

Analyses of Risk, Racial Disparity, and Outcomes Among US Patients With Cancer and COVID-19 Infection

QuanQiu Wang, MS; Nathan A. Berger, MD; Rong Xu, PhD

IMPORTANCE Patients with specific cancers may be at higher risk than those without cancer for coronavirus disease 2019 (COVID-19) and its severe outcomes. At present, limited data are available on the risk, racial disparity, and outcomes for COVID-19 illness in patients with cancer.

OBJECTIVES To investigate how patients with specific types of cancer are at risk for COVID-19 infection and its adverse outcomes and whether there are cancer-specific race disparities for COVID-19 infection.

DESIGN, SETTING, AND PARTICIPANTS This retrospective case-control analysis of patient electronic health records included 73.4 million patients from 360 hospitals and 317 000 clinicians across 50 US states to August 14, 2020. The odds of COVID-19 infections for 13 common cancer types and adverse outcomes were assessed.

EXPOSURES The exposure groups were patients diagnosed with a specific cancer, whereas the unexposed groups were patients without the specific cancer.

MAIN OUTCOMES AND MEASURES The adjusted odds ratio (aOR) and 95% CI were estimated using the Cochran-Mantel-Haenszel test for the risk of COVID-19 infection.

RESULTS Among the 73.4 million patients included in the analysis (53.6% female), 2 523 920 had at least 1 of the 13 common cancers diagnosed (all cancer diagnosed within or before the last year), and 273 140 had recent cancer (cancer diagnosed within the last year). Among 16 570 patients diagnosed with COVID-19, 1200 had a cancer diagnosis and 690 had a recent cancer diagnosis of at least 1 of the 13 common cancers. Those with recent cancer diagnosis were at significantly increased risk for COVID-19 infection (aOR, 7.14 [95% CI, 6.91-7.39]; $P < .001$), with the strongest association for recently diagnosed leukemia (aOR, 12.16 [95% CI, 11.03-13.40]; $P < .001$), non-Hodgkin lymphoma (aOR, 8.54 [95% CI, 7.80-9.36]; $P < .001$), and lung cancer (aOR, 7.66 [95% CI, 7.07-8.29]; $P < .001$) and weakest for thyroid cancer (aOR, 3.10 [95% CI, 2.47-3.87]; $P < .001$). Among patients with recent cancer diagnosis, African Americans had a significantly higher risk for COVID-19 infection than White patients; this racial disparity was largest for breast cancer (aOR, 5.44 [95% CI, 4.69-6.31]; $P < .001$), followed by prostate cancer (aOR, 5.10 [95% CI, 4.34-5.98]; $P < .001$), colorectal cancer (aOR, 3.30 [95% CI, 2.55-4.26]; $P < .001$), and lung cancer (aOR, 2.53 [95% CI, 2.10-3.06]; $P < .001$). Patients with cancer and COVID-19 had significantly worse outcomes (hospitalization, 47.46%; death, 14.93%) than patients with COVID-19 without cancer (hospitalization, 24.26%; death, 5.26%) ($P < .001$) and patients with cancer without COVID-19 (hospitalization, 12.39%; death, 4.03%) ($P < .001$).

CONCLUSIONS AND RELEVANCE In this case-control study, patients with cancer were at significantly increased risk for COVID-19 infection and worse outcomes, which was further exacerbated among African Americans. These findings highlight the need to protect and monitor patients with cancer as part of the strategy to control the pandemic.

JAMA Oncol. 2021;7(2):220-227. doi:10.1001/jamaoncol.2020.6178
Published online December 10, 2020.

← Invited Commentary page 227

+ Supplemental content

+ CME Quiz at
jamacmelookup.com and CME
Questions page 315

Author Affiliations: Center for Artificial Intelligence in Drug Discovery, School of Medicine, Case Western Reserve University, Cleveland, Ohio (Wang, Xu); Center for Science, Health, and Society, School of Medicine, Case Western Reserve University, Cleveland, Ohio (Berger); Case Comprehensive Cancer Center, School of Medicine, Case Western Reserve University, Cleveland, Ohio (Berger, Xu).

Corresponding Authors: Nathan A. Berger, MD, Center for Science, Health, and Society, School of Medicine, Case Western Reserve University, 2103 Cornell Rd, Cleveland, OH 44106 (nab@case.edu); Rong Xu, PhD, Center for Artificial Intelligence in Drug Discovery, School of Medicine, Case Western Reserve University, 2103 Cornell Rd, Cleveland, OH 44106 (rxu@case.edu).

Coronavirus disease 2019 (COVID-19) is caused by severe acute respiratory syndrome coronavirus 2 and has rapidly escalated into a global pandemic.¹ Severe illness of COVID-19 predominantly occurs in individuals with underlying medical comorbidities.²⁻⁵ Patients with cancer have weakened immune status; often have multiple comorbid conditions, including type 2 diabetes, obesity, and cardiovascular diseases; and are more likely to get infections.⁶⁻¹¹ Reports from China and the US described clinical characteristics and outcomes of patients with cancer and COVID-19 and examined risk factors for mortality in this population; they suggested a higher rate of death due to COVID-19 and severe outcomes in patients with cancer.¹²⁻¹⁶ However, there is a knowledge gap regarding the susceptibility of patients with cancer to COVID-19 and its severe outcomes in the US. It remains unknown how race and other demographic factors such as age and sex affect the risk of COVID-19 infection and outcomes among patients with cancer, given that the cancer burden in the US varies by race and other demographic factors.

