-19) | 22 (16-32) | 16 (9-20) | 25 (17-44) |
|-------|-----------|------------|-----------|------------|

CMML, chronic myelomonocytic leukemia; SEER, Surveillance, Epidemiology, and End Results.

## Supplemental Table S5

5-year Overall Survival % for CMML patients diagnosed 2001-2017- SEER (n=3929).

|                    | 5-year Overall Survival <sup>a,b</sup> % (95% CI) |                                 |                                 |                                    | Difference <sup>c</sup> |
|--------------------|---------------------------------------------------|---------------------------------|---------------------------------|------------------------------------|-------------------------|
|                    | 2001-2004                                         | 2005-2009                       | 2010-2013                       | 2014-2017                          |                         |
| <b>Total</b>       | 13.2 (10.9-15.6)                                  | 17.3 <sup>**</sup> (15.1-19.6)  | 22.2 <sup>*</sup> (19.6-25.0)   | 22.3 <sup>†††</sup> (19.2-25.5)    | $\chi^2_3=43.7^{***}$   |
| <b>Sex</b>         |                                                   |                                 |                                 |                                    |                         |
| Male               | 10.2 (7.7-13.1)                                   | 16.9 <sup>**</sup> (14.2-21.9)  | 22.1 <sup>*</sup> (18.8-25.6)   | 23.1 <sup>†††</sup> (19.3-27.1)    | $\chi^2_3=37.1^{***}$   |
| Female             | 17.6 (13.7-22.0)                                  | 18.0 (14.5-21.9)                | 22.5 (18.8-25.6)                | 20.7 <sup>††</sup> (15.6-26.4)     | $\chi^2_3=9.6^*$        |
| <b>Age group</b>   |                                                   |                                 |                                 |                                    |                         |
| 20-64              | 24.8 (17.7-32.5)                                  | 27.4 (21.8-33.2)                | 39.6 <sup>**</sup> (32.6-46.6)  | 36.4 <sup>†</sup> (28.5-44.3)      | $\chi^2_3=13.9^{***}$   |
| 65-74              | 11.2 (7.6-15.7)                                   | 21.9 <sup>***</sup> (17.5-26.6) | 25.1 (20.1-30.3)                | 22.3 <sup>†</sup> (17.1-27.9)      | $\chi^2_3=19.7^{***}$   |
| 75-84              | 13.7 (10.1-17.8)                                  | 12.8 (9.8-16.3)                 | 15.7 (11.9-19.9)                | 20.5 <sup>††</sup> (15.2-26.3)     | $\chi^2_3=12.1^{**}$    |
| 85+                | 3.9 (1.5-8.2)                                     | 4.9 <sup>*</sup> (2.3-9.0)      | 8.1 (4.3-13.4)                  | 10.4 <sup>†††</sup> (5.5-17.1)     | $\chi^2_3=15.4^{**}$    |
| <b>Race</b>        |                                                   |                                 |                                 |                                    |                         |
| White              | 13.2 (10.8-15.8)                                  | 16.8 <sup>*</sup> (14.5-19.2)   | 23.6 <sup>***</sup> (20.7-26.7) | 21.3 <sup>***†††</sup> (18.0-24.8) | $\chi^2_3=40.4^{***}$   |
| Black              | 13.8 (5.6-25.5)                                   | 17.6 (9.5-27.9)                 | 15.0 (7.4-25.1)                 | 20.5 (11.25-31.7)                  | $\chi^2_3=2.4$          |
| Other <sup>b</sup> | 12.4 (4.6-24.4)                                   | 22.4 <sup>*</sup> (14.6-32.3)   | 12.3 <sup>*</sup> (5.8-21.5)    | 34.6 <sup>***††</sup> (22.3-47.2)  | $\chi^2_3=13.3^{**}$    |

CMML, chronic myelomonocytic leukemia; SEER, Surveillance, Epidemiology, and End Results.

