



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Review

When hematologic malignancies meet COVID-19 in the United States: Infections, death and disparities

QuanQiu Wang^a, Nathan A. Berger^{b,c,**}, Rong Xu^{a,c,*}

^a Center for Artificial Intelligence in Drug Discovery, School of Medicine, Case Western Reserve University, Cleveland, OH, USA

^b Center for Science, Health, and Society, School of Medicine, Case Western Reserve University, Cleveland, OH, USA

^c Case Comprehensive Cancer Center, School of Medicine, Case Western Reserve University, Cleveland, OH, USA



ARTICLE INFO

Keywords:

COVID-19
Hematologic malignancies
Risk
Health outcomes
Disparity

ABSTRACT

Scientific data is limited on the risks, adverse outcomes and racial disparities for COVID-19 illness in individuals with hematologic malignancies in the United States. To fill this void, we screened and analyzed a nation-wide database of patient electronic health records (EHRs) of 73 million patients in the US (up to September 1st) for COVID-19 and eight major types of hematologic malignancies. Patients with hematologic malignancies had increased odds of COVID-19 infection compared with patients without hematologic malignancies for both all-time diagnosis (malignancy diagnosed in the past year or prior) (adjusted Odds ratio or AOR: 2.27 [2.17–2.36], $p < 0.001$) and recent diagnosis (malignancy diagnosed in the past year) (AOR: 11.91 [11.31–12.53], $p < 0.001$), with strongest effect for recently diagnosed acute lymphoid leukemia (AOR: 31.03 [25.87–37.27], $p < 0.001$), essential thrombocythemia (AOR: 20.65 [19.10–22.32], $p < 0.001$), acute myeloid leukemia (AOR: 18.94 [15.79–22.73], $p < 0.001$), multiple myeloma (AOR: 14.21 [12.72–15.89], $p < 0.001$). Among patients with hematologic malignancies, African Americans had higher odds of COVID-19 infection than Caucasians with largest racial disparity for multiple myeloma (AOR: 4.23 [3.21–5.56], $p < 0.001$). Patients with recently diagnosed hematologic malignancies had worse outcomes (hospitalization: 51.9%, death: 14.8%) than COVID-19 patients without hematologic malignancies (hospitalization: 23.5%, death: 5.1%) ($p < 0.001$) and hematologic malignancy patients without COVID-19 (hospitalization: 15.0%, death: 4.1%) ($p < 0.001$).

1. Introduction

Coronavirus disease 2019 (COVID-19) is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and has rapidly escalated into a global pandemic [1]. Comorbidities associated with severe illness and mortality of COVID-19 include cardiovascular disease, diabetes mellitus, hypertension, chronic lung disease, chronic kidney disease and obesity [2–5]. Patients with cancer, including hematologic malignancies, often have compromised immune status and multiple comorbid conditions [6–10], many of which are also risk factors for severe COVID-19 illness. Preliminary reports showed that COVID-19 patients with cancer had higher risks for severe outcomes [11–14]. With the onset of the COVID-19 pandemic, oncologists and oncology centers have become concerned with the increased vulnerability of their patients. Currently, oncologic decision making in the time of COVID-19

epidemic is often guided by limited scientific evidence and made on the basis of opinion rather than data [15–17].

Hematologic malignancies are caused by uncontrolled growth of abnormal blood cells, which prevent blood from performing many of its functions, like fighting off infections or preventing serious bleeding [18]. The effects of these malignancies and their therapies on the immune system may put patients with hematologic malignancies at particularly high risk of COVID-19 infection and its severe outcomes. An early study from Wuhan, China showed that hospitalized COVID-19 patients with hematologic malignancies ($n = 13$) had similar infection rate as patients without hematologic malignancies [19]. Studies also showed that patients with hematologic malignancies from Wuhan, China ($n = 9$) and New York, US ($n = 54$) had higher death rates and other severe outcomes [20,21]. Clinical characteristics of COVID-19 patients with multiple myeloma ($n = 8$) and chronic myeloid

* Correspondence to: R. Xu, Center for Artificial Intelligence in Drug Discovery, Case Comprehensive Cancer Center, 2103 Cornell Rd, Cleveland, OH 44106, United States of America.

** Correspondence to: N.A. Berger, Center for Science, Health, and Society, Case Comprehensive Cancer Center, 2103 Cornell Rd, Cleveland, OH 44106, USA.

E-mail addresses: nab@case.edu (N.A. Berger), rxx@case.edu (R. Xu).

leukemia (n = 5) were recently described [22,23]. However, there is a knowledge gap regarding how COVID-19 affects patients with hematologic malignancies in specific populations stratified by age, gender and race.

2. Methods

2.1. Database description

We performed a case-control study using de-identified population-level EHR data collected by the IBM Watson Health Explorers from 360 hospitals and 317,000 providers across 50 states in the US since 1999 [24]. The EHRs are de-identified according to the Health Insurance Portability and Accountability Act (HIPAA) and the Health Information Technology for Economic and Clinical Health (HITECH) Act standards. After the de-identification process, curation process normalizes the data through mapping key elements to widely-accepted standards [25]. Specifically, disease terms are coded using the Systematized Nomenclature of Medicine-Clinical Terms (SNOMED-CT), a global standard for health terms that provides the core general terminology for EHRs [26]. The details of de-identification and normalization were described in an early study using this database [27]. More than 160 published studies have used this large-scale and standardized EHR database and the cloud-based Explorers Cohort Discovery informatics tools to study a variety of diseases including cardiovascular diseases, cancers, neurologic diseases, infectious diseases, substance use disorders among others [28], including our recent studies of COVID-19 in patients with substance use disorders [29].

2.2. Study population

At the time of this study (September 1st, 2020), the database had 73,668,830 patients, among whom 517,580 diagnosed with hematologic malignancies, 17,130 diagnosed with COVID-19, and 420 with both COVID-19 and hematologic malignancies. The status of COVID-19 was based on diagnosis code “Coronavirus infection (disorder)” (Concept Code 186747009) and the diagnosis time frame was limited within the past year to capture the timing of new cases arising during the COVID-19 pandemic. The outcome measures were COVID-19 diagnosis, rates of and hospitalization and death. The odds of COVID-19 infections were examined in two populations: (1) all patients with hematologic malignancies, who were previously diagnosed with and are living with or in remission from hematologic malignancies; (2) patients with recent hematologic malignancies patients (i.g. new cases) who were diagnosed with the cancer within the past year. Eight specific types of hematologic malignancies were examined: acute lymphoid leukemia, acute myeloid leukemia, chronic lymphoid leukemia, essential thrombocythemia, multiple myeloma, myelodysplastic syndrome, Non-Hodgkin lymphoma, and polycythemia vera. For HIPAA-compliant, statistical de-identification, the IBM Watson Health Explorers Cohort Discovery platform does not report cohort counts less than 10., therefore, hematologic malignancies types with less than 10 cases for COVID-19 were not investigated, including Hodgkin’s lymphoma, chronic myeloid leukemia (CML), myelodysplastic syndrome, malignant histiocytic disorder, mast cell malignancy, among other hematologic malignancies types.