Methods

Database Description

We performed a retrospective case-control study using deidentified electronic health record (EHR) data collected by the IBM Watson Health Explorys from 360 hospitals and 317 000 clinicians across 50 states in the US since 1999, representing 20% of the US population.¹⁷ The EHRs are deidentified according to the Health Insurance Portability and Accountability Act and the Health Information Technology for Economic and Clinical Health Act standards.¹⁸ After the deidentification process, a curation process normalizes the data through mapping key elements to widely accepted standards.^{18,19} 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.²⁰ More than 100 published studies showed that with this large-scale, standardized EHR database and the cloud-based Explorys Cohort Discovery informatics tools, large case-control studies can be undertaken efficiently,²¹ including 2 recent studies by Zhou et al.^{22,23} The EHR data are deidentified, and the institutional review board of Case Western Reserve University determined that this study was exempt from approval and informed consent. This study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline.

Study Population

At the time of this study (August 14, 2020), the study population consisted of 73 449 510 patients, including 2 523 920 patients diagnosed with at least 1 of 13 common cancers. A total of 273 140 cancer diagnoses occurred within the past year; 16 570 patients were diagnosed with COVID-19, and 1200 with both COVID-19 and cancer. The status of COVID-19 was based on the concept *coronavirus infection (disorder)* (Concept Code 186747009), and we further limited the diagnosis time frame to within the past year to capture the timing of new cases arising

Key Points

Question Are patients with cancer at increased risk for coronavirus disease 2019 (COVID-19) infection and its adverse outcomes?

Findings In this case-control analysis of electronic medical records from 73.4 million unique patients, patients with a recent diagnosis of cancer were at significantly increased risk for COVID-19 infection and its adverse outcomes, especially in African Americans.

Meaning Based on these findings, it is important to closely monitor patients with cancer and protect them from exposure to severe acute respiratory syndrome coronavirus 2 and the severe outcomes of COVID-19.

during the COVID-19 pandemic. The outcome measures were COVID-19 diagnosis, rates of death, and hospitalization. We examined the risk association between COVID-19 and 13 common cancers.²⁴

The status of the 13 cancer types was based on the diagnosis codes from SNOMED-CT.²⁰ The status of bladder cancer was based on the diagnosis of *malignant tumor of urinary bladder (disorder)* (SNOMED-CT code 399326009); breast cancer, on *malignant neoplasm of breast (disorder)* (code 254837009); colorectal cancer, on *malignant neoplasm of colon (disorder)* (code 363406005) or *malignant tumor of rectum (disorder)* (code 363351006); endometrial cancer, on *malignant neoplasm of endometrium of corpus uteri (disorder)* (code 188192002); kidney cancer, on *malignant tumor of kidney (disorder)* (code 363518003); leukemia, on *leukemia, disease (disorder)* (code 93143009); non-Hodgkin lymphoma (NHL), on *non-Hodgkin's lymphoma (disorder)* (code 118601006); prostate cancer, on *malignant tumor of prostate (disorder)* (code 399068003); liver cancer, on *malignant neoplasm of liver (disorder)* (code 93870000); lung cancer, on *malignant tumor of lung (disorder)* (code 363358000); melanoma, on *malignant melanoma (disorder)* (code 372244006); pancreatic cancer, on *malignant tumor of pancreas (disorder)* (code 363418001); and thyroid cancer, on *malignant tumor of thyroid gland (disorder)* (code 363478007).

The following analyses were performed. First, we examined how cancer was associated with the risk of COVID-19, adjusted for age, sex, race, comorbidities, cancer treatments, transplant procedures, and nursing home stay. The exposure groups were patients diagnosed with cancer, the unexposed groups were patients without cancer, and the outcome measure was the diagnosis of COVID-19. A separate analysis was performed for each of the 13 cancer types. Second, we examined how demographic factors affected COVID-19 risk among patients with cancer. The case groups were patients with cancer and one of the following demographic factors: female, senior (aged >65 years), and African American. The comparison groups were patients with cancer and 1 of the following corresponding demographic factors (male, adult [aged 18-65 years], and White race). The outcome measure was diagnosis of COVID-19. Third, the rates of death and hospitalization among patients with COVID-19 and cancer were compared with

Table. Patient Characteristics

Characteristic	Study group, No. (%) ^a					
	All patients (N = 73 449 510)	Patients with all common cancers (n = 2 523 920)	Patients with recent common cancers (n = 273 140)	Patients with COVID-19 (n = 16 570)	Patients with COVID-19 and all common cancer (n = 1200)	Patients with COVID-19 plus recent common cancers (n = 690)
Sex						
Female	39 400 370 (53.64)	1 337 230 (52.98)	140 230 (51.34)	9700 (58.54)	690 (57.50)	390 (56.52)
Male	33 544 260 (45.67)	1 178 990 (46.71)	129 220 (47.31)	6830 (41.22)	500 (41.67)	300 (43.48)
Unknown	505 680 (0.69)	7700 (0.31)	3700 (1.35)	40 (0.24)	10 (0.83)	0
Age						
Adult (18-65 y)	43 980 750 (59.88)	797 060 (31.58)	105 870 (38.76)	11 610 (70.07)	490 (40.83)	300 (43.48)
Senior (>65 y)	17 935 910 (24.42)	1 697 890 (67.27)	165 730 (60.68)	3900 (23.54)	690 (57.50)	370 (53.62)
Junior (<18 y)	10 553 420 (14.37)	8590 (0.34)	1990 (0.73)	1050 (6.34)	20 (1.67)	20 (2.90)
Race/ethnicity						
White	40 205 760 (54.74)	1 936 590 (76.73)	204 950 (75.03)	8200 (49.49)	670 (55.83)	380 (55.07)
African American	7 584 550 (10.33)	258 240 (10.23)	36 790 (13.47)	6590 (39.77)	480 (40.00)	270 (39.13)
Asian	1 189 390 (1.62)	46 330 (1.84)	4060 (1.49)	150 (0.91)	0	0
Hispanic/Latino	1 054 410 (1.44)	15 140 (0.60)	1440 (0.53)	10 (0.06)	0	0
Unknown	9 023 560 (12.29)	295 140 (11.69)	23 510 (8.61)	890 (5.37)	80 (6.67)	50 (7.25)
Insurance						
Private	25 822 090 (35.16)	1 089 500 (43.17)	128 620 (47.09)	2520 (15.21)	210 (17.50)	140 (20.29)
Medicare	7 630 360 (10.39)	1 026 120 (40.66)	106 510 (38.99)	780 (4.71)	160 (13.33)	90 (13.04)
Medicaid	6 114 970 (8.33)	140 790 (5.58)	17 510 (6.41)	660 (3.98)	40 (3.33)	30 (4.35)
Self-pay	4 899 910 (6.67)	124 070 (4.92)	6970 (2.55)	190 (1.15)	0	0
Unknown	6 190 770 (8.43)	162 660 (6.44)	9610 (3.52)	270 (1.63)	20 (1.67)	10 (1.45)

Abbreviation: COVID-19, coronavirus disease 2019.

