**Cancer and COVID-19: Epidemiological Analysis of 31,880 Patients from a 3.8 Million Clinical Cohort**

**Liang Zhu1, Huili Zhu2, Xiaojin Li1, Yan Huang1, Youngran Kim1, Heather Bush3, Caroline Chung4, Guo-Qiang Zhang1,\***

1 The University of Texas Health Science Center at Houston, Houston, Texas, USA

2 Icahn School of Medicine at Mount Sinai, New York, New York, USA

3 The University of Kentucky, Lexington, Kentucky, USA

4 The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

\*Corresponding author.

**SUMMARY**

**Background** We characterized and compared the outcomes of subgroups of COVID-19 individuals with and without cancer diagnosis, and identified potential prognostic factors for mortality and severe illness in each cohort.

**Methods** We analyzed a de-identified Electronic Health Records (EHR) data on individuals with and without cancer diagnosis, and with confirmed severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection from the US Optum® Covid-19 EHR dataset with records dating back from January 1st, 2007. This EHR dataset contains information about baseline clinical condition, medication, cancer diagnosis and treatment, and COVID-19 disease status. The primary endpoint was all-cause mortality within 30 days of diagnosis of COVID-19. In both cancer and non-cancer subgroups, we assessed the association between the outcome and potential prognostic variables using Cox regression analysis. The secondary endpoint was a composite variable for COVID-19 severity consisting of mortality, severe illness requiring admission to hospital, admission to an intensive care unit (ICU), and mechanical ventilation. We assessed the association between severity and potential prognostic variables using logistic and multinomial regression analysis. We further performed matching of COVID-19 individuals with cancer diagnosis to those without cancer diagnosis on the variables of age, gender, race and ethnicity, region, and the numbers of established and possible comorbidities. After matching, effects of cancer on COVID-19 outcomes were analyzed using Cox regression and logistic regression.

**Findings** Of the 598,817 COVID-19 positive individuals in the 3.8 million Optum® dataset, 546,418 met inclusion criteria for our analysis. Among them, 31,880 had cancer diagnosis (median age, 67 years; interquartile range [IQR]: 57-77, male: 45.5%), and 514,538 did not (median age, 48 years; IQR: 33-61, male: 43.8%). Among individuals with cancer diagnosis, 85.3% had a solid tumor, 13.7% had hematologic malignancy, and 19.1% had survived for more than 5 years. Compared to individuals without cancer diagnosis, those with cancer diagnosis were older, more likely to be white, had more comorbidities, and were more likely to have recent surgery. They also had poorer outcomes in all aspects: 30-day mortality rate (6.4% vs 2.1%), admission to hospital (36.1% vs 16.7%), admission to ICU (6.6% vs 3.0%), and need for mechanical ventilation (3.3% vs 1.6%). Older age, male, non-Hispanic black, higher number of comorbidities, and recent surgery were risk factors for both mortality and other severe outcomes. For individuals with cancer diagnosis, having hematologic malignancy or having recent chemotherapy or radiation therapy were risk factors for mortality and severe outcomes as well. However, after matching on the common risk factors, the difference of outcomes between the two subgroups significantly decreased: 30-day mortality rate (6.5% vs 6.1%), admission to hospital (36.0% vs 31.6%), admission to ICU (6.6% vs 6.5%), and need for mechanical ventilation (3.3% vs 3.7%). The effect of cancer on mortality are not significant (hazard ratio [HR]: 1.03, 95% 0.97-1.10, p=0.324). The effect of cancer diagnosis on combined severity is significant but with a smaller effect (odds ratio [OR]: 1.21, 95% 1.17-1.26, p<0.0001).

**Interpretation** Cancer survivors had higher 30-day all-cause mortality and poorer outcomes after getting COVID-19 compared to individuals without cancer diagnosis. However, the differences were mainly due to the older age and more comorbidities in individuals with cancer diagnosis instead of cancer itself.

**INTRODUCTION**

An estimated 17 million individuals with or survived cancer live in the United States1 and millions more around the world. With nearly 125 million SARS-CoV-2 infection cases identified to date2, understanding the relationship between cancer and Covid-19 has become a challenging but pressing topic. Early studies from China and Italy showed that Covid-19 had a disproportionate effect on cancer patients3-6, particularly those who were undergoing chemotherapy or had recent surgery4,7. Factors found to be associated with increased mortality include increased age, male sex, smoking, number of comorbidities, and active or metastatic cancer8-10, especially lung cancer11-13. Advanced age, male gender, and comorbidities such as hypertension and diabetes have shown in the general population to be associated with greater risk of developing acute respiratory distress syndrome and death14. More recent studies have identified healthcare disparity and non-Hispanic black and Hispanic ethnicities as additional risk factors for cancer and poorer COVID-19 outcomes15. While the time-sensitive nature of these studies is appreciated, the majority of such studies involves a small number of patients from a single hospital system. Therefore, findings from such studies need to be interpreted in appropriate contexts, and studies involving larger cohorts may provide unique insights.

A particular question that can benefit from a larger-scale epidemiological analysis is whether cancer serves as an independent risk factor for acquiring COVID-19 and for its associated negative outcomes16. We consider the extent at which cancer diagnosis contributes to poorer Covid-19 outcomes as a distinct topic from how Covid-19 impacts those who are undergoing treatment. Before attributing poorer COVID-19 outcome to a patient’s cancer diagnosis, we need to first clarify whether it acts as an independent factor. The answer to this question will provide additional insight into the role of cancer for COVID-19 severity and help inform clinical management of this patient population. Because of the inherent heterogeneity of the population, a sufficiently large cancer patient population is required to match age, sex, and other known COVID-19 risk factors with adequate control of potential confounders. In this study, we use the Optum®US Covid-19 Electronic Health Record (EHR) data to elucidate Covid-19 outcomes between cancer and non-cancer subgroups.

The primary objective of the study is check whether just a simple knowledge of a history of cancer is a risk factor to outcome after coronavirus infection for a patient. If yes, we want to know which cancer-related factors (active treatment, cancer type) contributes to this result (doctors will need to pay more attention if a COVID patient has such a characteristic). The primary outcome of interest is all-cause comorbidity after 30 days after COVID diagnosis date. The secondary outcome of interest will be severity.

Literature review:

The advantage of EHR cancer studies and shortcoming compared with localized studies or observational studies specifically designed for this purpose.

Concern about big data. If systemic bias exists, big data doesn’t solve it (exhausting sampling solve it). Missing doesn’t necessarily lead to systemic bias, it depends on the mechanism of missing and how missing is handled. Big data itself may not solve missing issue (if we still use naïve missing handling which just create subgroup from the big data (similar to another sampling introduce systemic bias)).

Analysis methods

A descriptive table will be given for cancer and non-cancer group. P-value for association between continuous variable and cancer group by chi-square tests and by student’s t-test for continuous variable. The mortality is a time-to-event outcome. Cox proportional hazard model (cox-ph model). A propensity score matching was also used to check the. Matching was based on covariates (details). A cox-ph model was then used to explore potential important cancer-related factor to outcome in the cancer group.

Supplementary Methods

Matching

Missing data handling

Pairwise deletion

**METHODS**

**Data Source**

This study uses Optum®de-identified COVID-19 Electronic Health Record (EHR) data. To meet the urgent need to understand the clinical impact of SARS-CoV-2 infection amidst this global pandemic, Optum® developed a data pipeline with minimal time lag while preserving as much clinical information as possible. The COVID-19 data are sourced from Optum’s longitudinal EHR repository derived from more than 700 hospitals and 7000 clinics in the U.S. with patient medical history dating back to January 1, 2007. The January 28, 2021 release used in this study includes EHR data for 3.8 million unique individuals with a documented COVID-19 test on or after February 1, 2020. Enrollees are included regardless of test results. The Optum® COVID-19 EHR data cover patient-level, longitudinal records including demographics, diagnoses, procedures, lab tests, care settings, medications prescribed or administered, and mortality.

