**Research in Context Panel**

**Evidence before this study:** The negative impact of COVID-19 on cancer outcomes has been reported in the literature, particularly highlighting the impact on patients needing active treatments but facing delays. Studies have also reported that cancer diagnosis is associated with increased incidence of SARS-CoV-2 infection and worse COVID-19 outcomes, but the majority of these publications have been case series or cohorts from single institutions. The impact of cancer diagnosis as an *independent contributor* to worse COVID-19 outcomes has not been studied in large cohorts.

**Added value of this study:** This study reports the results of a large-scale epidemiological study of 546,418 patients diagnosed with COVID-19: 31,880 individuals with cancer and 514,538 patients without cancer. While this report also confirmed prior reported findings that individuals with cancer history had higher 30-day mortality and poorer outcomes after SARS-CoV-2 infection compared to those without, it was noted that individuals with cancer and COVID-19 were generally older, more likely to be White, and had more CDC recognized comorbidities associated with worse COVID-19 outcomes. Following a matched-cohort analysis that accounted for age and these comorbidities, the independent impact of cancer diagnosis on mortality from COVID-19 was no longer significant, and the effect of cancer on COVID-19 severity was significant but with a smaller effect.

**Implications of all the available evidence:** These findings highlight the need to carefully untangle the potential shared risk factors between cancer and COVID-19 on patient outcomes to ensure the associated risk factors are considered in context. Particularly as future care management decisions are made amidst a world now dramatically impacted by COVID-19, these new and critical findings will alert the need that other comorbid risk factors are considered in the treatment and care of cancer patients. As we pursue research on the post-acute sequelae of COVID-19, this study reminds the community that contextual and comprehensive information of patients and populations is critical for making meaningful inferences and deciphering independent risk factors.

**Does Cancer History Drive COVID-19 Outcome? A Large-scale Matched Cohort Analysis 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** Few studies have addressed the question of whether cancer is as an independent risk factor for poorer COVID-19 outcomes. Here we characterize and compare COVID-19 outcomes between individuals with and without cancer diagnosis.

**Methods** We use de-identified Optum® COVID-19 Electronic Health Records (EHR) of 3.8 million unique individuals. The primary endpoint was all-cause mortality within 30 days of COVID-19 diagnosis. The secondary endpoint was COVID-19 severity, defined using indications of hospital admission, intensive care unit (ICU) admission, and mechanical ventilation. Individuals with cancer were matched to those without, on demographic and risk factors. The impact of cancer on COVID-19 outcomes was then analyzed using logistic regression.    
**Findings** Of the 546,418 COVID-19 positive individuals, 31,880 had cancer and 514,538 did not. Individuals with cancer diagnosis were older, more likely to be White, and had more comorbidities. They also had poorer outcomes: 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%). After matching however, difference in outcome significantly decreased: 30-day mortality rate (6.5% vs 6.2%); admission to hospital (36.0% vs 31.3%); admission to ICU (6.6% vs 6.5%); and need for mechanical ventilation (3.3% vs 3.6%). The effect of cancer on mortality is no longer significant (odds ratio [OR]: 1.04, 95% 0.97-1.11, p=0.260).

**Interpretation** Individuals with cancer had higher 30-day mortality and poorer outcomes after COVID-19 infection compared to those without. However, the differences were mainly due to older age and comorbidities associated with cancer, rather than cancer history itself.

**INTRODUCTION**

An estimated 17 million individuals with or survived cancer live in the U. S.1 and millions more around the world. With more than 150 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 impact on cancer patients3-6, particularly those who were undergoing chemotherapy4,7. 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, without consideration for matched controls. 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 COVID-19 and its associated negative outcomes16. The extent at which cancer diagnosis contributes to poorer COVID-19 outcomes is a distinct topic from how COVID-19 impacts those who are undergoing treatments. Before attributing poorer COVID-19 outcome to an individual’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 particular patient population. Because of the inherent heterogeneity of this population, a sufficiently large cancer patient population is required to match age, sex, and other known COVID-19 risk factors with adequate control for potential confounders. In this study, we use the Optum®de-identified COVID-19 Electronic Health Record (EHR) data to elucidate COVID-19 outcomes between cancer and non-cancer subgroups.