^a The 13 common cancers are bladder, breast, colorectal, endometrial, kidney, prostate, thyroid, liver, lung, and pancreatic cancer; leukemia; non-Hodgkin lymphoma; and melanoma. The categories of race and insurance did not total

100% for the following reasons: (1) the Table only shows major subcategories of race or insurance; (2) not everyone in the electronic health record database has race or insurance information; and (3) a patient can report more than 1 race or insurance type.

those for patients with COVID-19 but no cancer and for patients with cancer but no COVID-19.

Statistical Analysis

The adjusted odds ratios (aORs), 95% CIs, and P values were calculated using the Cochran-Mantel-Haenszel method²⁵ by controlling for age groups (juniors [<18 years], adults, or seniors), sex (female or male), race (White or African American), and common comorbidities considered risk factors for COVID-19,² including asthma, cardiovascular diseases, type 2 diabetes, obesity, chronic kidney diseases, chronic obstructive pulmonary disease, cancer treatments (chemotherapy, radiotherapy, and immunotherapy), transplants (bone marrow and solid organ), and nursing home stay. Other demographic groups were not included owing to insufficient sample sizes for cases with COVID-19. Two-sided, 2-sample tests for equality of proportions with continuity correction were used to compare prevalence of comorbidities and outcomes. Multiple comparisons were analyzed by Bonferroni correction. Statistical tests were conducted with significance set at P < .05 (2 sided). All analyses were performed using R, version 3.6.3 (R Project for Statistical Computing).

Results

Patient Characteristics

The baseline characteristics of the study population (as of August 14, 2020) are presented in the **Table**. Among 73 449 510 pa-

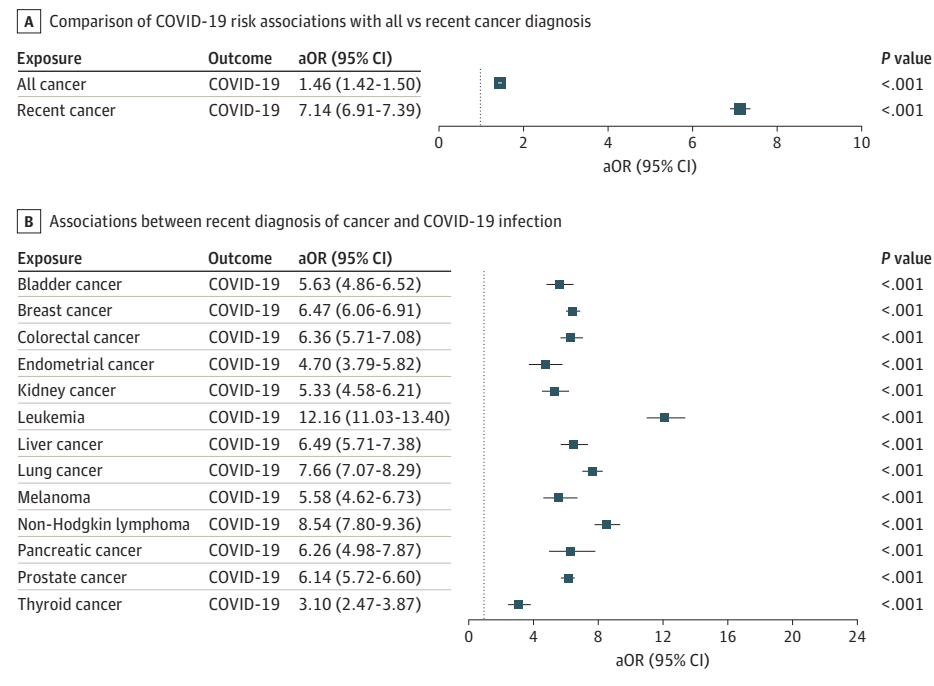
tients (among those with known data, 53.64% female and 45.67% male), 2 523 920 had at least 1 of the 13 common cancers diagnosed (all cancer diagnosed within or before the last year), and 273 140 had recent cancer (cancer diagnosed within the last year). These cancers included 138 890 bladder cancers, with 14 300 recent; 663 250 breast cancers, with 70 580 recent; 317 580 colorectal cancers, with 25 150 recent; 41 740 endometrial cancers, with 7750 recent; 124 170 kidney cancers, with 12 810 recent; 137 890 leukemias, with 16 930 recent; 193 140 liver cancers, with 15 070 recent; 419 050 lung cancers, with 34 830 recent; 198 890 melanomas, with 11 490 recent; 168 750 NHLs, with 26 460 recent; 70 950 pancreatic cancers, with 5280 recent; 487 560 prostate cancers, with 61 010 recent; and 109 870 thyroid cancers, with 14 140 recent.

Among 16 570 patients diagnosed with COVID-19, 1200 had a cancer diagnosis and 690 had a recent cancer diagnosis of at least 1 of the 13 common cancers. These included 50 any and 30 recent bladder cancer, 370 any and 180 recent breast cancer, 130 any and 60 recent colorectal cancer, 30 any and 20 recent endometrial cancer, 70 any and 30 recent kidney cancer, 100 any and 80 recent leukemia, 60 any and 40 recent liver cancer, 140 any and 100 recent lung cancer, 90 any and 20 recent melanoma, 120 any and 90 recent NHL, 20 any and 20 recent pancreatic cancer, 240 any and 140 recent prostate cancer, and 50 any and 20 recent thyroid cancer.