Our study included individuals aged 18 or older and those with a diagnosis of COVID-19 infection. Diagnosis of COVID-19 infection was determined if the individual had an EHR entry with a diagnosis of COVID-19 identified using the International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) codes (U07x) or had a confirmed polymerase chain reaction (PCR) test, antibody test, or antigen test. Enrollees were assigned to cancer groups according to any record of ICD-9 or 10 codes indicating cancer prior to COVID-19. To avoid misclassification, individuals were excluded if the first record a diagnosis code of cancer occurred on and after the date of COVID-19 infection confirmation.

This study used de-identified EHR data and was determined by the Committee for the Protection of Human Subjects (CPHS) at The University of Texas Health Science Center at Houston as non-human-subjects research.

**Outcomes**  
The primary outcome was all-cause mortality within 30 days of the initial date of COVID-19 infection. Secondary outcomes included development of severe illness, defined as death, hospital admission, intensive care unit (ICU) admission, mechanical ventilation, or a combination of these. We analyzed illness severity as both a dichotomous variable (none versus one of any of death, hospitalization, ICU, or ventilation) and a multinomial variable (0: none; 1: hospitalization only; 2: hospitalization + ICU/ventilation without death; 3: death).

**Cancer Definition**

Note: This is your primary “exposure.” This should include specifications of the definition of “cancer.” This should include codes and any subsequent cancer variables that are to be compared.

For individuals with cancer diagnosis, we also considered cancer type (LIST), recent chemotherapy (LIST), recent radiation therapy, age at cancer, and years of survival as potential risk factors. Recent surgery, recent chemotherapy, and recent radiation therapy were defined as within 4 weeks before the COVID-19 diagnosis.

**Potential risk factors**The potential risk factors included age, gender, race, ethnicity, number of comorbidities, and recent surgery.

Comorbidities were defined according to CDC’s classification [CITE] on established or possible risk factors for severe COVID-19. Specifically, established risk factors included cancer, chronic kidney disease, chronic obstructive pulmonary disease, Down syndrome, immunocompromised state from a solid organ transplant, obesity, pregnancy, serious cardiovascular disease, heart failure, coronary artery disease, cardiomyopathies, sickle cell disease, smoking, and type 2 diabetes mellitus. Since cancer was the primary focus of this study, we defined the number of comorbidities as the total number of conditions mentioned earlier other than cancer. For possible comorbidities, the list included asthma, cerebrovascular disease, cystic fibrosis, hypertension or high blood pressure, other immunodeficiencies, liver disease, dementia, overweight, pulmonary fibrosis, thalassemia, and type 1 diabetes mellitus.

**Matching**  
  
**Statistical Analysis (see comments about matching)**  
Descriptive statistics for demographic and patient characteristics are provided for matched and unmatched sample by cancer groups. Cox regression models are used to examine the association between time to death and cancer. For composite outcomes, logistic regression models and multinomial regression models are used. Kaplan-Meier survival curves are provided to visualize survival estimates between cancer and non-cancer groups. For those with a cancer diagnosis subgroup analyses were performed to examine the association of cancer-related treatments and COVID-19 outcomes. A p value less than 0.05 is considered as statistically significant. All data analyses use SAS software version 9.4 and all tests were conducted for both unmatched and matched samples.   
  
**RESULTS**

Out of 3,832,315 individuals in the Optum® COVID-19 HER data set, 598,817 were diagnosed with COVID-19 and reported one or more COVID-19 related outcomes. After excluding individuals under 18 (n=50,322) and those with cancer after COVID-10 diagnosis (n=2,077) among adults, the analytic sample included 546,418 individuals. These individuals were further grouped into two groups: those with a cancer diagnosis (31,880) and those without (514,538; see Figure 1).

Demographic, clinical, and tumor characteristics for the two groups are presented in Table 1. Individuals with cancer diagnosis were older than those without, with a median age of 67 years (IQR 57-77) versus 48 years (IQR 33-61). The percentage of individuals over 50 was 87.2% vs 46.7%. The distribution of gender was similar between the two groups (Male: 45.5% vs 43.8%). The percentage of non-Hispanic white was higher in individuals with cancer diagnosis (72.2% vs 59.4%), and the percentage of non-Hispanic black was similar (11.1% vs 10.8%). Individuals with cancer diagnosis had more established and possible comorbidities than those without cancer diagnosis, with mean numbers of established comorbidities to be 2.2 (SD=1.7) vs 1.2 (SD=1.3), and mean numbers of possible comorbidities as 2.1 (SD=1.1) vs 1.3 (SD=1.0). Individuals with a cancer diagnosis had a higher rate in each specific item of established and possible comorbidities except Down syndrome, overweight, obesity, and pregnancy (Table A.1).

After matching…

What are the percentages being presented in Table 3? Each column N should be the denominator but this doesn’t seem to be the case.

**Mortality:** The crude 30-day mortality rate in individuals with and without cancer diagnosis was 6.4% (95% CI: 6.17% - 6.71%) and 2.07% (95% CI: 2.03% - 2.10%), respectively (Table 2). The crude mortality rates for different levels of each variable are presented in Table 3 and the results from Cox regression model are presented in Table 4. Age older than 50, male, non-Hispanic black, number of established comorbidities, and recent surgery were risk factors of mortality in both groups (Table 4). Number of possible comorbidities is also a risk factor of mortality in individuals without cancer diagnosis, but not for individuals with cancer diagnosis. The effect of older age was stronger in individuals without cancer diagnosis, with the HR of 5.14 (95% CI: 4.62-5.72) for the age of 50 to 65, 16.45 (95% CI: 14.83-18.24) for the age of 65 to 75, 53.51 (95% CI: 48.44-59.12) for age ≥ 75 compared to age under 50. For comparison, the HR for individuals with cancer diagnosis are 1.80 (95% CI: 1.29-2.51) for age between 50 and 65, 3.79 (95% CI: 2.75-5.21) for age between 65 and 75, and 9.23 (95% CI: 6.75-12.61) for age ≥ 75 compared to those aged under 50. All other effects are comparable for individuals with and without cancer diagnosis.

For individuals with cancer diagnosis, hematologic malignancy had a higher risk of mortality (HR=1.47, 95% CI: 1.32-1.65) compared to solid tumor. Recent chemotherapy (HR=1.21, 95% CI: 1.06-1.37) and recent radiation therapy (HR=2.50, 95% CI: 1.93-3.24) were both risk factors for mortality. Individuals who survived over 5 years had a lower risk of mortality (HR=0.75, 95% CI: 0.66-0.84) (Table 4). After matching on related factors, the 30-day mortality rate in individuals with and without cancer diagnosis was 6.5% (95% CI: 6.18% - 6.73%) and 6.1% (95% CI: 5.88% - 6.41%), respectively. The hazard ratio of mortality for individuals with cancer diagnosis versus those without was 1.03(95% CI: 0.97-1.10) with a p value of 0.324 after adjusting for the other factors (Table 7). Kaplan-Meier plots for the two groups before and after matching are presented in Figure 2.

**Composite outcome:** For other severe outcomes, the rate of admission to hospital was 36.1% (95% CI: 35.54% - 36.59%) vs 16.7% (95% CI: 16.61% - 16.81%) in individuals with and without cancer diagnosis. The rate of admission to ICU was 6.6% (95% CI: 6.31% - 6.86%) vs 3.04% (95% CI: 2.99% - 3.09%). The rate of mechanical ventilation was 3.3% (95% CI: 3.11% - 3.50%) vs 1.56% (95% CI: 1.53% - 1.60%). The rate for composite outcome was 38.3% (95% CI: 37.74% - 38.81%) vs 17.6% (95% CI: 17.49% - 17.69%) (Table 2). The rates of these outcomes for different levels of each variable are presented in Table 3. Logistic regression result is presented in Table 4. In both groups, all the risk factors for mortality remained to be risk factors for the composite endpoint. However, individuals between 50 and 65 with cancer diagnosis did not have a significant risk of severe outcome compared to those younger than 50. Moreover, being Hispanic and having higher number of possible comorbidities were also significant risk factors for the composite endpoint in both groups (Table 4). We also analyzed severe outcomes using three levels: hospitalization only, hospitalization + ICU/ventilation without death, and death. The significant risk factors were similar to those identified by the logistic regression and were consistent among the three outcome levels. However, those aged between 50 and 65 in the cancer group and male in non-cancer group were associated with lower risk of hospitalization compared to no severe outcomes (Table 5).