The status of the broad term “hematologic malignancy” was based on the diagnosis of “Malignant tumor of lymphoid, hemopoietic AND/OR related tissue (disorder)” on SNOMED-CT Concept Code 26947500, acute lymphoid leukemia on the diagnosis of “Acute lymphoid leukemia, disease (disorder)” (SNOMED-CT code 91857003), acute myeloid leukemia on the diagnosis of “Acute myeloid leukemia, disease (disorder)” (SNOMED-CT code 91861009), chronic lymphoid leukemia on the diagnosis of “Chronic lymphoid leukemia, disease (disorder)” (SNOMED-CT code 92814006), essential thrombocythemia on the diagnosis of “Essential thrombocythemia (disorder)” (SNOMED-CT code 109994006), multiple myeloma on the diagnosis of “Multiple myeloma

(disorder)” (SNOMED-CT code 109989006), myelodysplastic syndrome on the diagnosis of “Myelodysplastic syndrome (disorder)” (SNOMED-CT code 109995007), Non-Hodgkin lymphoma on the diagnosis of “Non-Hodgkin’s lymphoma (disorder)” (SNOMED-CT code 118601006), and polycythemia vera (disorder) on the diagnosis of “Polycythemia vera (disorder)” (SNOMED-CT code 109992005).

The following analyses were performed: (1) odds of COVID-19 infection in patients with hematologic malignancies, adjusted for age, gender, race, and known COVID-19 risk factors, including cardiovascular diseases, type 2 diabetes, obesity, chronic kidney diseases, chronic obstructive pulmonary disease (COPD), asthma, substance use disorders, cancer therapy (chemotherapy, radiotherapy, immunotherapy), transplant procedure (bone marrow transplant, solid organ transplant) and nursing home stay status [2–5,29]. The exposure groups were patients diagnosed with a specific hematologic malignancy, the unexposed groups were patients without the specific malignancy, and the outcome measure was COVID-19 diagnosis. Separate analysis was done for all-time diagnosis of hematologic malignancy (diagnosed in the past year or prior) and recent diagnosis of hematologic malignancies (diagnosed within the past year). (2) effects of demographic factors on odds of COVID-19 infection among patients with hematologic malignancies. The case groups were patients with hematologic malignancies and one of the following demographic factors: Female, Senior (age > 65 yo), African American. The comparison groups were patients with hematologic malignancies and one of the following corresponding demographic factors (Male, Adult (age 18–65), Caucasian), adjusted for known COVID-19 risk factors. The outcome measure was COVID-19 diagnosis. (3) rates of hospitalization and death among three patient populations: patients with COVID-19 and hematologic malignancies, COVID-19 patients without hematologic malignancies, hematologic malignancy patients without COVID-19.

2.3. Statistical analysis

The EHR data are de-identified population-level (not patient-level) data, therefore we used odd ratios instead of performing regression analyses. For a given input set of patient characteristics (e.g., age, gender, race, diagnosis, finding, medications), the Explorers Cohort Discovery tool built a patient cohort by querying the underlying EHR databases for patients matching the inputs. Patients with missing values for the input queries were not included in the returned cohort. The Adjusted OR (AOR), 95% CI and p-values were calculated using the Cochran-Mantel-Haenszel (CMH) method [30] by controlling for age groups: 5-year age categories (e.g., 20–24, 25–29, ...0.65–69, 70–74, ...) were used for odds of COVID-19 infection analysis and three major age categories (juniors age < 18 years, adults age 18–65 years, senior age > 65 years) were used for age disparity analysis, gender (Female, Male), race (Caucasian, African American) and known risk factors for COVID-19 [2–5,29]. Multiple comparisons were corrected by Bonferroni correction. Two-sided, 2-sample test for equality of proportions with continuity correction were used to compare outcomes. Statistical tests were conducted with significance set at p-value < 0.05 (two-sided). All analyses were done using R, version 3.6.3.

3. Results

3.1. Patient characteristics

The baseline characteristics of the study population (as of September 1st, 2020) are presented in Table 1. Among 73,668,830 patients from the study population, 517,580 had an all-time diagnosis (diagnosis made within the last year or prior), 56,680 had a recent diagnosis (diagnosis made within the last year) of any hematologic malignancies (0.70% vs 0.08% of the study population), 17,130 patients diagnosed with COVID-19, 420 patients had all-time and 270 had recent diagnosis of hematologic malignancies (2.45% vs 1.58% of the COVID-19 population).

Table 1

Patient characteristics. Number of cases and percentage (%) are shown. Only major categories of races are shown.

Patient	Study population	Hematologic malignancies (all-time)	Hematologic malignancies (recent)	COVID-19	COVID-19 + hematologic malignancies (all-time)	COVID-19 + hematologic malignancies (recent)
Total	73,668,830	517,580	56,680	17,130	420	270
Gender						
Female	39,532,900 (54%)	272,010 (53%)	29,860 (53%)	10,050 (59%)	260 (62%)	160 (59%)
Male	33,642,290 (46%)	244,270 (47%)	26,190 (46%)	7060 (41%)	160 (38%)	110 (41%)
Unknown	507,990 (<1%)	1440 (<1%)	780 (1%)	40 (<1%)	0 (0%)	0 (0%)
Age						
Junior (< 18)	10,569,970 (14%)	10,800 (2%)	2180 (4%)	1110 (6%)	30 (7%)	20 (7%)
Adult (18 to 65)	44,101,640 (60%)	200,640 (39%)	23,580 (42%)	11,980 (70%)	210 (50%)	130 (48%)
Senior (> 65)	18,025,500 (24%)	302,110 (58%)	31,160 (55%)	4070 (24%)	180 (43%)	120 (44%)
Race						
Caucasian	40,253,700 (55%)	397,650 (77%)	42,290 (75%)	8460 (49%)	230 (55%)	150 (56%)
African American	7,599,730 (10%)	53,470 (10%)	7210 (13%)	6680 (39%)	160 (38%)	100 (37%)
Asian	1,190,850 (2%)	7480 (1%)	660 (1%)	160 (1%)	0 (0%)	0 (0%)
Hispanic/Latino	1,054,410 (1%)	4320 (<1%)	310 (<1%)	20 (<1%)	0 (0%)	0 (0%)
Unknown	9,030,700 (12%)	63,640 (12%)	4770 (8%)	930 (5%)	30 (7%)	20 (7%)

The numbers of patients with all-time and recent diagnosis of eight major hematologic malignancies as well as COVID-19 cases for all-time and recent cancer diagnosis are shown in Table 2. For example, a total of 61,580 patients were diagnosed with chronic lymphoid leukemia in the past 20 years, among whom 8789 had their diagnosis made within the past year. A total of 40 patients with all-time diagnosis of chronic lymphoid leukemia had COVID-19, among whom 20 had recent diagnosis of chronic lymphoid leukemia.