Odds of COVID-19 Infection for 13 Cancer Types

We examined the overall odds of COVID-19 infection for 13 common cancers with COVID-19 incidence for both all diagnoses

Figure 1. Associations of Coronavirus Disease 2019 (COVID-19) With Recent and All Cancer Diagnoses



and recent diagnoses after adjusting for age, sex, race, and potential COVID-19 risk factors,¹² including cardiovascular diseases, obesity, type 2 diabetes, asthma, chronic kidney diseases, chronic obstructive pulmonary disease, cancer treatments (chemotherapy, radiotherapy, and immunotherapy), transplant procedures (bone marrow and solid organ), and nursing home stay. Patients with cancer had a significantly increased risk for COVID-19 infection compared with patients without cancer, with the stronger effect for recent cancer (aOR, 7.14 [95% CI, 6.91-7.39]; $P < .001$) than all cancer diagnosis (aOR, 1.46 [95% CI, 1.42-1.50]; $P < .001$) (Figure 1), suggesting that patients with recent cancer diagnosis (ie, those with active cancer) are the most vulnerable among the population with any cancer (ie, both patients with active cancer and cancer survivors). Our subsequent analyses focused on patients with a recent cancer diagnosis.

We then examined associations of recent diagnosis of each of the 13 cancer types with COVID-19 infection after adjusting for age, sex, race, and COVID-19 risk factors. Patients with cancer had significantly increased risk for COVID-19 infection for all 13 cancer types, with the strongest effect for leukemia (aOR, 12.16 [95% CI, 11.03-13.40]; $P < .001$), NHL (aOR, 8.54 [95% CI, 7.80-9.36]; $P < .001$), lung cancer (aOR, 7.66 [95% CI, 7.07-8.29]; $P < .001$), liver cancer (aOR, 6.49 [95% CI, 5.71-7.38]; $P < .001$), and pancreatic cancer (aOR, 6.26 [95% CI, 4.98-7.87]; $P < .001$) and weakest but still significant effects for thyroid cancer (aOR, 3.10 [95% CI, 2.47-3.87]; $P < .001$) and endometrial cancer (aOR, 4.70 [95% CI, 3.79-5.82]; $P < .001$) (Figure 1).

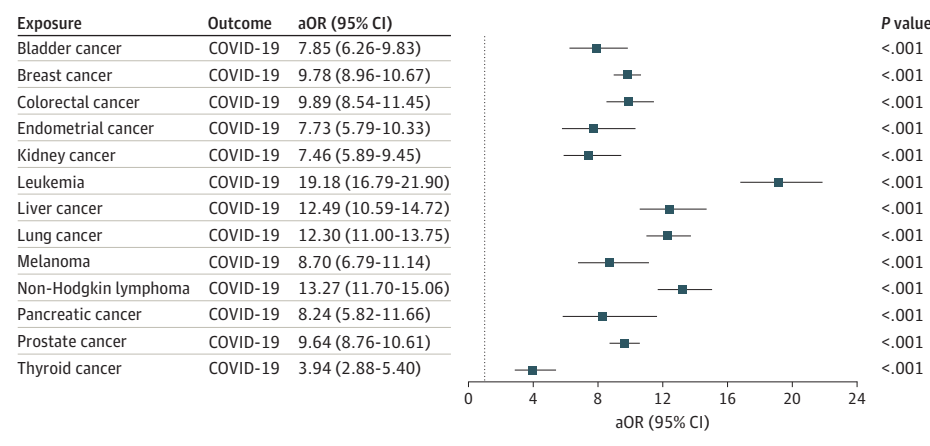
To examine how cancer itself and risk factors for COVID-19 (comorbidities, cancer treatments, transplant procedures, and nursing home stay), alone and together, affected the odds of

COVID-19 infection in patients with cancer, we calculated the odds of COVID-19 infection for each cancer type before adjusting for COVID-19 risk factors (Figure 2). These medical conditions and procedures indeed contributed to the increased risk for COVID-19 infection in patients with cancer, as evidenced by the reductions of aORs after adjusting for these comorbidities (Figure 2 vs Figure 1). For example, the odds of COVID-19 infection for leukemia decreased from 19.18 (95% CI, 16.79-21.90) to 12.16 after adjustment. However, even after adjusting for COVID-19 risk factors, patients with cancer still had a high risk for COVID-19 infection for all 13 cancer types, with aORs ranging from 12.16 for leukemia to 3.10 for thyroid cancer (Figure 1).