After matching on risk factors, the rate of admission to a hospital was 36.0% (95% CI: 35.48% - 36.55%) vs 31.6% (95% CI: 31.05% - 32.08%) in individuals with and without cancer diagnosis. The rate of admission to ICU was 6.6% (95% CI: 6.31% - 6.86%) vs 6.5% (95% CI: 6.27% - 6.82%). The rate of mechanical ventilation was 3.3% (95% CI: 3.10% - 3.50%) vs 3.7% (95% CI: 3.49% - 3.91%). Composite endpoint rate was 38.2% (95% CI: 37.67% - 38.75%) vs 33.8% (95% CI: 33.32% - 34.37%). The odds ratio for cancer versus non-cancer was 1.21 (95% CI: 1.17 - 1.26) after adjusting for risk factors, with a p value <0.0001 (Table 7).

**DISCUSSION**

Patient stratification is paramount in the management of both cancer and COVID-19 as it allows caregivers to tailor treatment and appropriately allocate resources. With cancer among the leading causes of death worldwide, identifying risk factors in cancer survivors has become a focus during the global SARS-CoV-2 pandemic. We conducted one of the largest clinical cohort studies of COVID-19 positive individuals with and without cancer diagnosis to date. Our analysis showed that cancer survivors had higher 30-day all-cause mortality and poorer outcomes after contracting COVID-19 compared to individuals without cancer diagnosis. However, the differences were mainly due to the older age and having more comorbidities in individuals with a cancer diagnosis instead of cancer itself.

The 31,880 individuals with a cancer diagnosis in our study were older and had more established and possible comorbidities than did the 514,538 patients without a cancer diagnosis. In the general population, many existing studies14,17,18 and the analytics platform OpenSAFELY19 have consistently identified age, male gender, smoking status, and comorbidities (such as hypertension, diabetes, cardiovascular diseases, or chronic lung diseases) to be risk factors for severe COVID-19 outcomes. We also found age older than 50, male gender, number of established comorbidities (smoking is one of comorbidities), and recent surgery to be risk factors of mortality in both individuals with and without cancer. Of note, advanced age had a stronger effect in individuals without cancer diagnosis. These findings demonstrated similar risk factors in cancer patients as predictors of severe COVID-19 as in the general population.

The primary outcome of crude 30-day mortality in individuals with cancer diagnosis was higher than in those without. We aimed to clarify whether this effect is due to cancer history alone or due to the risk factors aforementioned. Many meta-analysis and systematic reviews have found individuals with cancer to be more likely to experience severe COVID-19 illness and death.10,13,20-22 However, an early Italian study3 and another North London study23 concluded that prevalence of cancer was not associated with risk of infection. Another meta-analysis noted in subgroup analysis of patients older than 65 that all-cause mortality was comparable between those with and without cancer.24 After matching on related factors, we found the 30-day mortality rate in individuals with and without cancer diagnosis to be comparable. Thus, our findings add support to the hypothesis that cancer patients are more likely to have poorer COVID-19 outcome as a result of older age and having more comorbidities.

Among patients with cancer diagnosis, we found hematologic malignancies and recent chemotherapy to be risk factors of mortality, consistent with established literature.7,9,16,25-28 In addition, our study is one of the first at such a large scale to identify recent radiation therapy as a risk factor of mortality. We defined “recent” as within the past 4 weeks of COVID-19 diagnosis, but currently there is no consistent definition of “recent” in existing studies, commonly seen as varying from 2 weeks to 3 months.

Literature shows that lung cancer is unsurprisingly associated with increased severity and mortality.11,12 In patients with a diagnosis of lung cancer in our dataset, we found that both the mortality and severity rates are higher, i.e., 12.5% (95% CI: 11.10% -13.99%) for the crude mortality rate, 55.1% (95% CI: 52.88% - 57.23%) for hospitalization, 12.6% (95% CI: 11.19% - 14.10%) for ICU, 5.6% (95% CI: 4.57% - 6.58%) for ventilation, and 58.8% (95% CI: 56.68% - 60.99% for the composite severity endpoint. After matching on age, gender, race and ethnicity, region, numbers of established and possible comorbidities, the effect of lung cancer remains significant for both mortality (HR=1.64, 95% CI: 1.34 – 2.01, p<0.0001) and severe outcome (OR=1.86, 95% CI: 1.62 – 2.13, p<0.0001).

We found that admission to hospital, admission to ICU, and mechanical ventilation rates to be higher in patients with cancer diagnosis. When outcomes were further categorized into hospitalization only, hospitalization with ICU and mechanical ventilation without death, and death, significant risk factors were similar across the three levels. After matching on related factors, the differences became less prominent. Overall, the cancer group had slightly higher rates than non-cancer group (OR: 1.21, 95% 1.17-1.26, p<0.0001). We reaffirmed previous studies7,24,29-31 and contributed to the risk stratification of COVID-19 patients with and without cancer diagnosis. These findings can help oncologists, intensivists, and other providers better grasp important clinical outcomes, such as mortality and need for ICU admission.

Social determinants of health—including poverty, physical environment (i.e., smoke exposure, crowded living spaces, poor access to healthcare facilities), and race or ethnicity—undoubtedly affect COVID-19 outcomes.32 Appropriate alarm and caution should be taken to address these healthcare disparities.33,34 Studies have reported disproportionately high fatalities and low access to clinical trials within black communities.15,35,36 In our study, the percentage of non-Hispanic black was similar between the two cohorts and it was identified as a risk factor of mortality in both populations. Racial and healthcare disparities are multifactorial; therefore, additional data on socioeconomic status should be collected alongside medical history to further examine the inequities in treatment exposures and outcomes.

For quality control purposes, we also studied the population where the COVID positive was defined by a positive PCR test, regardless of the result of ICD diagnose code, antibody, or antigen test. We identified 414,889 patients as COVID positive with this new definition, which is 69% of patients previously identified as positive. Though the absolute number of patients reduced, the proportion of patients with a cancer diagnosis, the distributions of demographic, comorbidities, and cancer treatment, and the proportion of the outcomes for groups with and without cancer diagnosis remained similar. As a result, the hazard ratios from the Cox regression model and the odds ratios from the logistic regression model were similar to those based on COVID positive patients identified previously. After matching on age, gender, race and ethnicity, region, numbers of established and possible comorbidities, the effect of cancer were not significant for mortality (HR=1.04, 95% CI: 0.96 – 1.11, p=0.351), but significant for severity (OR=1.20, 95% CI: 1.15 – 126, p<0.0001), similar to our previous conclusion based on the general definition of COVID positive.

We acknowledge the considerable heterogeneity in cancer diagnosis and types of chemotherapy. Further large-scale prospective studies are needed to stratify between solid and hematologic malignancies as well as the various categories of chemotherapy (i.e., cytotoxic, targeted, immunologic, or hormonal). Another limitation is that the exact death date was not recorded. We only know the month of death. We had to assume the death date to be in the middle of the month to calculate the time to death and set the minimum time to death as 0. We performed sensitivity analysis by assuming that the death all occurred at the beginning or end of the month. The resulting death rate ranged from 1.4% to 2.3% in the non-cancer group and 4.2% to 7.3% in the cancer group. The results from regression models were very similar, except when we assumed the death all occurring at the end of the month, the smaller death rate led to non-significance of chemotherapy effect on death. After matching, the cancer effect on mortality has an HR of 1.05 (95% CI: 0.98-1.12) with a p-value of 0.138, or 1.01 (95% CI: 0.93-1.09) with a p-value of 0.850, under the two assumptions respectively. The rate of composite endpoint, on the other hand, almost had no change with different assumptions on the death date, and therefore, was very robust.In summary, this is one of the largest retrospective cohort study designed to investigate the various ways cancer history can impact COVID-19 outcomes. Important aspects of patient characteristics, clinical outcomes such as severity of illness and mortality were explored. Patient characteristics appeared to be the main determinants of mortality rather than the history of cancer itself. Cancer and COVID-19 are heterogeneous diseases; thus, we encourage prospective studies and more systematic efforts to understanding the effects of cancer subtypes and antitumor treatments in the context of studying COVID-19 outcomes.