**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, regardless of the test results. The Optum® COVID-19 EHR data cover patient-level, longitudinal clinical records including demographics, diagnoses, procedures, lab tests, care settings, medications prescribed or administered, and mortality.

The inclusion criteria for our study are: 1) Subject is aged 18 or older; 2) Subject has a positive result of polymerase chain reaction (PCR) test, antibody test, or antigen test for COVID-19, or has 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). Individuals are assigned to cancer groups according to any record of ICD-9 or 10 codes indicating cancer before positive COVID-19 diagnosis. To avoid misclassification, individuals are excluded if the first record of diagnosis for cancer appeared on and after the first date of COVID-19 confirmation. This study was determined by the Committee for the Protection of Human Subjects at The University of Texas Health Science Center at Houston (UTHealth) as non-human-subjects research.

**Cancer and Subtypes**

Both ICD-9 and ICD-10 codes are used to identify cancer diagnosis. ICD-10 codes include: C50 (breast cancer), C61 (prostate cancer), C18 (colon and rectal cancer), and C34 (lung cancer). Additional variables related to cancer diagnosis are considered. These include cancer subtypes (solid tumor, hematologic malignancy, multiple cancers); recent systemic therapy; recent radiation therapy; age at first cancer diagnosis; and years of survival. Recent systemic therapy and recent radiation therapy are defined as such therapies being performed within 4 weeks before COVID-19 diagnosis.

**Outcomes**

The primary outcome is all-cause mortality (death) within 30 days of the initial date of COVID-19 infection. The secondary outcome is severe illness, defined as one or a combination of: hospital admission, intensive care unit (ICU) admission, mechanical ventilation, and death. Severity of illness is treated as both a dichotomous variable (none versus one of any of hospitalization, ICU, ventilation, or death) and a multinomial variable (0: none; 1: hospitalization only; 2: hospitalization + ICU/ventilation without death; 3: mortality).

**Risk Factors**

We use potential risk factors for severe COVID-19 illness according to those provided by The Center for Disease Control (CDC). They include age, gender, race, ethnicity, and established or possible comorbidities.  Established comorbidities 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. Possible comorbidities 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.

To characterize and compare COVID-19 outcomes between individuals with and without cancer diagnosis, we consider cancer type, recent chemotherapy, recent radiation therapy, age at cancer, and years of survival as exposure variables. Recent surgery, recent chemotherapy, and recent radiation therapy were defined as within 4 weeks before the COVID-19 diagnosis. To avoid confounding, we exclude cancer diagnosis from the list of comorbidities in our analysis.

**Matching**

Cancer and non-cancer subgroups are matched (1-1) based on age, gender, race and ethnicity, region, and the numbers of established and possible comorbidities. Nearest neighbor method is used to pick matching individuals based on propensity scores.

**Statistical Analysis**

Descriptive statistics for demographic and cancer characteristics are used to illustrate the effects of matching (before and after). Logistic regression models are used for analyzing the association between cancer and COVID-19 outcomes such as mortality and severity in the matched subgroup. Multinomial regression analysis is used to account for levels of severity. Subgroup analyses are performed for those with cancer to examine the association between cancer-related treatments and COVID-19 outcomes. A p-value less than 0.05 is considered statistically significant. Propensity score matching is performed using *matchit,* a part of R package. The rest of the analyses are performed in SAS 9.4.

Data imputation

Listwise deletion was used.

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 30-days mortality is a time-to-event outcome and patients without confirmed death within 30 days will be noted as censored. The

The average effect in COVID outcome of cancer will be estimated by propensity score matching (PSM) analysis and the details of matching will be shown in the section below.

Matching was performed using logistic regression,

After matching, cancer group and control group. Methods that were used for propensity-score matched data. Cox proportional hazard model (Cox-ph model) with cluster-robust standard errors with 30-days mortality as outcome and cancer as the only predictor was used to estimate the cancer effect in mortality. Conditional logistic regression was used for the

For the secondary study objective (disintegrate the effect of cancer), a sub-group analysis will be conducted in the cancer group before matching to explore the association between cancer-related treatment and COVID-19 outcome. Cox proportional hazard model was used to estimate the effect of cancer-related treatment in 30-days mortality after adjusting potential confounders. Logistic regression was used for the.