Demographic Disparity of COVID-19 Risk Among Patients With Cancer

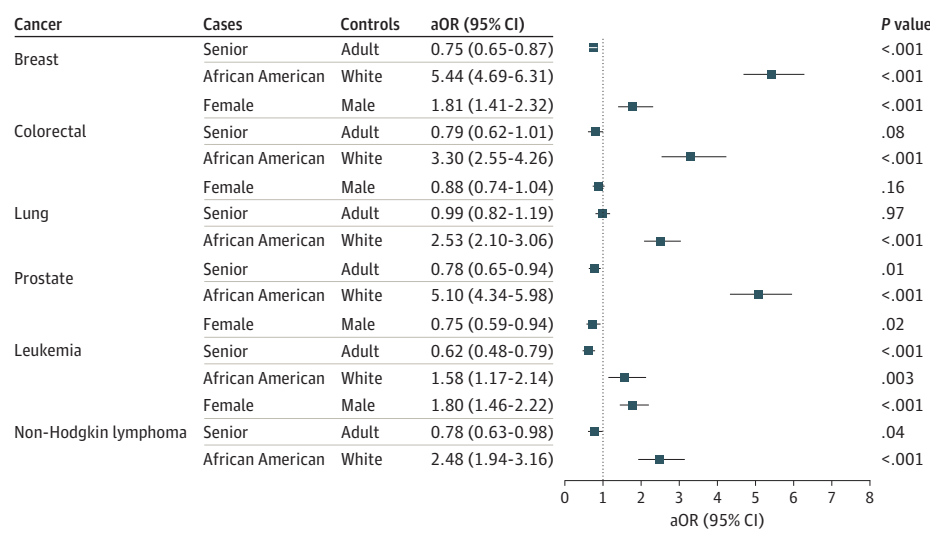
We examined how demographic factors (age, sex, and race) were associated with COVID-19 risk among patients with a recent cancer diagnosis after adjusting for common comorbidities, cancer treatments, transplant procedures, and nursing home stay. Six cancer types with a sufficient number of patients and COVID-19 cases were examined, including breast cancer, colorectal cancer, lung cancer, prostate cancer, leukemia, and NHL. Among patients with a recent diagnosis of these cancers, African American patients were more likely to be infected by COVID-19 than White patients after adjusting for age, sex, and COVID-19 risk factors, with the largest race disparity for breast cancer (aOR, 5.44 [95% CI, 4.69-6.31]; $P < .001$), prostate cancer (aOR, 5.10 [95% CI, 4.34-5.98]; $P < .001$), colorectal cancer (aOR, 3.30 [95% CI, 2.55-4.26]; $P < .001$) and lung cancer (aOR, 2.53 [95% CI, 2.10-3.06]; $P < .001$) and weakest but still significant for leukemia (aOR, 1.58 [95% CI, 1.17-

Figure 2. Association of Coronavirus Disease 2019 (COVID-19) Infection With Recent Cancer Diagnosis



Each of the 13 common cancer types is included as an exposure. COVID-19 was included as the outcome. Odds ratios were adjusted for age, sex, and race only. aOR indicates adjusted odds ratio.

Figure 3. Coronavirus Disease 2019 (COVID-19) Risk Disparity Among Patients With Recent Diagnoses of Cancer



Effects of demographics on odds of COVID-19 infection among patients with recent cancer (diagnosis made within the past year), after adjusting for age, sex, and risk factors for COVID-19, including asthma, cardiovascular diseases, type 2 diabetes, obesity, chronic kidney diseases, chronic obstructive pulmonary disease, cancer treatments (chemotherapy, radiotherapy, and immunotherapy), transplants (bone marrow and solid organ), and nursing home stay. Adult indicates patients aged 18 to 65 years; senior, patients older than 65 years.

2.14]; $P = .003$). Women had a higher risk of COVID-19 than men for colorectal cancer (aOR, 1.81 [95% CI, 1.41-2.33]; $P < .001$) and NHL (aOR, 1.80 [95% CI, 1.46-2.22]; $P < .001$). Age in general had no association (range of aORs for seniors, 0.62 [95% CI, 0.48-0.79] to 0.99 [95% CI, 0.82-1.19]) (Figure 3). Similar racial disparity was observed without adjusting for COVID-19 risk factors (eFigure in the Supplement).

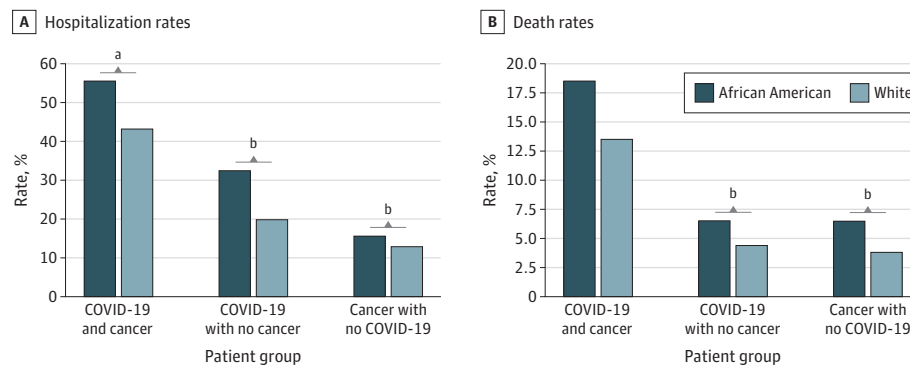
Hospitalization and Death Rates in Patients With Cancer and COVID-19

The overall hospitalization rate for 15 510 adult or senior patients with COVID-19 (aged ≥ 18 years) was 25.27%. Among 670 adult or senior patients with COVID-19 and recent cancer, 320 were hospitalized (47.76%), with higher rates for African American (150 of 270 [55.56%]) than White patients (160 of 370 [43.24%]) ($P = .003$). Among 14 840 adult or senior patients with COVID-19 and no cancer, 3600 were hospitalized (24.26%), with higher rates for African American (1900 of 5910 [32.15%]) than White patients (1370 of 7160 [19.13%]) ($P < .001$).

The hospitalization rate for the 270 380 adult or senior patients with cancer but no COVID-19 was 33 490 of 270 380 (12.39%), with higher rates for African American (5690 of 36 360 [15.65%]) than White patients (26 280 of 203 870 [12.89%]) ($P < .001$) (Figure 4). Overall, the hospitalization rate for patients with a recent cancer diagnosis and COVID-19 (320 of 670 [47.76%]) was higher than for patients with COVID-19 but no recent cancer diagnosis (3600 of 14 840 [24.26%]; $P < .001$) and that for patients with recent cancer diagnosis but no COVID-19 (33 490 of 270 380 [12.39%]; $P < .001$). COVID-19 and cancer had a synergistic effect on hospital rate, because the rate was greater than the sum of their individual effects.