a: log-rank test for the survival function - difference with previous cohort: \*: <.05; \*\*: <.01; \*\*\*: <.001

b: log-rank test for the survival function - difference between the 1st and 4th cohorts: †: <.05; ††: <.01; †††: <.001

c: test between the 4 periods: \*: <.05; \*\*: <.01; \*\*\*: <.001

## Supplemental Table S6

Comparison of Study population and Excluded population<sup>a</sup> - CMML patients diagnosed 2001-2017- SEER (n=5918)

|                           | Study Population |       | Excluded Population     |       |
|---------------------------|------------------|-------|-------------------------|-------|
|                           | N                | %     | N                       | %     |
| <b>N</b>                  | 3,929            |       | 1,989                   |       |
| <b>Median OS (95% CI)</b> | 17 (16-18)       |       | 14 (13-15) <sup>b</sup> |       |
| <b>Sex</b>                |                  |       |                         |       |
| Male                      | 2,429            | 61.8% | 1,283                   | 64.5% |
| Female                    | 1,500            | 38.2% | 706                     | 35.5% |
| <b>Median age - IQR</b>   | 75 (67-82)       |       | 78 (71-84)              |       |
| <b>Age Group</b>          |                  |       |                         |       |

|             |                  |       |                  |       |
|-------------|------------------|-------|------------------|-------|
| 20-64       | 759              | 19.3% | 212              | 10.7% |
| 65-74       | 1,196            | 30.4% | 495              | 24.9% |
| 75-84       | 1,372            | 34.9% | 865              | 43.5% |
| 85+         | 602              | 15.3% | 417              | 21.0% |
| <b>Race</b> |                  |       |                  |       |
| White       | 3,378            | 86.0% | 1,772            | 89.1% |
| Black       | 243              | 6.2%  | 114              | 5.7%  |
| Other       | 308 <sup>c</sup> | 7.8%  | 103 <sup>d</sup> | 5.2%  |

CMML, chronic myelomonocytic leukemia; IQR: interquartile range; SEER, Surveillance, Epidemiology, and End Results

a: diagnostic not microscopically confirmed: 408; not first primary only: 1556; diagnostic only from death certificate and autopsy: 23; alive with no survival time: 2.

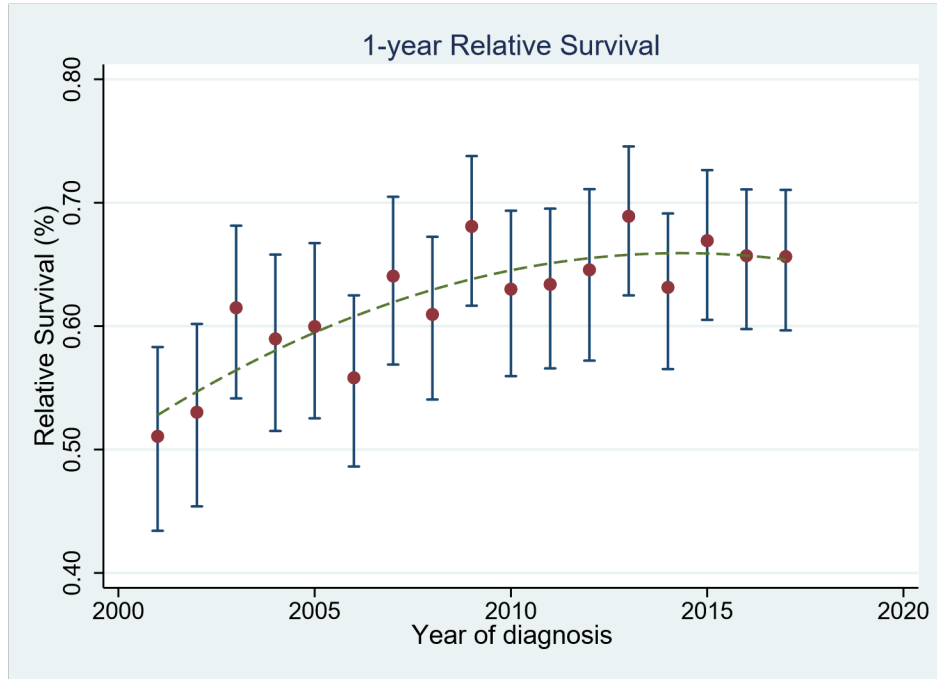
b: excludes 23 subjects with survival month=9999 and 2 subjects with zero months of survival

c: includes: n=291 Asian/Pacific-Islander, American Indians/Alaska Natives; n=17: unknown