Propensity score matching

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.

Logistic regression was used to estimate the probability of being in cancer group (propensity score) based on baseline characteristics.

**RESULTS**

Of the 3,832,315 individuals in the Optum® COVID-19 EHR 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-19 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; Figure 1).

Demographic and clinical characteristics for the two groups before and after matching are summarized in Table 1. Before matching, individuals with cancer diagnosis were older than those without, with a median age of 67 (IQR 57-77) versus 48 (IQR 33-61). 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%), although 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, 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, the distributions of each variable used for matching were well balanced between patients with and without a cancer history (Table 1). Although we did not match at individual comorbidity level, the distributions of comorbidities became more similar after matching (Table A.1).

**Mortality.** Before matching, the crude 30-day mortality rates in individuals with and without cancer diagnosis were 6.4% (95% CI: 6.17% - 6.71%) and 2.07% (95% CI: 2.03% - 2.10%), respectively (Table 2). After matching, the 30-day mortality rates in individuals with and without cancer diagnosis were 6.5% (95% CI: 6.18% - 6.73%) and 6.2% (95% CI: 5.93% - 6.47%), respectively (Table 2).

Logistic regression analysis using matched controls shows that older age, male, non-Hispanic Black, higher number of established and possible comorbidities are associated with higher risk of mortality (Table 3). However, the hazard ratio of mortality for individuals with cancer diagnosis versus those without was 1.04 (95% CI: 0.97-1.11) with a p-value of 0.260 after adjustment. This indicates a non-significant effect of cancer history on COVID-19 mortality.

**Severity.** With respect to other types of severe outcomes, the rate of admission to hospital before matching (Table 4a) 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%).

After matching (Table 4b), the rates of admission to a hospital were 36.0% (95% CI: 35.48% - 36.55%) vs 31.3% (95% CI: 30.82% - 31.85%) in individuals with and without cancer diagnosis. The rates of admission to ICU were 6.6% (95% CI: 6.31% - 6.86%) vs 6.5% (95% CI: 6.23% - 6.77%). The rates of mechanical ventilation were 3.3% (95% CI: 3.10% - 3.50%) vs 3.6% (95% CI: 3.39% - 3.81%). Composite endpoint rates were 38.2% (95% CI: 37.67% - 38.75%) vs 33.6% (95% CI: 33.08% - 34.13%). Older age, male, non-Hispanic Black, higher number of established and possible comorbidities were consistently associated with higher risk of hospitalization, ventilation, and ICU (Table 4b), as confirmed by logistic regression (Table 3). After adjusting for these variables, the odds ratio for cancer versus non-cancer was 1.25 (95% CI: 1.21 - 1.30), with p-value <0.0001 (Table 3).

**Characteristics of cancer.** Among individuals with cancer diagnosis, 85.3% had solid tumor, 13.7% had hematologic malignancy, and 1% had multiple cancers (Table A.2). The most common cancer sites are (Figure A.1): breast (20.9%), prostate (16.2%), colon and rectal (6.9%), lymphoma (6.4%), lung (6.3%), leukemia (4.9%), thyroid (4.7%), skin (melanoma) (4.5%), kidney (4.0%), and bladder (3.9%). The majority of initial cancer diagnosed happened after 60 years of age (61.4%). 19.1% survived cancer for more than 5 years. 9.4% received systemic therapy and 1.2% received the therapy within 4 weeks before COVID-19 diagnosis.

**Outcome by cancer types.** Hematologic malignancy, recent systemic therapy, recent radiation therapy and older age at cancer diagnosis were associated with higher rates for worse outcomes (Table 5). Survival of 5 or more years seemed to be associated with lower rates for these outcomes. Logistic regression analysis confirmed that these risk factors were significantly associated with higher odds ratio of mortality and severe outcomes (Table 6).

**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 analyses 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, using matched controls, our analysis demonstrated that the differences were mainly due to older age and having more comorbidities in individuals with a cancer diagnosis, instead of cancer itself.