The overall death rate for 15 510 adult and senior patients with COVID-19 was 5.61%. The death rate in 670 adult patients with COVID-19 and cancer was 100 of 670 (14.93%), similar for African Americans (50 of 270 [18.52%]) and White patients (50 of 370 [13.51%]; $P = .11$). The death rate in 14 840 adult patients with COVID-19 without cancer was 5.26%, higher for African American (400 of 6110 [6.55%]) than White patients

Figure 4. Hospitalization and Death Rates Among Adults



Patients were adults and seniors older than 18 years stratified by having both recent diagnosis of cancer and coronavirus disease 2019 (COVID-19), COVID-19 without a recent diagnosis of cancer, and a recent diagnosis of cancer but no COVID-19. The Systematized Nomenclature of Medicine–Clinical Terms concepts *hospital admission (procedure)* (code 32485007) was used to obtain hospitalization status from patient electronic health records. Exploratory

regularly imports from the Social Security Death Index for the deceased status. Outcomes for specific cancer types were not examined owing to their small sample sizes for COVID-19 cases and outcome information.

^a $P < .01$.

^b $P < .001$.

(330 of 7450 [4.43%]; $P < .001$). The death rate for the 270 380 adult and senior patients with cancer but no COVID-19 was 4.03%, with higher rates for African American (2360 of 36 360 [6.49%]) than White patients (7910 of 203 870 [3.88%]; $P < .001$) (Figure 4). Cancer had a synergistic effect on death rate because the rate was greater than the sum of their individual effects.

Discussion

Based on this analysis of a nationwide EHR database in the US, we found that patients with recently diagnosed cancer, particularly leukemia, lung cancer, and NHL, had significantly increased risk of COVID-19. Compared with White patients with cancer, African American patients with cancer had significantly higher risk of COVID-19. Cancer and COVID-19 had synergistic effects on patient outcome as measured by rates of hospitalization and death. This study identified high-risk groups of patients with cancer based on demographic groups and cancer types that are most vulnerable to COVID-19.

We showed that patients with cancer in general are at significantly increased risk for COVID-19 infection compared with patients without cancer after adjusting for COVID-19 risk factors. The odds of COVID-19 infection for patients with cancer decreased after adjusting for COVID-19 risk factors, indicating that these factors contributed to their risk for COVID-19 infections. However, even after adjusting for COVID-19 risk factors, patients with cancer still were at high risk for COVID-19 for all 13 cancer types. These results, especially the widely different aORs for different types of cancer and the high risk for blood cancers, indicate that differential systems affected by the cancer may affect a patient's risk for COVID-19 infection. In addition, a recent study showed that front-line health care workers were at increased risk for COVID-19.²⁶ Patients with cancer may be at high risk for COVID-19 because they often have high levels of contact with health care workers. In our study,

patients with more deadly cancers (cancers with poorer prognosis), such as pancreatic cancer and liver cancer (aORs, 6.26 and 6.49), did not have higher risk for COVID-19 infection compared with less deadly cancer such as leukemia and NHL (aORs, 12.16 and 8.54), suggesting that although high levels of health care contact might have contributed to increased risk for COVID-19 in patients with cancer, cancer itself had direct effect on a patient's risk for COVID-19 infection. This is further supported by the finding among 13 cancer types that blood cancers, which can directly change the way the immune system blood cells work, had highest risk associations with COVID-19 infection. Our finding is consistent with the fact that people with cancer are in general more likely to get infections because of changes in the immune system that control their body's defense systems.¹¹ Other factors, such as socioeconomic status and behavioral and lifestyle factors, may also contribute to COVID-19 infection in patients with cancer. However, owing to limited information available in the EHR database, we are unable to assess how these factors affected the risk for COVID-19 infection among patients with cancer.

African American patients with cancer were more likely to be infected by COVID-19 than White patients. This finding is consistent with data showing that COVID-19 affects African American individuals at a disproportionately high rate.^{27,28} Our study showed similar racial disparity for COVID-19 infection before and after controlling for COVID-19 risk factors, suggesting that other factors, such as social adversity, economic status, access to health care, and lifestyle, may have contributed to this profound racial disparity. Owing to limited socioeconomic, behavioral, and lifestyle information available in the EHR database, we are unable to assess how these factors contributed to this profound racial disparity in patients with cancer. The underlying mechanisms for the observed cancer-specific race disparity, with the largest disparity seen for breast cancer (aOR, 5.44), remain unclear and warrant further investigation. Although advanced age is a risk factor for cancer, our study showed that age had no additional effect on the risk for

COVID-19 infection among patients with cancer. Women had a higher risk of COVID-19 than men among patients with colorectal cancer and NHL. Reasons for this cancer-specific sex disparity are unclear and could reflect either a higher risk for infection or a higher likelihood of being tested.

Our study found that 5.61% of patients with COVID-19 died, consistent with the death rate of 5.9% reported by Centers for Disease Control and Prevention.¹ In our study, we found that patients with cancer and COVID-19 had significantly worse outcomes than patients with COVID-19 without cancer and patients with cancer without COVID-19. We observed synergistic effects between COVID-19 and cancer on both death and hospitalization rates. Biological mechanisms underlying this synergy warrant further investigation. African American patients with cancer had higher hospitalization rates than White patients, which is consistent with a recent report of disproportionately high rates of COVID-19 hospitalizations for the Black population.²⁹ However, African American patients with cancer were not more likely to die of COVID-19, as shown in our study.