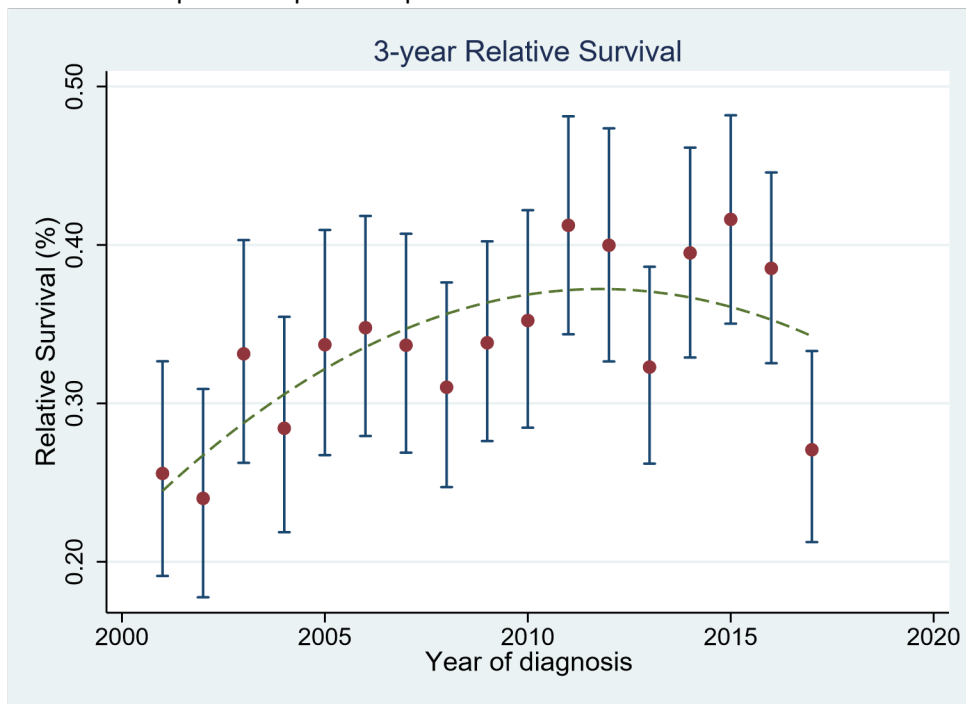
d: includes: n=99 Asian/Pacific-Islander, American Indians/Alaska Natives; n=4: unknown

## Supplemental Figure F1

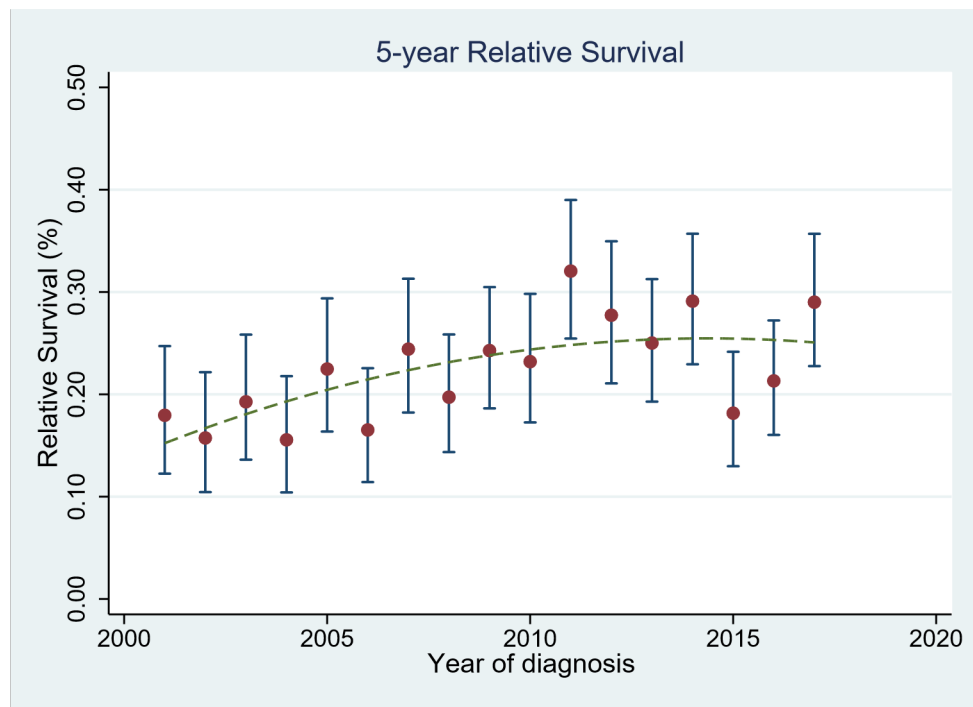
Relative Survival trend CMML patients diagnosed 2001-2017- SEER (n=3929).