**Contributors**

**GQZ, LZ conceived and designed the study. LZ, HB, CC designed and refined the statistical analysis plan. XJL, YH, YK, LZ, GQZ extracted and curated the data and developed the figures and tables. HZ, CC provided oncological expertise for interpretation of results. All authors contributed intellectual content during the drafting and revision of the work and approved the final version.**

**Declaration of interests**

**None**

**Data sharing**

**The dataset used for this study is provided by OPTUM, a third-party vendor. The University of Texas Health Science Center at Houston licensed this dataset.**

**Acknowledgments**

**This study was partly supported by grants from the National Cancer Institute (**R21CA231904 to GQZ**).**

**Table 1. Patient demographic and clinical characteristics**

| **Variable** | **No cancer**  **(n=514538)** | **Cancer**  **(n=31880)** |
| --- | --- | --- |
| Age, years |  |  |
| Mean (SD) |  |  |
| 18-50 | 274031 (53.3%) | 4096 (12.8%) |
| 50-65 | 139655 (27.1%) | 9947 (31.2%) |
| 65-75 | 55877 (10.9%) | 8387 (26.3%) |
| >=75 | 44975 (8.7%) | 9450 (29.6%) |
| Gender |  |  |
| Female | 288743 (56.1%) | 17363 (54.5%) |
| Male | 225289 (43.8%) | 14501 (45.5%) |
| Unknown | 506 (0.1%) | 16 (0.1%) |
| Race and ethnicity |  |  |
| Hispanic | 56331 (10.9%) | 2089 (6.6%) |
| Non-hispanic black | 55626 (10.8%) | 3530 (11.1%) |
| Non-hispanic white | 305755 (59.4%) | 23004 (72.2%) |
| Others/Unknown | 96826 (18.8%) | 3257 (10.2%) |
| Region of patient residence |  |  |
| Midwest | 238679 (46.4%) | 14137 (44.3%) |
| Northeast | 128132 (24.9%) | 9869 (31.0%) |
| South | 94738 (18.4%) | 5355 (16.8%) |
| West | 34601 (6.7%) | 1872 (5.9%) |
| Other/Unknown | 18388 (3.6%) | 647 (2.0%) |
| Number of established comorbidities |  |  |
| 0 | 187668 (36.5%) | 5061 (15.9%) |
| 1 | 166821 (32.4%) | 8134 (25.5%) |
| 2 | 90419 (17.6%) | 7152 (22.4%) |
| 3 | 36223 (7.0%) | 4770 (15.0%) |
| >=4 | 33407 (6.5%) | 6763 (21.2%) |
| Number of possible comorbidities |  |  |
| 0 | 120116 (23.3%) | 2224 (7.0%) |
| 1 | 187719 (36.5%) | 7395 (23.2%) |
| 2 | 141150 (27.4%) | 12370 (38.8%) |
| 3 | 51046 (9.9%) | 6916 (21.7%) |
| >=4 | 14507 (2.8%) | 2975 (9.3%) |
| Recent surgery |  |  |
| No | 443188 (86.1%) | 22801 (71.5%) |
| Yes | 71350 (13.9%) | 9079 (28.5%) |
| Type of malignancy |  |  |
| Solid tumor |  | 27195 (85.3%) |
| Hematologic malignancy |  | 4382 (13.7%) |
| Multiple cancers |  | 303 (1.0%) |
| Recent chemotherapy |  |  |
| No |  | 28884 (90.6%) |
| Yes |  | 2996 (9.4%) |
| Recent radiation therapy |  |  |
| No |  | 31483 (98.8%) |
| Yes |  | 397 (1.2%) |
| Age at cancer diagnosis, years |  |  |
| 0-20 |  | 149 (0.5%) |
| 20-40 |  | 2033 (6.4%) |
| 40-60 |  | 10103 (31.7%) |
| 60-80 |  | 15696 (49.2%) |
| >=80 |  | 3899 (12.2%) |
| Survival years >=5 |  |  |
| No |  | 25799 (80.9%) |
| Yes |  | 6081 (19.1%) |

**Table 2. Primary and secondary outcomes**

|  |  |  |
| --- | --- | --- |
| **Outcome** | **No cancer**  **(n=564642)** | **Cancer**  **(n=32080)** |
| Died | 10631 (2.1%) | 2052 (6.4%) |
| Admitted to a hospital | 85990 (16.7%) | 11498 (36.1%) |
| Admitted to an ICU | 15644 (3.0%) | 2099 (6.6%) |
| Required mechanical ventilation | 8049 (1.6%) | 1053 (3.3%) |
| Composite endpoint (yes/no) | 90511 (17.6%) | 12202 (38.3%) |
| Level of severity |  |  |
| None | 425030 (82.6%) | 19829 (62.2%) |
| Admitted to a hospital alone | 65039 (12.6%) | 8247 (25.9%) |
| Admitted to a hospital+ICU/ventilation | 13838 (2.7%) | 1752 (5.5%) |
| Died | 10631 (2.1%) | 2052 (6.4%) |