Limitations

Our study is based on retrospective analysis of patient EHR data. Patient EHR data have been widely used for observational studies, including use of health care services, drug use, epidemiology (incidence and prevalence), risk factors, and safety surveillance.³⁰⁻³² However, patient EHR data have inherent limitations when used for research purposes: data are collected for billing purposes; often reflect underdiagnosis, overdiagnosis, or misdiagnosis; do not include all confounding factors; have limited time-series information for patients; have limited information on socioeconomic and lifestyle determinants; and lack finer-

grained information of admitting diagnoses of hospitalization and contributing causes of death, among others. The new concept COVID-19 was first included in the March 2020 SNOMED-CT International Edition Interim Release under the parent term *coronavirus infection*³³; however, at the time of this study (August 14, 2020), this update had not yet been incorporated in the Explorys EHR database. In our study, the status of COVID-19 was based on the parent term *coronavirus infection*. The number of COVID-19 cases in the database is significantly lower than the number of reported cases in the US. Because COVID-19 is regularly tested at drive-up and pop-up testing locations, it is possible that many of the reported cases may not be captured by diagnosis codes in EHRs. Despite these limitations, this large nationwide database allows us to identify early trends in risks, disparities, and outcomes of COVID-19 in patients with cancer engaged with health care systems on a nationwide basis.

Conclusions

Findings from this EHR-based case-control study on a nationwide population basis are associational, not causal. These associational findings need to be replicated in and compared with other EHR databases and patient registries such as the National Patient-Centered Clinical Research Network (PCORnet), Accrual to Clinical Trials (ACT) network, TriNetX, National COVID Cohort Collaborative (N3C), and the COVID-19 and Cancer Consortium (CCC19) registry. Our study provides an independent extension of 2 recent reports by the COVID-19 and Cancer Consortium registry^{14,15} and can serve as a baseline study of initial COVID-19 risk, racial disparity, and outcomes observations in patients with cancer across the US.

ARTICLE INFORMATION

Accepted for Publication: August 21, 2020.

Published Online: December 10, 2020.
doi:10.1001/jamaoncol.2020.6178

Author Contributions: Mr Wang and Dr Xu had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Berger, Xu.

Acquisition, analysis, or interpretation of data: Wang, Xu.

Drafting of the manuscript: Berger, Xu.

Critical revision of the manuscript for important intellectual content: All authors.

Statistical analysis: Wang, Xu.

Administrative, technical, or material support: Wang, Xu.

Supervision: Berger, Xu.

Conflict of Interest Disclosures: None reported.

Funding/Support: This study was supported by New Innovator Award DP2HD084068 from the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health (NIH); grants R01 AG057557, R01 AG061388, and R56 AG062272 from the NIH National Institute of Aging; Research Scholar Grant R5G-16-049-01-MPC from the American Cancer Society; grant IUL1TRO02548-01 from the Clinical and Translational Science Collaborative of

Cleveland; grant P30 CA043703 from the Case Comprehensive Cancer Center; and Cancer Health Disparities SPORE Planning Grant P20 CA2332216 from the Case Comprehensive Cancer Center.

Role of the Funder/Sponsor: The sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

REFERENCES

- Centers for Disease Control and Prevention (CDC). Cases in the US. Published August 2020. Accessed August 17, 2020. <https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html>
- Centers for Disease Control and Prevention (CDC). Evidence used to update the list of underlying medical conditions that increase a person's risk of severe illness from COVID-19. Updated October 6, 2020. Accessed August 14, 2020. <https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/evidence-table.html>
- Richardson S, Hirsch JS, Narasimhan M, et al; the Northwell COVID-19 Research Consortium. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area. *JAMA*. 2020; 323(20):2052-2059. doi:10.1001/jama.2020.6775
- Myers LC, Parodi SM, Escobar GJ, Liu VX. Characteristics of hospitalized adults with COVID-19 in an integrated health care system in California. *JAMA*. 2020;323(21):2195-2198. doi:10.1001/jama.2020.7202
- Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*. 2020;395(10229):1054-1062. doi:10.1016/S0140-6736(20)30566-3
- Schreiber RD, Old LJ, Smyth MJ. Cancer immunoeediting: integrating immunity's roles in cancer suppression and promotion. *Science*. 2011; 331(6024):1565-1570. doi:10.1126/science.1203486
- Ogle KS, Swanson GM, Woods N, Azzouz F. Cancer and comorbidity: redefining chronic diseases. *Cancer*. 2000;88(3):653-663. doi:10.1002/(SICI)1097-0142(20000201)88:3<653::AID-CNCR24>3.0.CO;2-1
- Lauby-Secretan B, Scoccianti C, Loomis D, Grosse Y, Bianchini F, Straif K; International Agency for Research on Cancer Handbook Working Group. Body fatness and cancer: viewpoint of the IARC Working Group. *N Engl J Med*. 2016;375(8):794-798. doi:10.1056/NEJMs1606602