Dashed line represents quadratic prediction



Dashed line represents quadratic prediction



Dashed line represents quadratic prediction

## PAPER SUMMARIES

**PAPER #1.** We reported an association between race and overall survival favoring Black patients. Given the literature consistently pointing to inferior outcomes for Black patients for many cancers, this was unexpected in the context of no known biological differences between races associated with outcomes in MDS. After controlling for all measured confounders reported to SEER, the association was maintained although attenuated. Outcomes also favored Black patients in a matching analysis. The main novel finding of our work is that the association appears to change in direction and magnitude based on various MDS histological subtypes as reported to SEER. RA, RARS, MDS-NOS favored Black patients, while RAEB favored White patients. There was no difference for RCMD and del5(q) MDS. Our paper also places our data within the important context of inherent limitations in the SEER database, and points to the challenges of disease classification in cancer registries for a disease such as MDS, where diagnosis is complex (resulting in ~50% of cases having a “provisional diagnosis” in SEER) and to potential cases misclassification impacting the cohort of Black patients [7]. Lastly, we confirmed the importance of the cardiovascular risk as the 2<sup>nd</sup> most frequent causes of death for all patients with MDS, in particular for Black patients and concluded on the importance of continued monitoring of this risk in clinical practice.

**PAPER #2.** Our study reported a clinically meaningful and statistically significant improvement in survival outcomes over time for CMML between 2001 to 2017 in the US, as reported to SEER. We used net survival, in a relative survival framework, and overall survival as measures and observed that the most pronounced improvement was associated with the time-periods following the availability of hypomethylating agents.

We also made interesting and clinically relevant observations with most sub-groups of patients benefiting, with patients  $\geq 85+$  years of age appearing to have benefited the most in relative terms, while Black race was not associated with an increase in survival over the study period. We further reported an increase in long-term survivors, a clinically useful observation, not previously reported, that is not readily available from the limited clinical trials dedicated to CMML to date and their short follow-up.

## **DISCUSSION, PERSPECTIVE, AND POSSIBLE FUTURE DIRECTIONS**

**PAPER #1.** Our study reported unexpected results of longer survival associated with race, favoring Black patients. Our main finding was consistent with another recent MDS study based on SEER and with trends from other smaller studies. However, in the absence of a biological hypothesis, we need to interpret our observations with caution. The NCI SEER database is a recognized tool for clinical cancer investigation with some robust quality control processes that makes it a worldwide reference. The database is particularly useful in the investigation of rare diseases, long-term trend and national demographic differences in the diagnosis, treatment, and outcome of cancer. It can also be used for hypotheses generating with acknowledgment of its inherent limitations. In particular, the SEER dataset does not contain information on clinical, laboratory values, molecular, therapy, co-morbidity, disease specific prognostic classification for patients diagnosed with MDS that would allow us to interrogate the data more rigorously and control for potential residual confounders. Moreover, a selection bias is possible in the case ascertainment between the Black and White cohorts of patients due to known barriers and inequalities in access to care in the US healthcare system. Also, race as self-reported in SEER is more a social construct than a biological entity.

Our study invites further research. Initial steps could be to replicate these results in other cancer registries in the US and to look more specifically in the elderly group of higher-risk MDS (RAEB) and also compare the Black cohort with other races than White. A further initiative could be to interrogate RCTs and more comprehensively annotated dataset such as the “National Myelodysplastic Syndromes Study” from the NHLBI. Our findings could also represent a “call to action” for the MDS community to

improve the accuracy of MDS classification at diagnosis to reduce the proportion of patients categorized as “MDS-NOS” and improve the utility of population-based disease surveillance.