**Table 3. Primary and secondary outcomes by potential factors**

| **Variable** |  |  | **No cancer**  **(n=564642)** | | | | **Cancer**  **(n=32080)** | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | **Hospitalization**  **(n=88997)** | **Ventilation**  **(n=8098)** | **ICU**  **(n=15867)** | **Death**  **(n=10685)** | **Hospitalization**  **(n=11585)** | **Ventilation**  **(n=1060)** | **ICU**  **(n=2108)** | **Death**  **(n=2055)** |
| Age, years |  |  |  |  |  |  |  |  |  |  |
| 18-50 |  |  | 26312 (9.6%) | 1274 (0.5%) | 2987 (1.1%) | 507 (0.2%) | 951 (23.2%) | 46 (1.1%) | 119 (2.9%) | 43 (1.0%) |
| 50-65 |  |  | 23033 (16.5%) | 2645 (1.9%) | 4799 (3.4%) | 1648 (1.2%) | 2660 (26.7%) | 249 (2.5%) | 475 (4.8%) | 230 (2.3%) |
| 65-75 |  |  | 16252 (29.1%) | 2215 (4.0%) | 3851 (6.9%) | 2381 (4.3%) | 3184 (38.0%) | 356 (4.2%) | 642 (7.7%) | 478 (5.7%) |
| >=75 |  |  | 20393 (45.3%) | 1915 (4.3%) | 4007 (8.9%) | 6095 (13.6%) | 4703 (49.8%) | 402 (4.3%) | 863 (9.1%) | 1301 (13.8%) |
| Gender |  |  |  |  |  |  |  |  |  |  |
| Female |  |  | 46008 (15.9%) | 3092 (1.1%) | 6594 (2.3%) | 4636 (1.6%) | 5536 (31.9%) | 405 (2.3%) | 899 (5.2%) | 841 (4.8%) |
| Male |  |  | 39930 (17.7%) | 4951 (2.2%) | 9041 (4.0%) | 5988 (2.7%) | 5959 (41.1%) | 648 (4.5%) | 1199 (8.3%) | 1211 (8.4%) |
| Unknown |  |  | 52 (10.3%) | 6 (1.2%) | 9 (1.8%) | 7 (1.4%) | 3 (18.8%) | 0 (0.0%) | 1 (6.3%) | 0 (0.0%) |
| Race and ethnicity |  |  |  |  |  |  |  |  |  |  |
| Hispanic |  |  | 11706 (20.8%) | 1120 (2.0%) | 2200 (3.9%) | 814 (1.4%) | 853 (40.8%) | 81 (3.9%) | 173 (8.3%) | 97 (4.6%) |
| Non-hispanic black |  |  | 14402 (25.9%) | 1272 (2.3%) | 2998 (5.4%) | 1338 (2.4%) | 1732 (49.1%) | 177 (5.0%) | 402 (11.4%) | 263 (7.5%) |
| Non-hispanic white |  |  | 46908 (15.3%) | 4077 (1.3%) | 7992 (2.6%) | 6705 (2.2%) | 7822 (34.0%) | 665 (2.9%) | 1308 (5.7%) | 1491 (6.5%) |
| Others/Unknown |  |  | 12974 (13.4%) | 1580 (1.6%) | 2454 (2.5%) | 1774 (1.8%) | 1091 (33.5%) | 130 (4.0%) | 216 (6.6%) | 201 (6.2%) |
| # of established comorbidities |  |  |  |  |  |  |  |  |  |  |
| 0 |  |  | 10932 (5.8%) | 636 (0.3%) | 1351 (0.7%) | 1178 (0.6%) | 935 (18.5%) | 39 (0.8%) | 115 (2.3%) | 114 (2.3%) |
| 1 |  |  | 22609 (13.6%) | 1509 (0.9%) | 3193 (1.9%) | 1913 (1.1%) | 1980 (24.3%) | 116 (1.4%) | 262 (3.2%) | 300 (3.7%) |
| 2 |  |  | 20563 (22.7%) | 1723 (1.9%) | 3522 (3.9%) | 2170 (2.4%) | 2384 (33.3%) | 176 (2.5%) | 385 (5.4%) | 364 (5.1%) |
| 3 |  |  | 12902 (35.6%) | 1471 (4.1%) | 2756 (7.6%) | 1841 (5.1%) | 2079 (43.6%) | 221 (4.6%) | 394 (8.3%) | 376 (7.9%) |
| >=4 |  |  | 18984 (56.8%) | 2710 (8.1%) | 4822 (14.4%) | 3529 (10.6%) | 4120 (60.9%) | 501 (7.4%) | 943 (13.9%) | 898 (13.3%) |
| # of possible comorbidities |  |  |  |  |  |  |  |  |  |  |
| 0 |  |  | 5296 (4.4%) | 282 (0.2%) | 604 (0.5%) | 683 (0.6%) | 444 (20.0%) | 37 (1.7%) | 72 (3.2%) | 61 (2.7%) |
| 1 |  |  | 23259 (12.4%) | 1579 (0.8%) | 3385 (1.8%) | 2069 (1.1%) | 2054 (27.8%) | 144 (1.9%) | 331 (4.5%) | 348 (4.7%) |
| 2 |  |  | 33369 (23.6%) | 3601 (2.6%) | 6606 (4.7%) | 4343 (3.1%) | 4278 (34.6%) | 378 (3.1%) | 779 (6.3%) | 763 (6.2%) |
| 3 |  |  | 17252 (33.8%) | 1837 (3.6%) | 3533 (6.9%) | 2534 (5.0%) | 3027 (43.8%) | 310 (4.5%) | 565 (8.2%) | 581 (8.4%) |
| >=4 |  |  | 6814 (47.0%) | 750 (5.2%) | 1516 (10.5%) | 1002 (6.9%) | 1695 (57.0%) | 184 (6.2%) | 352 (11.8%) | 299 (10.1%) |
| Recent surgery |  |  |  |  |  |  |  |  |  |  |
| No |  |  | 56409 (12.7%) | 4086 (0.9%) | 9187 (2.1%) | 6421 (1.4%) | 6295 (27.6%) | 458 (2.0%) | 1100 (4.8%) | 1026 (4.5%) |
| Yes |  |  | 29581 (41.5%) | 3963 (5.6%) | 6457 (9.0%) | 4210 (5.9%) | 5203 (57.3%) | 595 (6.6%) | 999 (11.0%) | 1026 (11.3%) |
| Type of malignancy |  |  |  |  |  |  |  |  |  |  |
| Solid tumor |  |  |  |  |  |  | 9441 (34.7%) | 817 (3.0%) | 1709 (6.3%) | 1642 (6.0%) |
| Hematologic malignancy |  |  |  |  |  |  | 1944 (44.4%) | 221 (5.0%) | 363 (8.3%) | 388 (8.9%) |
| Multiple cancers |  |  |  |  |  |  | 113 (37.3%) | 15 (5.0%) | 27 (8.9%) | 22 (7.3%) |
| Recent chemotherapy |  |  |  |  |  |  |  |  |  |  |
| No |  |  |  |  |  |  | 9731 (33.7%) | 896 (3.1%) | 1778 (6.2%) | 1770 (6.1%) |
| Yes |  |  |  |  |  |  | 1767 (59.0%) | 157 (5.2%) | 321 (10.7%) | 282 (9.4%) |
| Recent radiation |  |  |  |  |  |  |  |  |  |  |
| No |  |  |  |  |  |  | 11257 (35.8%) | 1028 (3.3%) | 2058 (6.5%) | 1991 (6.3%) |
| Yes |  |  |  |  |  |  | 241 (60.7%) | 25 (6.3%) | 41 (10.3%) | 61 (15.4%) |
| ~~Age at cancer diagnosis, years~~ |  |  |  |  |  |  |  |  |  |  |
| ~~0-20~~ |  |  |  |  |  |  | 41 (27.5%) | 1 (0.7%) | 2 (1.3%) | 0 (0.0%) |
| ~~20-40~~ |  |  |  |  |  |  | 427 (21.0%) | 17 (0.8%) | 49 (2.4%) | 16 (0.8%) |
| ~~40-60~~ |  |  |  |  |  |  | 2464 (24.4%) | 213 (2.1%) | 428 (4.2%) | 180 (1.8%) |
| ~~60-80~~ |  |  |  |  |  |  | 6329 (40.3%) | 664 (4.2%) | 1278 (8.1%) | 1128 (7.2%) |
| ~~>=80~~ |  |  |  |  |  |  | 2237 (57.4%) | 158 (4.1%) | 342 (8.8%) | 728 (18.7%) |
| Survival years >=5 |  |  |  |  |  |  |  |  |  |  |
| No |  |  |  |  |  |  | 9712 (37.6%) | 880 (3.4%) | 1777 (6.9%) | 1730 (6.7%) |
| Yes |  |  |  |  |  |  | 1786 (29.4%) | 173 (2.8%) | 322 (5.3%) | 322 (5.3%) |

**Table 4. Multivariable regression models of potential factors associated with outcomes**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Death** | | **Severity (Yes/No)** | |
| **Variable** | **No Cancer** | **Cancer** | **No Cancer** | **Cancer** |
| Age, years |  |  |  |  |
| 50-65 | 5.14(4.62,5.72) | 1.80(1.29,2.51) | 1.21(1.19,1.24) | 0.94(0.86,1.03) |
| 65-75 | 16.45(14.83,18.24) | 3.79(2.75,5.21) | 2.12(2.07,2.17) | 1.35(1.22,1.48) |
| >=75 | 53.51(48.44,59.12) | 9.23(6.75,12.61) | 4.87(4.75,5.00) | 2.40(2.19,2.64) |
| 18-50 (ref) |  |  |  |  |
| Gender |  |  |  |  |
| Male | 1.54(1.48,1.60) | 1.30(1.19,1.42) | 1.09(1.07,1.11) | 1.22(1.16,1.28) |
| Female (ref) |  |  |  |  |
| Race and ethnicity |  |  |  |  |
| Hispanic | 1.25(1.16,1.35) | 0.98(0.79,1.20) | 2.08(2.02,2.13) | 1.52(1.38,1.69) |
| Non-hispanic black | 1.35(1.27,1.43) | 1.27(1.11,1.45) | 1.98(1.93,2.03) | 1.84(1.70,1.99) |
| Others/Unknown | 1.36(1.29,1.44) | 1.16(1.00,1.35) | 1.36(1.33,1.39) | 1.17(1.08,1.28) |
| Non-hispanic white (ref) |  |  |  |  |
| # of established comorbidities | 1.20(1.19,1.21) | 1.16(1.13,1.19) | 1.45(1.44,1.46) | 1.32(1.29,1.34) |
| # of possible comorbidities | 1.09(1.07,1.11) | 1.04(0.99,1.08) | 1.29(1.27,1.30) | 1.15(1.12,1.18) |
| Recent surgery |  |  |  |  |
| Yes | 1.65(1.58,1.72) | 1.84(1.68,2.01) | 2.36(2.32,2.41) | 2.53(2.40,2.68) |
| No (ref) |  |  |  |  |
| Type of malignancy |  |  |  |  |
| Hematologic malignancy |  | 1.47(1.32,1.65) |  | 1.58(1.47,1.69) |
| Multiple type |  | 1.17(0.76,1.80) |  | 1.07(0.83,1.40) |
| Solid tumor (ref) |  |  |  |  |
| Recent chemotherapy |  |  |  |  |
| Yes |  | 1.21(1.06,1.37) |  | 2.44(2.23,2.66) |
| No (ref) |  |  |  |  |
| Radiation therapy |  |  |  |  |
| Yes |  | 2.50(1.93,3.24) |  | 2.55(2.02,3.22) |
| No (ref) |  |  |  |  |
| Survival years >=5 |  |  |  |  |
| Yes |  | 0.75(0.66,0.84) |  | 0.64(0.60,0.68) |
| No (ref) |  |  |  |  |