3.2. Odds of COVID-19 infection in patients with hematologic malignancies

Patients with a recent diagnosis of hematologic malignancies had higher odds of COVID-19 infection than patients without hematologic malignancies after adjusting for age, gender, race and known COVID-19 risk factors (adjusted Odds ratio or AOR: 11.91 [11.31–12.53], $p < 0.001$). Of the eight hematologic malignancies types, patients with recent diagnosis of acute lymphoid leukemia had the largest odds of COVID-19 infection, followed by essential thrombocythemia, acute myeloid leukemia, multiple myeloma, chronic lymphoid leukemia, Non-Hodgkin lymphoma, myelodysplastic syndrome and polycythemia vera (Fig. 1 top). Compared with patients with recent diagnosis of hematologic malignancies, patients with all-time diagnosis had lower odds of COVID-19 infection (AOR: 2.27 [2.17–2.36], $p < 0.001$), with the strongest effect for acute lymphoid leukemia followed by essential

thrombocythemia, multiple myeloma, acute myeloid leukemia, Non-Hodgkin lymphoma, chronic lymphoid leukemia, polycythemia vera and myelodysplastic syndrome (Fig. 1 bottom). Since 10.9% of patients with all-time diagnosis of hematologic malignancies had their malignancies diagnosed within the past year, their modest increased risk for COVID-19 infection is likely contributed by the sub-population with recent diagnosis.

3.3. Effects of demographics on odds of COVID-19 infection in patients with hematologic malignancies

We examined how demographic factors (age, sex, race) affected COVID-19 risk among patients with recently diagnosed hematologic malignancies. Among eight hematologic malignancy types, three with the most numbers of COVID-19 cases were examined: essential thrombocythemia, multiple myeloma and non-Hodgkin lymphoma. Among patients with recent diagnosis of these cancers, African Americans had more than 2-fold increase of odds for COVID-19 infection than Caucasians, adjusted for age, gender and COVID-19 risk factors, with largest race disparity for multiple myeloma (AOR: 4.23 [3.21–5.56]), $p < 0.001$, followed by essential thrombocythemia (AOR: 3.74 [3.12–4.48], $p < 0.001$) and non-Hodgkin lymphoma (AOR: 2.43 [1.91–3.09], $p < 0.001$). Compared with men, women with essential thrombocythemia had lower odds of COVID-19 infection and women with non-Hodgkin lymphoma had higher odds for COVID-19 infection. Advanced age had

Table 2

The numbers of patients with all-time and recent diagnosis of eight major hematologic malignancies as well as COVID-19 cases for all-time and recent cancer diagnosis.

Hematologic malignancy	All		COVID-19	
	All-time cancer diagnosis	Recent cancer diagnosis	All-time cancer diagnosis	Recent cancer diagnosis
Acute lymphoid leukemia	18,360	2040	30	20
Acute myeloid leukemia	27,850	2430	20	20
Chronic lymphoid leukemia	61,580	8780	40	30
Essential thrombocythemia	121,200	13,780	210	120
Multiple myeloma	61,350	7790	70	50
Myelodysplastic syndrome	50,370	4580	20	10
Non-Hodgkin lymphoma	169,390	26,190	130	90
Polycythemia vera	72,150	6880	40	20

Odds of COVID-19 infection in patients with recent diagnosis of hematologic malignancies

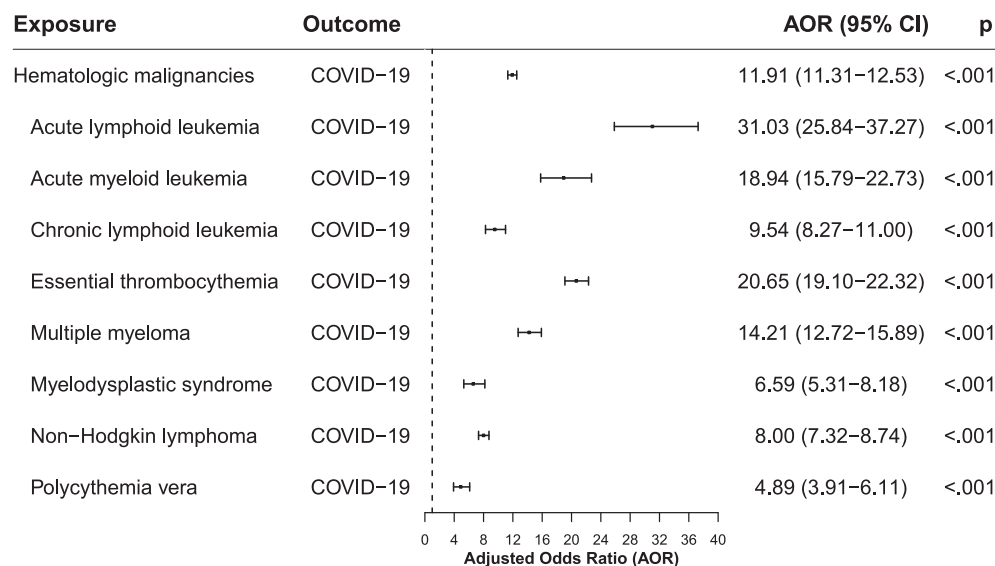
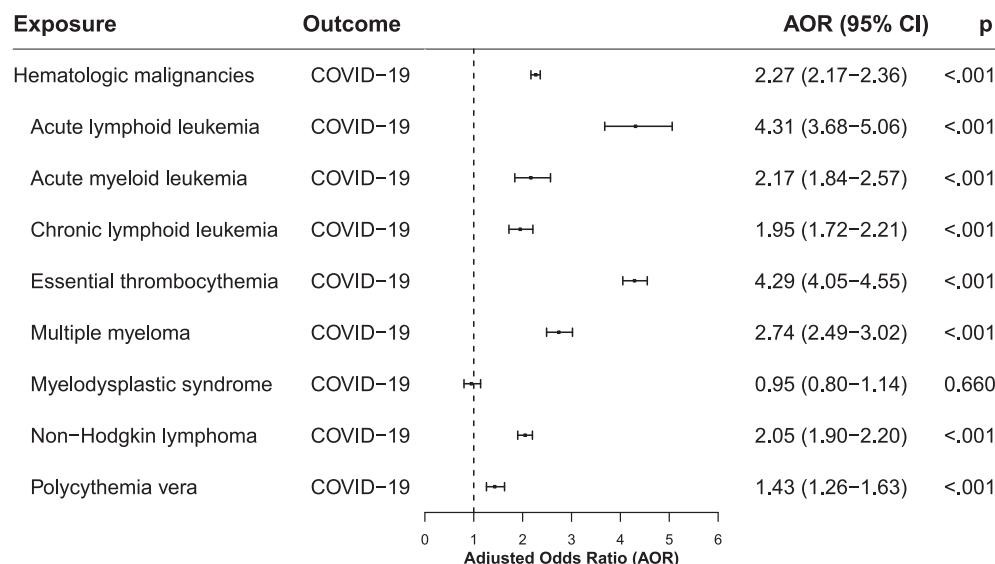