**PAPER #2.** Our study sheds an encouraging light on improving outcomes for CMML, an uncommon disease with generally poor outcomes. Such a trend of observed progress, at the population level, would be welcome news for patients and their oncologists.

Consistent with others, we hypothesized and provided supporting evidence for the important role played by hypomethylating agents in this positive trend. Important limitations of our study stems from a lack of information on clinical, laboratory, molecular, disease specific prognostic classification, co-morbidity and other therapeutic interventions (investigational agents, transplants, supportive care) for patients with CMML in SEER and the resulting inability to control for these possible additional confounders. Additionally, life tables according to our coding of race were not available in 2018 and 2019, and we used the same estimates as for 2017, hence the precision of the RS estimates may be slightly affected, although we don’t believe it affects the overall findings of the study as changes in mortality in the general population tend to be very stable year to year (our study did not include the Covid-19 pandemic years where changes in mortality rate in the general population and various age groups may have been more significant).

In our study, we defined 4 time-based cohorts based on year of diagnosis; period 1: 2001-2004; period 2: 2005-2009; period 3: 2010-2013; period 4: 2014-2017. Based on the FDA approvals of Azacitidine and Decitabine in May 2004 and May 2006 respectively, we considered that period 2 corresponded to the time where HMAs

effectively became part of the standard of care, with the typical “early adopters” from tertiary and academic centers upon FDA approval and a progressive broadening of use thereafter to other centers and to the community. Our choice of breaking down the study in 4 periods of approximately similar size allowed us to analyze both an overall temporal trend over 17 years including 2 “post-HMA approval” time periods but also a comparison of the 2 periods proximal to FDA approval, period 1 immediately before and period 2 as the period of adoption of HMAs. An alternative approach to compare pre and post HMA could have been to study years 2001-2003 with 2007-2010 for example, thereby excluding the time corresponding to the approval and adoption in clinical practice in order to compare 2 “steady-state” periods. This was done by Zeidan et al in their SEER-Medicare published in 2017 study where they reported the first evidence supporting an association between HMAs and improved outcomes. Our analysis added to their study and showed that most of the improvement in survival appears to stem from the period proximal to the FDA approvals of HMA and tend to plateau thereafter, providing this new perspective to the literature.

A follow-on study with a larger and more comprehensive database (such as SEER-Medicare) to decipher receipt and type of treatment may further refine the findings of our study, including potential differential outcomes by race.

## BIBLIOGRAPHY

- [1] K. Kirtane, S.J. Lee, Racial and ethnic disparities in hematologic malignancies, *Blood* 130(15) (2017) 1699–1705.
- [2] L.A. Siminoff, L. Ross, Access and equity to cancer care in the USA: a review and assessment, *Postgrad. Med. J.* 81(961) (2005) 674–679.
- [3] American Cancer Society, Cancer Facts & Figures for African Americans 2019–2021, 2019. <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-andstatistics/cancer-facts-and-figures-for-african-americans/cancer-facts-and-figures-forafrican-americans-2019-2021.pdf>. (Accessed August 8, 2022).
- [4] American Society of Hematology, Message from the ASH President on Diversity, Equity, and Inclusion., 2020. <https://www.hematology.org/newsroom/press-releases/2020/messagefrom-the-ash-president-on-diversity-equity-and-inclusion>. (Accessed September 2, 2020).
- [5] E. Goss, A.M. Lopez, C.L. Brown, D.S. Wollins, O.W. Brawley, D. Raghavan, American Society of Clinical Oncology policy statement: disparities in cancer care, *J. Clin. Oncol.* 27(17) (2009) 2881–2885.
- [6] Surveillance, Epidemiology, and End Results (SEER) Program ([www.seer.cancer.gov](http://www.seer.cancer.gov)) SEER\*Stat Database, Number of Persons by Race and Hispanic Ethnicity for SEER Participants (2020 Census Data). .
- [7] Lesegretain A, Brunner A, King AJ, Laadem A, Fell G, Fathi AT. Comparison of demographics, disease characteristics, and outcomes between Black and White patients with myelodysplastic syndromes: A population-based study. *Leuk Res.* 2023;125:107006.
- [8] Patnaik MM, Tefferi A. Chronic Myelomonocytic leukemia: 2020 update on diagnosis, risk stratification and management. *Am J Hematol.* 2020 Jan;95(1):97-115.