**Table 5. Multivariable multinomial regression models of potential factors associated with levels of severity**

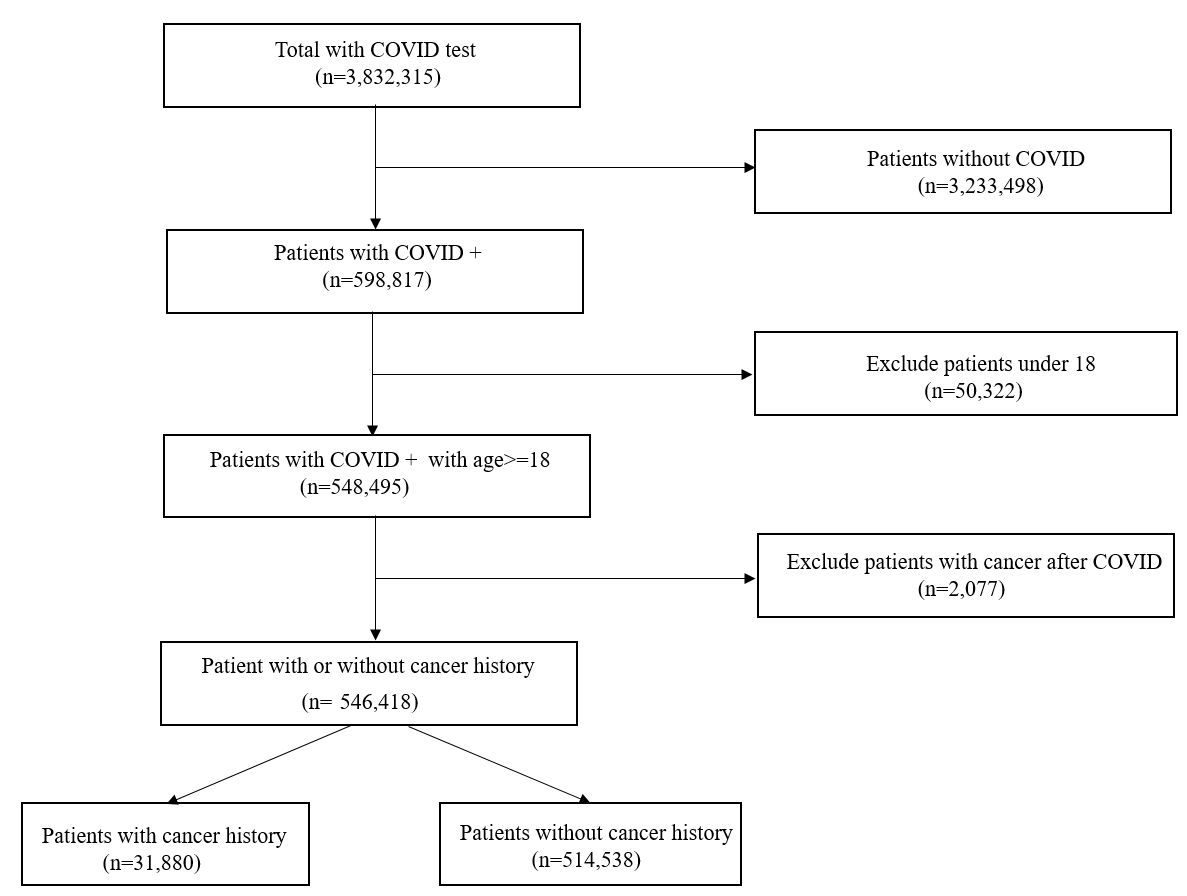
|  | **No Cancer** | | | **Cancer** | | |
| --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Hospitalization**  **Only vs**  **Not severe** | **Hospitalization +**  **ICU or ventilation vs**  **not severe** | **Death vs**  **not severe** | **Hospitalization**  **only vs**  **not severe** | **Hospitalization +**  **ICU or ventilation vs**  **not severe** | **Death vs**  **not severe** |
| Age, years |  |  |  |  |  |  |
| 50-65 | 1.06(1.03,1.08) | 1.76(1.67,1.84) | 4.42(3.99,4.89) | 0.86(0.78,0.95) | 1.23(0.99,1.52) | 1.66(1.19,2.32) |
| 65-80 | 1.68(1.64,1.73) | 2.82(2.67,2.98) | 15.32(13.87,16.92) | 1.17(1.06,1.29) | 1.54(1.24,1.92) | 3.92(2.85,5.41) |
| >=80 | 3.23(3.14,3.32) | 4.14(3.90,4.39) | 67.21(61.11,73.92) | 1.83(1.65,2.02) | 2.02(1.62,2.51) | 12.14(8.87,16.63) |
| 0-50 (ref) |  |  |  |  |  |  |
| Gender |  |  |  |  |  |  |
| Male | 0.96(0.94,0.98) | 1.63(1.58,1.69) | 1.62(1.56,1.69) | 1.14(1.08,1.21) | 1.45(1.31,1.61) | 1.44(1.31,1.59) |
| Female (ref) |  |  |  |  |  |  |
| Race and ethnicity |  |  |  |  |  |  |
| Hispanic | 2.01(1.96,2.07) | 2.66(2.52,2.81) | 1.77(1.63,1.91) | 1.52(1.36,1.70) | 2.00(1.65,2.42) | 1.24(0.99,1.55) |
| Non-hispanic black | 1.89(1.85,1.95) | 2.52(2.40,2.64) | 1.91(1.80,2.04) | 1.76(1.61,1.92) | 2.48(2.15,2.85) | 1.79(1.54,2.08) |
| Others/Unknown | 1.25(1.22,1.28) | 1.79(1.70,1.88) | 1.62(1.53,1.72) | 1.07(0.98,1.18) | 1.53(1.29,1.81) | 1.27(1.07,1.50) |
| Non-hispanic white (ref) |  |  |  |  |  |  |
| # of established comorbidities | 1.42(1.41,1.43) | 1.58(1.56,1.60) | 1.46(1.44,1.48) | 1.27(1.25,1.30) | 1.49(1.44,1.53) | 1.36(1.32,1.40) |
| # of possible comorbidities | 1.29(1.28,1.31) | 1.36(1.33,1.39) | 1.20(1.17,1.23) | 1.15(1.11,1.18) | 1.16(1.10,1.22) | 1.12(1.07,1.18) |
| Recent surgery |  |  |  |  |  |  |
| Yes | 2.29(2.24,2.34) | 2.76(2.65,2.86) | 2.50(2.39,2.61) | 2.45(2.31,2.60) | 2.57(2.32,2.86) | 3.02(2.73,3.34) |
| No (ref) |  |  |  |  |  |  |
| Type of malignancy |  |  |  |  |  |  |
| Hematologic malignancy |  |  |  | 1.51(1.40,1.64) | 1.63(1.42,1.87) | 1.89(1.66,2.15) |
| Multiple type |  |  |  | 0.97(0.73,1.31) | 1.41(0.89,2.24) | 1.18(0.73,1.90) |
| Solid tumor (ref) |  |  |  |  |  |  |
| Recent chemotherapy |  |  |  |  |  |  |
| Yes |  |  |  | 2.48(2.27,2.72) | 2.46(2.11,2.86) | 2.15(1.85,2.50) |
| No (ref) |  |  |  |  |  |  |
| Recent radiation therapy |  |  |  |  |  |  |
| Yes |  |  |  | 2.28(1.79,2.92) | 1.97(1.29,3.01) | 4.50(3.21,6.31) |
| No (ref) |  |  |  |  |  |  |
| Survival years >=5 |  |  |  |  |  |  |
| Yes |  |  |  | 0.66(0.61,0.71) | 0.62(0.54,0.72) | 0.59(0.52,0.68) |
| No (ref) |  |  |  |  |  |  |