Fig. 1. Odds of COVID-19 infection in patients with recent versus all-time diagnosis of hematologic malignancies, adjusted for age, gender, race and potential COVID-19 risk factors (cardiovascular diseases, type 2 diabetes, obesity, chronic kidney diseases, chronic obstructive pulmonary disease (COPD), asthma, substance use disorders, cancer therapy (chemotherapy, radiotherapy, immunotherapy), transplant procedure (bone marrow transplant, solid organ transplant) and nursing home stay status.

Odds of COVID-19 infection in patients with all-time diagnosis of hematologic malignancies



no statistically significant effects on patients' risk for COVID-19 infection (Fig. 2). Similar strong racial disparities were observed for patients with all-time diagnosis of essential thrombocythemia, multiple myeloma and non-Hodgkin lymphoma (Supplementary Fig. S1).

3.4. Hospitalizations and death rates among patients with hematologic malignancies and COVID-19

Among 17,130 COVID-19 patients, 4110 were hospitalized (24.0%), higher for African Americans (32.8%) than Caucasians (20.3%)($p < 0.001$). Among 270 patients with COVID-19 and recent diagnosis of hematologic malignancies, 140 were hospitalized (51.9%), not

statistically different between African Americans (60.0%) and Caucasians (53.3%). Among 16,860 COVID-19 patients without recently diagnosed hematologic malignancies, 3960 were hospitalized (23.5%), higher for African Americans (32.4%) than Caucasians (19.7%) ($p < 0.01$). Among 56,410 patients with recently diagnosed hematologic malignancies but without COVID-19, 8460 were hospitalized (15.0%), higher for African Americans (19.5%) than Caucasians (15.7%) ($p < 0.001$) (Fig. 3, top). Overall the hospitalization rate for patients with recently diagnosed hematologic malignancies and COVID-19 (51.9%) was higher than for COVID-19 patients without hematologic malignancies (23.5%) ($p < 0.001$) and that for hematologic malignancy patients without COVID-19 (15.0%) ($p < 0.001$), with synergistic effect

Demographic disparities of COVID-19 infection in patients with recent diagnosis of hematologic malignancy

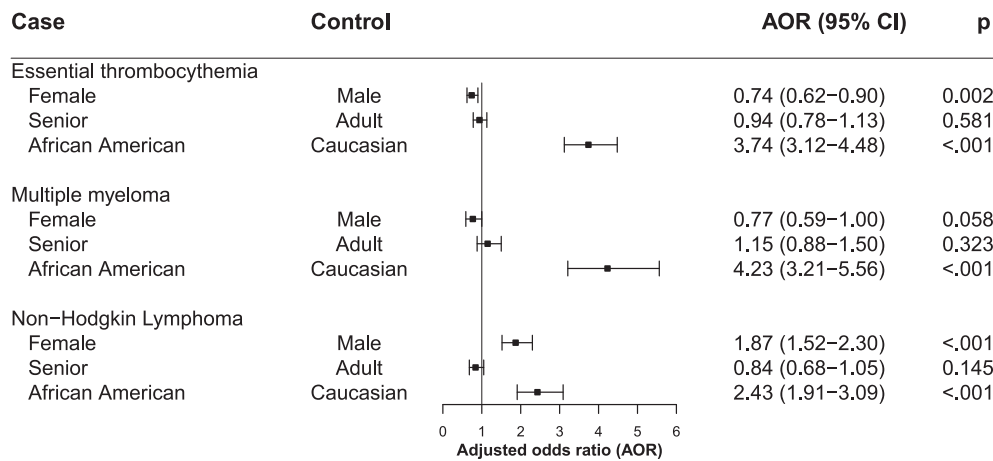


Fig. 2. Effects of demographics on odds of COVID-19 infection among patients with recently diagnosed hematologic malignancies, adjusted for other demographics and potential COVID-19 risk factors (cardiovascular diseases, type 2 diabetes, obesity, chronic kidney diseases, chronic obstructive pulmonary disease (COPD), asthma, substance use disorders, cancer therapy (chemotherapy, radiotherapy, immunotherapy), transplant procedure (bone marrow transplant, solid organ transplant) and nursing home stay status.

observed. The hospitalization rates for patients with all-time diagnosis of hematologic malignancies are similar to these for recent diagnosis (Supplementary Fig. S2).

Among 17,130 COVID-19 patients, 900 died (5.3%), higher for African Americans (6.9%) than Caucasians (4.6%)($p < 0.001$). Among 270 patients with COVID-19 and recent diagnosis of hematologic malignancies, 40 died (14.8%), not statistically different from African Americans (20.0%) and Caucasians (13.3%). Among 16,860 COVID-19 patients without recently diagnosed hematologic malignancies, 860 died (5.1%), higher for African Americans (6.8%) than Caucasians (4.5%) ($p < 0.01$). Among 56,410 patients with recently diagnosed hematologic malignancies but without COVID-19, 2340 died (4.1%), higher for African Americans (5.8%) than Caucasians (4.3%) ($p < 0.001$) (Fig. 3, bottom). Overall the death rate for patients with recently diagnosed hematologic malignancies and COVID-19 (14.8%) was higher than for COVID-19 patients without hematologic malignancies (5.1%) ($p < 0.001$) and that for hematologic malignancy patients without COVID-19 (4.1%) ($p < 0.001$), with synergistic effect observed. The death rates for patients with all-time diagnosis of hematologic malignancies are similar to these for recent diagnosis (Supplementary Fig. S3).