The 31,880 individuals with cancer history 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 including 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, and number of established comorbidities (smoking is one of them) 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. Our findings confirmed similar risk factors in cancer patients as predictors of severe COVID-19 outcome 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 itself. 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, but the reason is due to older age and having more comorbidities.

Among patients with cancer diagnosis, we found hematologic malignancies and recent systemic therapy 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. As the interpretation of “recent” ranged from 2 weeks to 3 months in published literature, we defined “recent” as within the past 4 weeks of COVID-19 diagnosis.

Literature shows that lung cancer is, unsurprisingly, associated with increased severity and mortality.11,12 For individuals with a diagnosis of lung cancer, we found, before matching, that both the mortality and severity rates are higher, i.e., 12.5% (95% CI: 11.10% -13.99%) for mortality, 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 (OR=1.64, 95% CI: 1.33 – 2.03, p<0.0001) and severe outcomes (OR=1.83, 95% CI: 1.60 – 2.09, p<0.0001).

We found that the rates of admission to hospital, admission to ICU, and mechanical ventilation to be higher in patients with cancer diagnosis. Further categorization of outcomes into hospitalization only, hospitalization with ICU and mechanical ventilation without death, and death, significant risk factors did not change the results of the analysis. After matching on related risk factors, the differences became less prominent. Overall, the cancer group had slightly higher rates than non-cancer group (OR: 1.25, 95% 1.21-1.30, p<0.0001). We reaffirmed previous studies7,24,29-30 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. 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.

To make sure that our analytical results are not sensitive to different methods for COVID-19 diagnosis, we performed parallel analyses using PCR test for the identification of COVID-19. PCR test identified 414,889 COVID-19 positive individuals, accounting for 69% of previously identified individuals (n=598,817). Although the total number of the base cohort is 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 odds ratios from logistic regression analyses were also similar. In particular, after matching on age, gender, race and ethnicity, region, numbers of established and possible comorbidities, the effect of cancer were not significant for mortality (OR=1.06, 95% CI: 0.98 – 1.14, p=0.174), but significant for severity (OR=1.24, 95% CI: 1.19 – 1.30, p<0.0001), similar to our previous conclusion based on a more inclusive criterion for COVID-19 positivity.

In summary, this large-scale epidemiological study of 31,880 individuals with cancer diagnosis and COVID-19 infection used, to our knowledge, the largest COVID-19 related population with comprehensive Electronic Health Record information dating back to 2007. Our results revealed that while individuals with cancer history had higher 30-day mortality and poorer outcomes after COVID-19 infection compared to those without cancer, as previously reported, the differences in outcome were mainly due to older age and non-cancer specific comorbidities within the cohort of patients with cancer rather than the cancer diagnosis itself. After matching for these risk factors for worse COVID-19 outcome, the effect of cancer on 30-day mortality from COVID-19 infection was no longer significant, and the effect of cancer on COVID-19 severity (other than mortality) although statistically significant had a small relative effect. These findings add new, critical knowledge and highlights the need to untangle the potentially shared risk factors between cancer and COVID-19 outcomes in order to guide future clinical management and research.

**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 and provided further content. 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 de-identified EHR data 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**).**