**Table 6. Primary and secondary outcomes after matching**

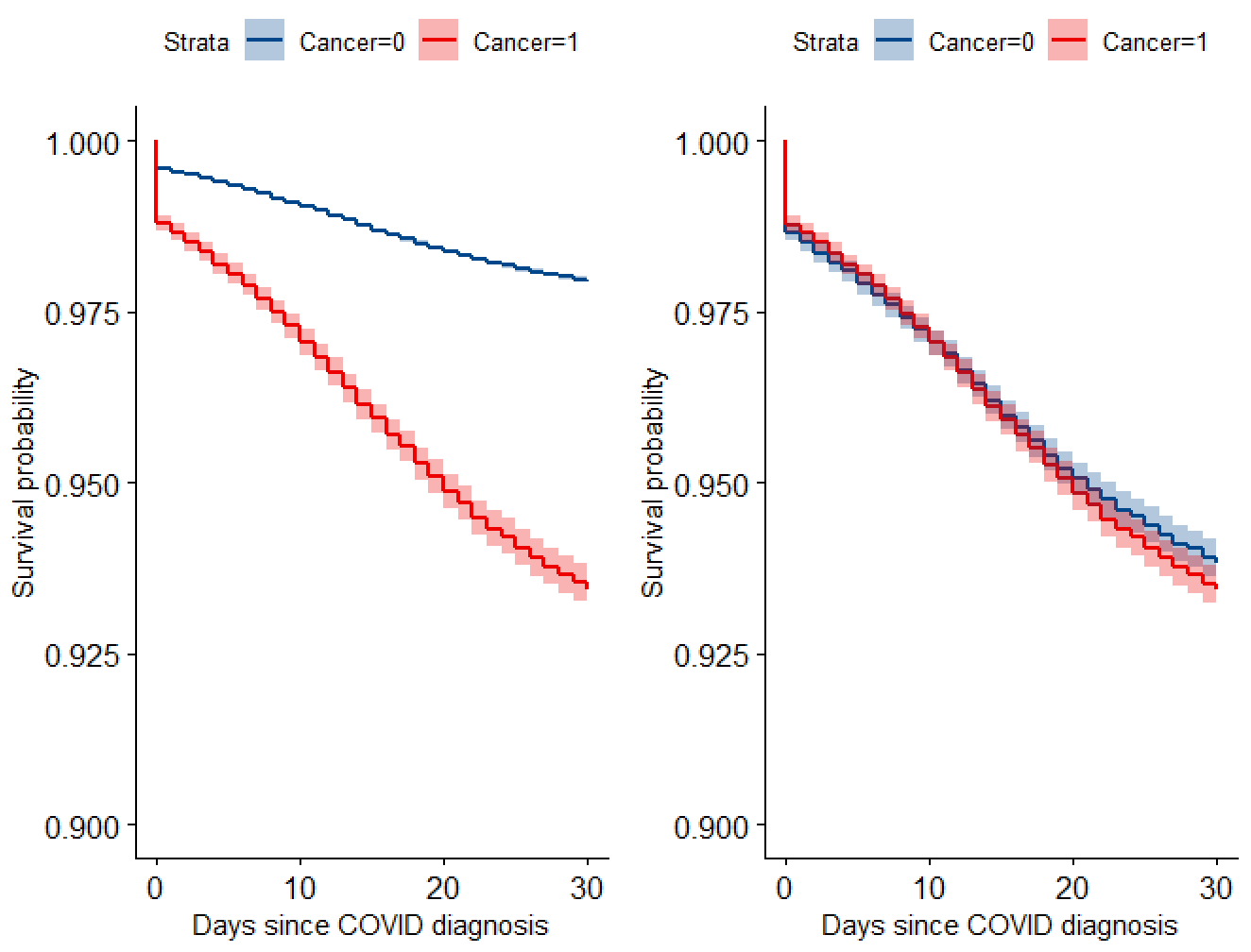
|  |  |  |
| --- | --- | --- |
| **Outcome** | **No cancer**  **(n= 31219)** | **Cancer**  **(n= 31219)** |
| Died | 1918 (6.1%) | 2015 (6.5%) |
| Admitted to a hospital | 9854 (31.6%) | 11244 (36.0%) |
| Admitted to an ICU | 2044 (6.5%) | 2057 (6.6%) |
| Required mechanical ventilation | 1156 (3.7%) | 1030 (3.3%) |
| Composite endpoint (yes/no) | 10566 (33.8%) | 11929 (38.2%) |
| Level of severity |  |  |
| None | 20812 (66.7%) | 19434 (62.3%) |
| Admitted to a hospital alone | 6818 (21.8%) | 8056 (25.8%) |
| Admitted to a hospital+ICU/ventilation | 1671 (5.4%) | 1714 (5.5%) |
| Died | 1918 (6.1%) | 2015 (6.5%) |

**Table 7. Multivariable regression models of cancer associated with outcomes adjusted for factors using matched data**.

|  | **Death** | | **Severity** | |
| --- | --- | --- | --- | --- |
| **Variable** | **Hazard ratio** | **P value** | **Odds ratio** | **P value** |
| Cancer |  |  |  |  |
| Yes | 1.03(0.97,1.10) | 0.324 | 1.21(1.17,1.26) | <0.0001 |
| No (ref) |  |  |  |  |
| Age, years |  |  |  |  |
| 50-65 | 2.29(1.69,3.08) | <0.0001 | 1.04(0.97,1.12) | 0.242 |
| 65-80 | 5.65(4.23,7.55) | <0.0001 | 1.62(1.51,1.74) | <0.0001 |
| >=80 | 14.80(11.13,19.70) | <0.0001 | 3.10(2.89,3.32) | <0.0001 |
| 0-50 (ref) |  |  |  |  |
| Gender |  |  |  |  |
| Male | 1.45(1.36,1.55) | <0.0001 | 1.30(1.25,1.35) | <0.0001 |
| Female (ref) |  |  |  |  |
| Race and ethnicity |  |  |  |  |
| Hispanic | 1.06(0.91,1.23) | 0.474 | 1.71(1.58,1.84) | <0.0001 |
| Non-Hispanic black | 1.27(1.16,1.41) | <0.0001 | 1.94(1.83,2.05) | <0.0001 |
| Non-Hispanic white (ref) | 1.24(1.11,1.38) | <0.0001 | 1.26(1.18,1.34) |  |
| # of established comorbidities | 1.17(1.15,1.19) | <0.0001 | 1.33(1.31,1.34) | <0.0001 |
| # of possible comorbidities | 1.05(1.01,1.08) | 0.007 | 1.16(1.14,1.19) | <0.0001 |
| Recent surgery |  |  |  |  |
| Yes | 1.71(1.60,1.83) | <0.0001 | 2.42(2.32,2.52) | <0.0001 |
| No (ref) |  |  |  |  |



**Figure 1: CONSORT diagram on data acquisition.**



**Figure 2. Survival plots for patients with and without cancer diagnosis. Left panel: before matching. Right panel: after matching.**