4. Summary and future directions

We screened and analyzed a nation-wide EHR database in the US and showed that patients with hematologic malignancies diagnosed within the last year, particularly acute lymphoid leukemia, essential thrombocythemia, multiple myeloma and acute myeloid leukemia, had increased odds for COVID-19 infection compared with patients without hematologic malignancies. Among patients with hematologic malignancies, African Americans had more than 2-fold increase in odds of COVID-19 infection than Caucasians. This study identified high risk hematologic malignancies patient groups based on diagnostic and demographic groups who are most vulnerable to COVID-19 illness and adverse outcomes. Since our EHR-based study represents nationwide population basis, it provides an independent extension of two recent reports by the COVID-19 and Cancer Consortium (CCC19) registry, the majority of which are based in academic medical centers, mostly in the Northeast US [13,16].

Our study shows that patients with hematologic malignancies are at increased risk for developing COVID-19 as compared to patients without hematologic malignancies, after adjusting for age, gender, race and known COVID-19 risk factors. These findings are consistent for the eight common hematologic malignancies for which there was sufficient data

for us to examine. Patients with hematologic malignancies often had more comorbidities and cancer treatments, and higher contacts with medical systems than those without cancers. After adjusting for age, gender, race and known COVID-19 risk factors that include common comorbidities such as cardiovascular diseases, type 2 diabetes and obesity, cancer treatments and transplant procedures, patients with hematologic malignancies still had high odds of COVID-19 infection, with marked difference across different cancer types. These results suggest that these cancers might have direct effects on patients' risk for COVID-19 infection. Hematologic malignancies affect the production and function of blood cells in fighting off infections [31]. Patients with hematologic malignancies often have multiple immune dysfunctions of the innate and adaptive immune system including low immunoglobulin G levels in patients with chronic lymphocytic leukemia (CLL) or other B cell neoplasms, and functionally impaired immature or neoplastic dysfunctional granulocytes in patients with myeloid neoplasms [32,33]. Impaired immune functions predispose patients with hematologic neoplasms to diverse array of infections. Our study indicates that impaired immune functions predispose patients with hematologic neoplasms to COVID-19 infection.

Other unadjusted confounding factors (e.g., other comorbidities, symptoms, contacts with medical systems, socio-economic factors, life styles) might also play important roles. However, unadjusted potential confounders alone may not be sufficient to explain the widely differential odds of COVID-19 infection for different types of hematologic malignancies, as patients with acute lymphoid leukemia, who had the largest odds of COVID-19 infection, did not necessarily have more comorbidities or lower socio-economic status than those with other types of hematologic malignancies. Increased contacts with healthcare workers may affect a patient's risk for COVID-19 infection and patients with more aggressive cancers often need more healthcare contacts and testing. However we did not observe such trend as patients with more aggressive cancer (e.g., acute myeloid leukemia) had higher odds of COVID-19 infection than those with less aggressive cancer (e.g., acute lymphoid leukemia, essential thrombocythemia). Our study may be confounded by the likelihood that patients with hematologic malignancies might have been more likely to be tested in the setting of elective admission. This needs further exploration in future studies in other datasets since our current EHR database does not discern between planned admissions and those admissions related to the COVID-19 diagnosis. Taken together, our data suggest that the effects of hematologic malignancies on the immune system may put patients with these malignancies at particularly high risk of COVID-19 infection and this