9. Giovannucci E, Harlan DM, Archer MC, et al. Diabetes and cancer: a consensus report. *Diabetes Care*. 2010;33(7):1674-1685. doi:10.2337/dc10-0666
10. Kitsis RN, Riquelme JA, Lavandro S. Heart disease and cancer: are the two killers colluding? *Circulation*. 2018;138(7):692-695. doi:10.1161/CIRCULATIONAHA.118.033907
11. American Cancer Society. Why people with cancer are more likely to get infections. Published March 2020. Accessed August 14, 2020. <https://www.cancer.org/treatment/treatments-and-side-effects/physical-side-effects/low-blood-counts/infections/why-people-with-cancer-are-at-risk.html>
12. Yu J, Ouyang W, Chua MLK, Xie C. SARS-CoV-2 transmission in patients with cancer at a tertiary care hospital in Wuhan, China. *JAMA Oncol*. 2020;6(7):1108-1110. doi:10.1001/jamaoncol.2020.0980
13. Yang K, Sheng Y, Huang C, 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):904-913. doi:10.1016/S1470-2045(20)30310-7
14. Kuderer NM, Choueiri TK, Shah DP, et al; COVID-19 and Cancer Consortium. Clinical impact of COVID-19 on patients with cancer (CCC19): a cohort study. *Lancet*. 2020;395(10241):1907-1918. doi:10.1016/S0140-6736(20)31187-9
15. Rivera DR, Peters S, Panagiotou OA, et al; COVID-19 and Cancer Consortium. Utilization of COVID-19 treatments and clinical outcomes among patients with cancer: a COVID-19 and Cancer Consortium (CCC19) cohort study. *Cancer Discov*. 2020;CD-20-CD-0941.
16. Robilotti EV, Babady NE, Mead PA, et al. Determinants of COVID-19 disease severity in patients with cancer. *Nat Med*. 2020;26(8):1218-1223. doi:10.1038/s41591-020-0979-0
17. IBM Corporation. IBM Explorers EHR solutions. Published August 2020. Accessed August 14, 2020. <https://www.ibm.com/watson-health/about/explorers>
18. 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):965-972. doi:10.1136/amiajnl-2011-000782
19. Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Res*. 2004;32(Database issue):D267-D270. doi:10.1093/nar/gkh061
20. SNOMED International. The Systematized Nomenclature of Medicine—Clinical Terms (SNOMED CT). Published June 2020. Accessed August 14, 2020. <http://www.snomed.org/snomed-ct/why-snomed-ct>
21. IBM Watson Health. IBM Explorers EHR Database Bibliography Categorized by Therapeutic Area. Published June 2020. Accessed August 14, 2020. <https://www.ibm.com/downloads/cas/RPKNLLIM>
22. Zhou M, Xu R, Kaelber DC, Gurney ME. Tumor necrosis factor (TNF) blocking agents are associated with lower risk for Alzheimer's disease in patients with rheumatoid arthritis and psoriasis. *PLoS One*. 2020;15(3):e0229819. doi:10.1371/journal.pone.0229819
23. Zhou M, Zheng C, Xu R. Combining phenome-driven drug-target interaction prediction with patients' electronic health records-based clinical corroboration toward drug discovery. *Bioinformatics*. 2020;36(suppl 1):i436-i444. doi:10.1093/bioinformatics/btaa451
24. National Cancer Institute (NCI). Common cancer types. Updated September 25, 2020. Accessed June 28, 2020. <https://www.cancer.gov/types/common-cancers>
25. Kuritz SJ, Landis JR, Koch GG. A general overview of Mantel-Haenszel methods: applications and recent developments. *Annu Rev Public Health*. 1988;9:123-160. doi:10.1146/annurev.pu.09.050188.001011
26. Nguyen LH, Drew DA, Graham MS, et al. Risk of COVID-19 among front-line health-care workers and the general community: a prospective cohort study. *Lancet Public Health*. Published online July 30, 2020. doi:10.1016/S2468-2667(20)30164-X
27. Yancy CW. COVID-19 and African Americans. *JAMA*. 2020;323(19):1891-1892. doi:10.1001/jama.2020.6548
28. Mahajan UV, Larkins-Pettigrew M. Racial demographics and COVID-19 confirmed cases and deaths: a correlational analysis of 2886 US counties. *J Public Health (Oxf)*. 2020;42(3):445-447. doi:10.1093/pubmed/fdaa070
29. Karaca-Mandic P, Georgiou A, Sen S. Assessment of COVID-19 hospitalizations by race/ethnicity in 12 states. *JAMA Intern Med*. Published online August 17, 2020. doi:10.1001/jamainternmed.2020.3857
30. Cowie MR, Blomster JI, Curtis LH, et al. Electronic health records to facilitate clinical research. *Clin Res Cardiol*. 2017;106(1):1-9. doi:10.1007/s00392-016-1025-6
31. Coorevits P, Sundgren M, Klein GO, et al. Electronic health records: new opportunities for clinical research. *J Intern Med*. 2013;274(6):547-560. doi:10.1111/joim.12119
32. Ahmad FS, Chan C, Rosenman MB, et al. Validity of cardiovascular data from electronic sources: the Multi-Ethnic Study of Atherosclerosis and HealthLNK. *Circulation*. 2017;136(13):1207-1216. doi:10.1161/CIRCULATIONAHA.117.027436
33. SNOMED CT and COVID-19. Published March 2020. Accessed June 28, 2020. <https://www.snomed.org/snomed-ct/covid-19>

Invited Commentary

Use of Real-World Electronic Health Records to Estimate Risk, Risk Factors, and Disparities for COVID-19 in Patients With Cancer

Aakash Desai, MBBS, MPH; Ali Raza Khaki, MD; Nicole M. Kuderer, MD

The coronavirus disease 2019 (COVID-19) pandemic has introduced additional serious risks for patients with cancer and is deeply affecting current cancer care. Early data from observational studies have suggested high rates of hospitalization and early mortality.¹ However, even more granular data are needed to identify specific populations with cancer at greatest risk of serious outcomes from COVID-19, as well as to understand the risk factors that predispose patients to develop COVID-19.

In this issue of *JAMA Oncology*, Wang et al² present results from a retrospective case-control study using IBM Watson Health Explorers electronic health record (EHR) data to assess the association of cancer and demographic factors with COVID-19 outcomes. Among the 73 449 510 patients, 2 523 920 were identified with any cancer diagnosis and 273 140 with a recent cancer

diagnosis (within 1 year before infection). Included patients had 1 of 13 common malignant neoplasms. Among 16 570 patients identified with COVID-19 as of August 14, 2020, 1200 had any cancer diagnosis and 690 had a recent cancer diagnosis. The study found that patients with any or a recent cancer diagnosis were at substantially increased risk for COVID-19 compared with patients without cancer (adjusted odds ratios [AORs], 1.46 [95% CI, 1.42-1.50] and 7.14 [95% CI, 6.91-7.39], respectively, adjusted for major COVID-19 risk factors). The recent cancer diagnoses with the greatest risk were leukemia (AOR, 12.16 [95% CI, 11.03-13.40]), non-Hodgkin lymphoma (AOR, 8.54 [95% CI, 7.80-9.36]), and lung cancer (AOR, 7.66 [95% CI, 7.07-8.29]). Higher rates of hospitalization (47.46%) and mortality (14.93%) were observed in those with concurrent diagnoses of recent cancer and COVID-19 than expected with either diagnosis alone, suggesting a potential synergistic effect.²