**Appendix**

**Table A1. Details of comorbidities**

| **Variable** | **No cancer**  **(n=514538)** | **Cancer**  **(n=31880)** |
| --- | --- | --- |
| **Established comorbidities** |  |  |
| Chronic kidney disease | 38891 (7.6%) | 7798 (24.5%) |
| Chronic Obstructive pulmonary disease | 29462 (5.7%) | 6061 (19.0%) |
| Down syndrome | 395 (0.1%) | 9 (0.0%) |
| Immunocompromised state from  Solid organ transplant | 177 (0.0%) | 83 (0.3%) |
| Obesity | 212527 (47.2%) | 13796 (44.0%) |
| Pregnancy | 14868 (2.9%) | 188 (0.6%) |
| Heart failure | 30768 (6.0%) | 5820 (18.3%) |
| Coronary artery disease | 49553 (9.6%) | 8973 (28.1%) |
| Cardiomyopathies | 11044 (2.1%) | 2171 (6.8%) |
| Sickle cell disease | 1308 (0.3%) | 116 (0.4%) |
| Smoking | 140015 (27.2%) | 14140 (44.4%) |
| Type 2 diabetes mellitus | 87630 (17.0%) | 10731 (33.7%) |
| **Possible comorbidities** |  |  |
| Asthma | 69564 (13.5%) | 5491 (17.2%) |
| Cerebrovascular | 7757 (1.5%) | 1527 (4.8%) |
| Cystic fibrosis | 14780 (2.9%) | 1408 (4.4%) |
| Hypertension | 183861 (35.7%) | 22279 (69.9%) |
| Immunodeficiency, unspecified | 1757 (0.3%) | 710 (2.2%) |
| Liver disease | 29029 (5.6%) | 5198 (16.3%) |
| Dementia | 15695 (3.1%) | 2485 (7.8%) |
| Overweight | 346027 (76.9%) | 23786 (75.9%) |
| Pulmonary fibrosis | 6565 (1.3%) | 1679 (5.3%) |
| Thalassemia | 1096 (0.2%) | 139 (0.4%) |
| Type 1 diabetes mellitus | 7927 (1.5%) | 858 (2.7%) |

**References**

1. American Cancer Society. Cancer Statistics Center. <http://cancerstatisticscenter.cancer.org>. March 24, 2021.

2. The Johns Hopkins Coronavirus Resource Center. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). <https://coronavirus.jhu.edu/map.html>. March 24, 2021.

3. Rugge M, Zorzi M, Guzzinati S. SARS-CoV-2 infection in the Italian Veneto region: adverse outcomes in patients with cancer. *Nature Cancer* 2020; **1**(8): 784-8.

4. Liang W, Guan W, Chen R, et al. Cancer patients in SARS-CoV-2 infection: a nationwide analysis in China. *The lancet oncology* 2020; **21**(3): 335-7.

5. Yu J, Ouyang W, Chua ML, Xie C. SARS-CoV-2 transmission in cancer patients of a tertiary hospital in Wuhan. *MedRxiv* 2020.

6. Dai M, Liu D, Liu M, et al. Patients with cancer appear more vulnerable to SARS-CoV-2: a multicenter study during the COVID-19 outbreak. *Cancer discovery* 2020; **10**(6): 783-91.

7. Zhang L, Zhu F, Xie L, et al. Clinical characteristics of COVID-19-infected cancer patients: a retrospective case study in three hospitals within Wuhan, China. *Annals of oncology* 2020; **31**(7): 894-901.

8. Kuderer NM, Choueiri TK, Shah DP, et al. Clinical impact of COVID-19 on patients with cancer (CCC19): a cohort study. *The Lancet* 2020; **395**(10241): 1907-18.

9. Albiges L, Foulon S, Bayle A, et al. Determinants of the outcomes of patients with cancer infected with SARS-CoV-2: results from the Gustave Roussy cohort. *Nature Cancer* 2020; **1**(10): 965-75.

10. Zhang H, Han H, He T, et al. Clinical Characteristics and Outcomes of COVID-19–Infected Cancer Patients: A Systematic Review and Meta-Analysis. *JNCI: Journal of the National Cancer Institute* 2020.

11. Garassino MC, Whisenant JG, Huang L-C, et al. COVID-19 in patients with thoracic malignancies (TERAVOLT): first results of an international, registry-based, cohort study. *The Lancet Oncology* 2020; **21**(7): 914-22.

12. Luo J, Rizvi H, Preeshagul IR, et al. COVID-19 in patients with lung cancer. *Annals of Oncology* 2020; **31**(10): 1386-96.

13. Yang L, Chai P, Yu J, Fan X. Effects of cancer on patients with COVID-19: a systematic review and meta-analysis of 63,019 participants. *Cancer Biology & Medicine* 2021; **18**(1): 298.

14. Wu C, Chen X, Cai Y, et al. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA internal medicine* 2020; **180**(7): 934-43.

15. Rivera DR, Peters S, Panagiotou OA, et al. Utilization of COVID-19 treatments and clinical outcomes among patients with cancer: a COVID-19 and Cancer Consortium (CCC19) cohort study. *Cancer discovery* 2020; **10**(10): 1514-27.

16. Derosa L, Melenotte C, Griscelli F, et al. The immuno-oncological challenge of COVID-19. *Nature Cancer* 2020; **1**(10): 946-64.

17. Jordan RE, Adab P, Cheng K. Covid-19: risk factors for severe disease and death. British Medical Journal Publishing Group; 2020.

18. Zheng Z, Peng F, Xu B, et al. Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis. *Journal of Infection* 2020.

19. Williamson EJ, Walker AJ, Bhaskaran K, et al. OpenSAFELY: factors associated with COVID-19 death in 17 million patients. *Nature* 2020.

20. Tian Y, Qiu X, Wang C, et al. Cancer associates with risk and severe events of COVID‐19: A systematic review and meta‐analysis. *International journal of cancer* 2021; **148**(2): 363-74.

21. Ofori-Asenso R, Ogundipe O, Agyeman AA, et al. Cancer is associated with severe disease in COVID-19 patients: a systematic review and meta-analysis. *ecancermedicalscience* 2020; **14**.

22. Meng Y, Lu W, Guo E, et al. Cancer history is an independent risk factor for mortality in hospitalized COVID-19 patients: a propensity score-matched analysis. *Journal of hematology & oncology* 2020; **13**(1): 1-11.

23. Joharatnam-Hogan N, Hochhauser D, Shiu K-K, et al. Outcomes of the 2019 novel coronavirus in patients with or without a history of cancer: a multi-centre North London experience. *Therapeutic advances in medical oncology* 2020; **12**: 1758835920956803.

24. Giannakoulis VG, Papoutsi E, Siempos II. Effect of cancer on clinical outcomes of patients with COVID-19: a meta-analysis of patient data. *JCO global oncology* 2020; **6**: 799-808.

25. Park R, Lee SA, Kim SY, de Melo AC, Kasi A. Association of active oncologic treatment and risk of death in cancer patients with COVID-19: a systematic review and meta-analysis of patient data. *Acta Oncologica* 2021; **60**(1): 13-9.

26. Liu H, Yang D, Chen X, et al. The effect of anticancer treatment on cancer patients with COVID‐19: A systematic review and meta‐analysis. *Cancer medicine* 2021; **10**(3): 1043-56.

27. Li P, Li L, Wang S, Liu Y, Li Z, Xia S. Effect of antitumor therapy on cancer patients infected by SARS‐CoV‐2: A systematic review and meta‐analysis. *Cancer Medicine* 2021.

28. He W, Chen L, Chen L, et al. COVID-19 in persons with haematological cancers. *Leukemia* 2020; **34**(6): 1637-45.

29. Miyashita H, Mikami T, Chopra N, et al. Do patients with cancer have a poorer prognosis of COVID-19? An experience in New York City. *Annals of Oncology* 2020.

30. Salunke AA, Nandy K, Pathak SK, et al. Impact of COVID-19 in cancer patients on severity of disease and fatal outcomes: a systematic review and meta-analysis. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews* 2020.

31. ElGohary GM, Hashmi S, Styczynski J, et al. The risk and prognosis of COVID-19 infection in cancer patients: a systematic review and meta-analysis. *Hematology/oncology and stem cell therapy* 2020.

32. Abrams EM, Szefler SJ. COVID-19 and the impact of social determinants of health. *The Lancet Respiratory Medicine* 2020; **8**(7): 659-61.

33. Laurencin CT, McClinton A. The COVID-19 pandemic: a call to action to identify and address racial and ethnic disparities. *Journal of racial and ethnic health disparities* 2020; **7**(3): 398-402.

34. Chowkwanyun M, Reed Jr AL. Racial health disparities and Covid-19—caution and context. *New England Journal of Medicine* 2020; **383**(3): 201-3.

35. Yancy CW. COVID-19 and african americans. *Jama* 2020; **323**(19): 1891-2.

36. Price-Haywood EG, Burton J, Fort D, Seoane L. Hospitalization and mortality among black patients and white patients with Covid-19. *New England Journal of Medicine* 2020; **382**(26): 2534-43.