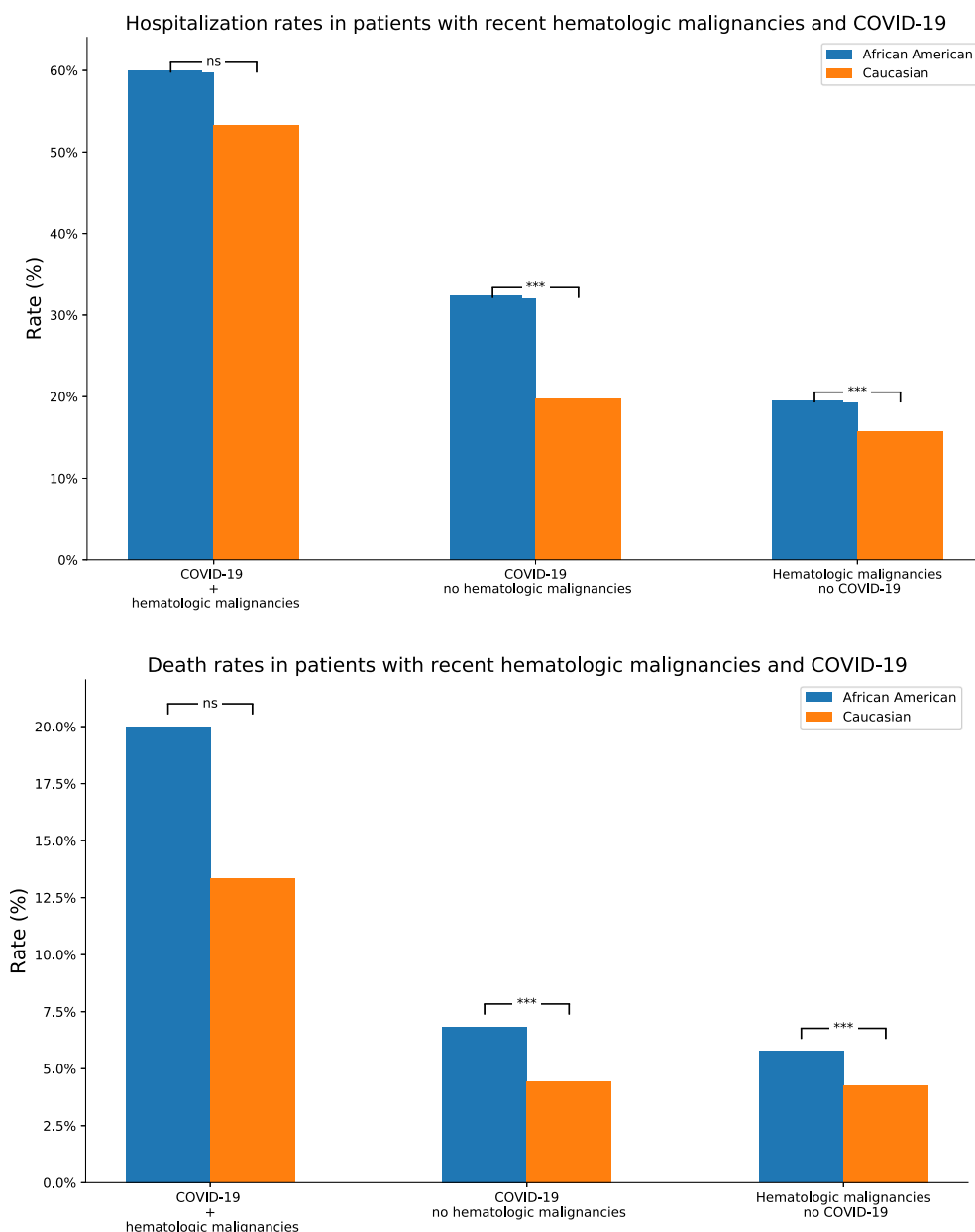


Fig. 3. Hospitalization and death rates among three populations: patients with both recent diagnosis of hematologic malignancies and COVID-19, COVID-19 patients without hematologic malignancies, hematologic malignancy patients without COVID-19. ***: $p < 0.0001$; **: $p < 0.001$; ns: not statistically significant. The SNOMED-CT concepts “Hospital admission (procedure)” (ID 32485007) was used to obtain hospitalization status from patient EHRs. Exploratory regularly imports from the Social Security Death index for the “deceased” status.

finding needs to be replicated in other independent datasets.

Patients with recent diagnosis of acute lymphoid leukemia had the largest odds of COVID-19 infection, followed by essential thrombocythemia, acute myeloid leukemia, multiple myeloma, chronic lymphoid leukemia, Non-Hodgkin lymphoma, myelodysplastic syndrome and polycythemia vera. Patients with acute hematologic malignancies had higher risk for COVID-19 infection than those with chronic hematologic malignancies patients, which may be associated with a more aggressive disease phenotype. The high risk for COVID-19 in patients with essential thrombocythemia suggests the possibility of direct interaction of the virus with a specific component of platelet structure and/or function and warrants further investigation. The relative absence of patients with other major hematologic malignancies types including chronic myeloid leukemia and Hodgkin lymphoma from the COVID-19 group compared to their prevalence in the study population suggests the possibility that either their disease status and/or their treatment may confer a degree of resistance.

In our study, African Americans with hematologic malignancies had higher risk for COVID-19 infection than Caucasians after adjusting for

age, gender and known COVID-19 risk factors. This is consistent with data for general population across the US showing that COVID-19 affects African Americans at a disproportionately high rate [34]. The more than 2-fold difference in odd ratio after adjusting for comorbidities and other COVID-19 risk factors suggests that other factors such as social, behavioral and lifestyle factors may have contributed to the profound racial inequality. The EHR data has limited social-economical information, which limited our ability of investigating how social adversity and other social-economical determinants may contribute to the observed profound racial inequality in patients with hematologic malignancies. While advanced age is a risk factor for COVID-19 severe illness in general population, our study showed that advanced age conferred no additional effect on the risk of getting COVID-19 among patients with hematologic malignancies. Compared with men, women with essential thrombocythemia had lower odds of COVID-19 infection and women with non-Hodgkin lymphoma had higher odds for COVID-19 infection. The reason(s) for this differential gender effects on patients with different types of hematologic malignancies warrant further investigation.

Previous studies showed that patients with hematologic malignancies and COVID-19 had more severe outcomes than non-cancer patients [20,21]. Since patients with hematologic malignancies in general had more severe outcomes than non-cancer patients, it was unclear if COVID-19 exerted additive or synergistic adverse outcome effects. In our study, we observed synergistic effects of COVID-19 and hematologic malignancies on both hospitalization and death outcomes.

It is possible that higher rates of hospitalization may be partially due to higher rate of planned COVID-19 testing in the setting of elective admissions. However, the limited data available in the EHR database does not allow us to discern the direct causes for hospital admission or death. African Americans had worse outcomes than Caucasians both in COVID-19 patients without hematologic malignancies and in hematologic malignancy patients without COVID-19. No statistically significant difference in rates of death and hospitalization was observed in African Americans and Caucasians with COVID-19 and hematologic malignancies. This may be due to the relatively small sample size of patients with both COVID-19 and hematologic malignancies ($n = 420$) at the time of this study.

Our study has several limitations. First, this is an analysis of patient EHR data. Patient EHR data were collected for billing purposes. Though EHR data have been widely used for research purposes, they have inherent limitations including under, over, or misdiagnosis, limited information of time-series, timing and adherence of medications and treatments, socio-economic and life style determinants, among others [35–37]. Second, in the EHR database, a total of 17,130 patients were diagnosed with COVID-19 at the time of our study, which is significantly lower than the total confirmed cases in US reported by CDC [1]. COVID-19 is regularly tested at drive-ups and popup testing locations other than hospitals, therefore many of the confirmed cases may have not been captured by EHRs. COVID-19 cases in our study were based on diagnosis codes in the EHR database for billing purpose, therefore they likely skewed towards more severe cases of COVID-19. Third, specific for the Explorys database, no detailed information was provided regarding rural vs urban, socioeconomic and geographic composition of the patient population as well as healthcare facility venue. In addition, patient ages were treated as a categorical instead of continuous variable. For HIPAA-compliant statistical de-identification purposes, the Explorys Explore tool does not allow reporting on sample sizes less than 10 and all population counts above 10 are rounded to the nearest 10. This rounding may affect the accuracy of the results, particularly when some categories of patients are very small. Fourth, though patients in this EHR database represent 20% of US population, they represent patients who had encounters with healthcare systems and are not necessarily representative of the US population. For example, among a total of 27,850 cases of all-time AML diagnoses, 2430 (8.7%) were diagnosed in the past year, which is lower than estimated 30% new cases of AML in 2020 (estimated 19,940 new diagnoses among 64,512 people living with AML in the United States) reported by the latest Surveillance, Epidemiology, and End Results Program (SEER), National Program of Cancer Registries (NPCR), and patient EHR databases and showed that there are general similarities in demographic and geographic distribution of cancer patients in these databases but there are also overarching differences across the populations they cover [39]. In addition, due to relatively lower adoption rate of health IT technologies in rural versus urban areas, patients from rural areas may be less represented in the this EHR database. These limitations should be considered when interpreting and extrapolating our results obtained with the EHR database. This study can serve as a baseline study of COVID-19 risk, disparities and outcomes in patients with hematologic malignancies and findings from this study need to be validated in other patient databases or other populations.

4.1. Future considerations

- Future studies will be required to validate these findings in other independent patient databases or other populations as well as those for which insufficient data was available such as Hodgkin lymphoma, chronic myeloid leukemia.
- Studies need to be conducted to determine which aspects of newly diagnosed hematologic malignancies, e.g., therapeutic agents, cytopenias, immune insufficiency, may contribute to increased susceptibility to COVID-19 infection and how these effects can be mitigated.
- Given the high risk for COVID-19 infection and mortality in patients with recently diagnosed hematologic malignancies, these patients are prime candidates for evaluation and administration for induction of prophylactic passive and/or active immunity.

4.2. Practice points

- All patients with hematologic malignancies should practice recommended guidelines for avoiding COVID-19 exposure.
- Patients with newly diagnosed hematologic malignancies (within last year) should be extra cautious regarding COVID-19 exposure.
- African American patients with hematologic malignancies should be extra cautious regarding COVID-19 exposure.
- hematologic malignancy patients should have therapeutic interventions directed at reducing COVID-19 risk factors.

4.3. Research agenda

- To interrogate patient electronic health records acquired on a nationwide basis to identify a broad population-based inventory of COVID-19 infected patients with hematologic malignancies.
- To evaluate infections and impact of COVID-19 in patients with specific hematologic malignancies.
- To identify impact of race and demographics on patients with COVID-19 infection and hematologic malignancies.
- To identify association of comorbidities with COVID-19 infection in patients with hematologic malignancies.
- To provide early nationwide data-driven indicators to inform development of precautionary strategies for management of patients with hematologic malignancies in era of COVID-19.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.blre.2020.100775>.

Authors' contributions

Q.W. and RX conceived the study, designed the experiments, conducted the analysis and authored the manuscript. N.A.B participated in study design, clinical interpretation, result analysis, and manuscript preparation.

Role of the funding source

The funding source had no role in the design of this study and will not have any role during its execution, analyses, interpretation of the data, or decision to submit results.

Funding

This work was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under the NIH Director's New Innovator Award (grant number DP2HD084068), NIH National Institute on Aging (grant numbers R01 AG057557, R01 AG061388, R56 AG062272), American Cancer Society Research Scholar Grant (grant number RSG-16-049-01 –

MPC), The Clinical and Translational Science Collaborative (CTSC) of Cleveland (grant number 1UL1TR002548-01), NIH National Cancer Institute (grant numbers CA043703, CA2332216).

Declaration of Competing Interest

Q.W., N.A.B, and R.X have no financial interests to disclose.

References

- [1] CDC. Cases in the U.S. <https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html> (accessed on September 1st, 2020).
- [2] CDC. Groups at Higher Risk for COVID-19 Severe Illness. <https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/groups-at-higher-risk.html#serious-heart-conditions> (accessed on June 28, 2020).
- [3] Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area [published correction appears in doi: 10.1001/jama.2020.7681]. *JAMA* 2020;323(20):2052–9. <https://doi.org/10.1001/jama.2020.6775>.
- [4] Myers LC, Parodi SM, Escobar GJ, Liu VX. Characteristics of Hospitalized Adults With COVID-19 in an Integrated Health Care System in California [published online ahead of print, 2020 Apr 24]. *JAMA*. 2020:e207202. <https://doi.org/10.1001/jama.2020.7202>.
- [5] Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 2020 Mar 28;395(10229):1054–62. [https://doi.org/10.1016/S0140-6736\(20\)30566-3](https://doi.org/10.1016/S0140-6736(20)30566-3) [Erratum in *Lancet*. 2020 Mar 28;395(10229):1038. doi: 10.1016/S0140-6736(20)30606-1. Epub 2020 Mar 12. PMID: 32171424] [Erratum in *Lancet*. 2020 Mar 28;395(10229):1038. doi: 10.1016/S0140-6736(20)30638-3. Epub 2020 Mar 17. PMID: 32192581].
- [6] Schreiber RD, Old LJ, Smyth MJ. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. *Science* 2011;331(6024):1565–70. <https://doi.org/10.1126/science.1203492>.
- [7] Ogle KS, Swanson GM, Woods N, Azzouz F. Cancer and comorbidity: redefining chronic diseases. *Cancer* 2000;88(3):653–63.
- [8] Lauby-Secretan B, Scoccianti C, Loomis D, Grosse Y, Bianchini F, Straif K. Body fatness and cancer—viewpoint of the IARC working group. *N Engl J Med* 2016;375(8):794–8. <https://doi.org/10.1056/NEJMs1606602>.
- [9] Giovannucci E, Harlan DM, Archer MC, Bergenstal M, Gapstur SM, Habel LA, et al. Diabetes and cancer: a consensus report. *Diabetes Care* 2010;33(7):1674–85. <https://doi.org/10.2337/dc10-0666>.
- [10] Kitis RN, Riquelme JA, Lavandro S. Heart disease and cancer: are the two killers colluding? *Circulation* 2018;138(7):692–5. <https://doi.org/10.1161/CIRCULATIONAHA.118.033907>.
- [11] Yu J, Ouyang W, Chua MLK, Xie C. SARS-CoV-2 Transmission in Patients With Cancer at a Tertiary Care Hospital in Wuhan, China [published online ahead of print, 2020 Mar 25]. *JAMA Oncol* 2020:e200980. <https://doi.org/10.1001/jamaoncol.2020.0980>.
- [12] Yang K, Sheng Y, Huang C, Jin Y, Xiong N, Jiang K, et al. Clinical characteristics, outcomes, and risk factors for mortality in patients with cancer and COVID-19 in Hubei, China: a multicentre, retrospective, cohort study. *Lancet Oncol* 2020;21(7):904–13. [https://doi.org/10.1016/S1470-2045\(20\)30310-7](https://doi.org/10.1016/S1470-2045(20)30310-7).
- [13] Kuderer NM, Choueiri TK, Shah DP, Shyr Y, Rubinstein SM, Rivera DR, et al. Clinical impact of COVID-19 on patients with cancer (CCC19): a cohort study. *Lancet*. 2020;395(10241):1907–18. [https://doi.org/10.1016/S0140-6736\(20\)31187-9](https://doi.org/10.1016/S0140-6736(20)31187-9).
- [14] Robilotti EV, Babady NE, Mead PA, Rolling T, Perez-Johnston R, Bernardes M, et al. Determinants of COVID-19 disease severity in patients with cancer [published online ahead of print, 2020 Jun 24]. *Nat Med* 2020. <https://doi.org/10.1038/s41591-020-0979-0>.
- [15] Hanna TP, Evans GA, Booth CM. Cancer, COVID-19 and the precautionary principle: prioritizing treatment during a global pandemic. *Nat Rev Clin Oncol* 2020;17(5):268–70. <https://doi.org/10.1038/s41571-020-0362-6>.
- [16] Rivera DR, Peters S, Panagiotou OA, Shah DP, Kuderer NM, Hsu CY, et al. Utilization of COVID-19 treatments and clinical outcomes among patients with cancer: A COVID-19 and Cancer Consortium (CCC19) cohort study [published online ahead of print, 2020 Jul 22]. *Cancer Discov* 2020. <https://doi.org/10.1158/2159-8290.CD-20-0941>. CD-20-0941.
- [17] van de Haar J, Hoes LR, Coles CE, Seamon K, Fröhling S, Jäger D, et al. Caring for patients with cancer in the COVID-19 era. *Nat Med* 2020;26(5):665–71. <https://doi.org/10.1038/s41591-020-0874-8>.
- [18] <https://www.hematology.org/education/patients/blood-cancers>.
- [19] He W, Chen L, Chen L, Yuan G, Fang Y, Chen W, et al. COVID-19 in persons with haematological cancers. *Leukemia* 2020;34(6):1637–45. <https://doi.org/10.1038/s41375-020-0836-7>.
- [20] Dai M, Liu D, Liu M, Zhou F, Li G, Chen Z, et al. Patients with cancer appear more vulnerable to SARS-CoV-2: a multi-center study during the COVID-19 outbreak. *Cancer Discov* 2020;10(6):783–91. <https://doi.org/10.1158/2159-8290.CD-20-0422>.
- [21] Mehta V, Goel S, Kabarriti R, Cole D, Goldfinger M, Acuna-Villaorduna A, et al. Case fatality rate of cancer patients with COVID-19 in a New York hospital system. *Cancer Discov* 2020;10(7):935–41. <https://doi.org/10.1158/2159-8290.CD-20-0516>.
- [22] Dhakal B, D'Souza A, Chhabra S, Hari P. Multiple myeloma and COVID-19. *Leukemia* 2020;34(7):1961–3. <https://doi.org/10.1038/s41375-020-0879-9>.
- [23] Li W, Wang D, Guo J, Yuan G, Yang Z, Gale RP, et al. COVID-19 in persons with chronic myeloid leukaemia. *Leukemia* 2020;34(7):1799–804. <https://doi.org/10.1038/s41375-020-0853-6>.
- [24] IBM Explorix EHR Solutions. <https://www.ibm.com/watson-health/about/explorix>; 2020 (accessed up to September 1st, 2020).
- [25] Bodenreider O. The unified medical language system (UMLS): integrating biomedical terminology. *Nucleic Acids Res* 2004;32(Database issue):D267–70. <https://doi.org/10.1093/nar/gkh061>.
- [26] SNOMED International. The Systematized Nomenclature of Medicine – Clinical Terms (SNOMED CT). <http://www.snomed.org/snomed-ct/why-snomed-ct>.
- [27] Kaelber DC, Foster W, Gilder J, Love TE, Jain AK. Patient characteristics associated with venous thromboembolic events: a cohort study using pooled electronic health record data. *J Am Med Inform Assoc* 2012;19(6):965–72. <https://doi.org/10.1136/amiajnl-2011-000782>.
- [28] IBM Explorix EHR Database Bibliography Categorized by Therapeutic Area. <https://www.ibm.com/downloads/cas/RPKNLLIM> (accessed on July 14, 2020).
- [29] Wang QQ, Kaelber DC, Xu R, Volkow ND. COVID-19 risk and outcomes in patients with substance use disorders: analyses from electronic health records in the United States. *Mol Psychiatry* 2020 Sep;14:1–10. <https://doi.org/10.1038/s41380-020-00880-7>.
- [30] Kuritz SJ, Landis JR, Koch GG. A general overview of mantel-Haenszel methods: applications and recent developments. *Annu Rev Public Health* 1988;9:123–60. <https://doi.org/10.1146/annurev.pu.09.050188.001011>.
- [31] American Society of Hematology. Blood Cancers. <https://www.hematology.org/education/patients/blood-cancers> (accessed on October 27, 2020).
- [32] Safdar A, Armstrong D. Infections in patients with hematologic neoplasms and hematopoietic stem cell transplantation: neutropenia, humoral, and splenic defects. *Clin Infect Dis* 2011;53(8):798–806. <https://doi.org/10.1093/cid/cir492>.
- [33] Hamblin AD, Hamblin TJ. The immunodeficiency of chronic lymphocytic leukaemia. *Br Med Bull* 2008;87:49–62. <https://doi.org/10.1093/bmb/ldn034>. Epub 2008 Aug 27. PMID: 18755702.
- [34] Yancy CW. COVID-19 and African Americans [published online ahead of print, 2020 Apr 15]. *JAMA* 2020. <https://doi.org/10.1001/jama.2020.6548>.
- [35] Cowie MR, Blomster JI, Curtis LH, Duclaux S, Ford I, Fritz F, et al. Electronic health records to facilitate clinical research. *Clin Res Cardiol* 2017;106(1):1–9. <https://doi.org/10.1007/s00392-016-1025-6>.
- [36] Coorevits P, Sundgren M, Klein GO, Bahr A, Claerhout B, Daniel C, et al. Electronic health records: new opportunities for clinical research. *J Intern Med* 2013;274(6):547–60. <https://doi.org/10.1111/joim.12119>.
- [37] Ahmad FS, Chan C, Rosenman MB, Post WS, Fort DG, Greenland P, et al. Validity of cardiovascular data from electronic sources: the multi-ethnic study of atherosclerosis and HealthLNK. *Circulation* 2017;136(13):1207–16. <https://doi.org/10.1161/CIRCULATIONAHA.117.027436>.
- [38] National Cancer Institute (NCI). The Surveillance, Epidemiology, and End Results (SEER) program. In: *Cancer Stat Facts: Leukemia — Acute Myeloid Leukemia (AML)*; 2020. <https://seer.cancer.gov/statfacts/html/amyl.html> (accessed on October 27, 2020).
- [39] Ma X, Long L, Moon S, Adamson BJ, Baxi SS. Comparison of population characteristics in real-world clinical oncology databases in the US. *Flatiron Health, SEER, and NPCR*; 2020. <https://doi.org/10.1101/2020.03.16.20037143> (medRxiv).