**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.**

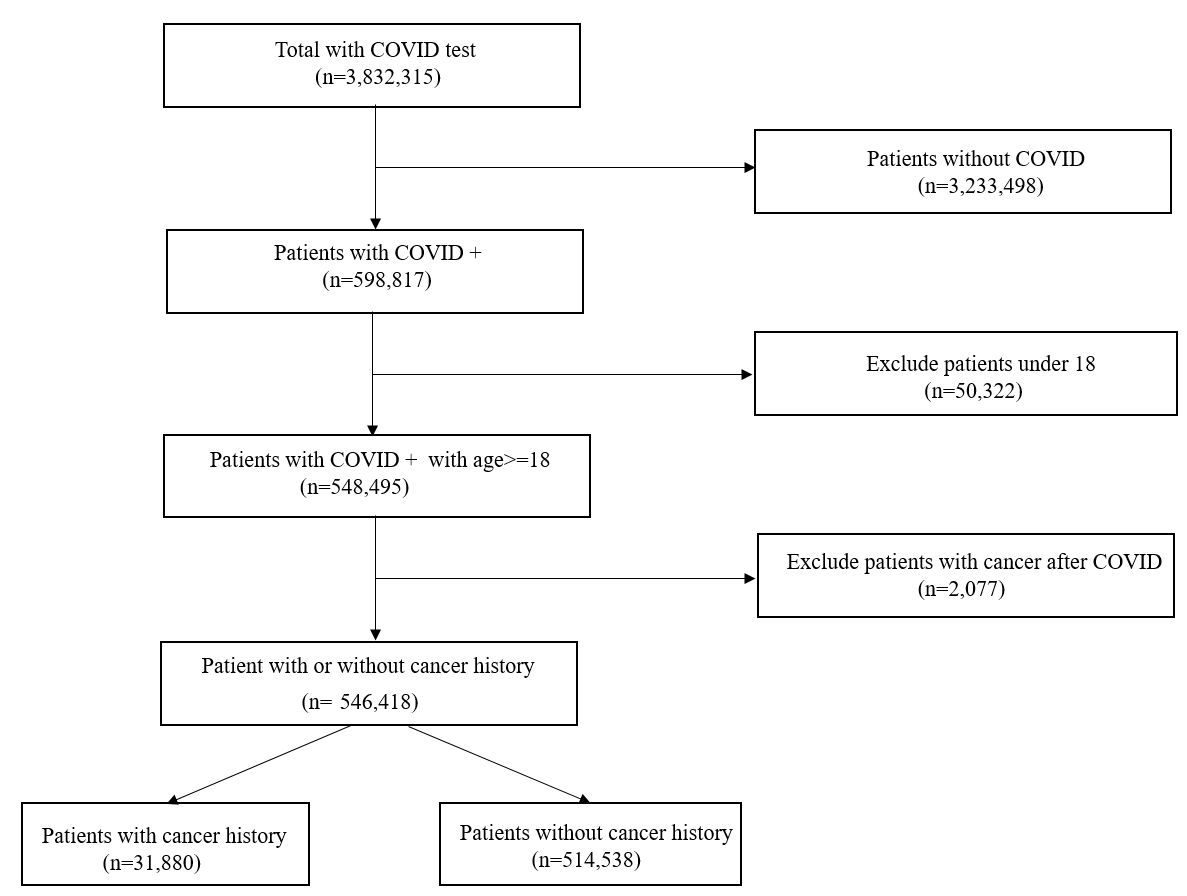
**27. Li P, Li L, Wang S, Liu Y, Li Z, Xia S. Effect of antitumor therapy on cancer patients infected by SARS‐CoV‐2: A systematic review and meta‐analysis. Cancer Medicine 2021.**

**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.**

**Figure 1: Study subgroup flowchart.**



**Table 1. Demographic and clinical characteristics: before and after matching.**

|  | **Before matching** | | **After matching** | |
| --- | --- | --- | --- | --- |
|  | **No cancer**  **(n=514538)** | **Cancer**  **(n=31880)** | **No cancer**  **(n=31219)** | **Cancer**  **(n=31219)** |
| **Age** | | | | |
| Median (IQR) | 48 (33 – 61) | 67 (55 – 77) | 67 (57 – 77) | 67 (57 – 77) |
| 18-50 | 274031 (53.3%) | 4096 (12.8%) | 3927 (12.6%) | 3958 (12.7%) |
| 50-65 | 139655 (27.1%) | 9947 (31.2%) | 9743 (31.2%) | 9752 (31.2%) |
| 65-75 | 55877 (10.9%) | 8387 (26.3%) | 8146 (26.1%) | 8249 (26.4%) |
| ≥75 | 44975 (8.7%) | 9450 (29.6%) | 9403 (30.1%) | 9260 (29.7%) |
| **Gender** | | | | |
| Female | 288743 (56.1%) | 17363 (54.5%) | 17051 (54.6%) | 16994 (54.4%) |
| Male | 225289 (43.8%) | 14501 (45.5%) | 14168 (45.4%) | 14225 (45.6%) |
| Unknown | 506 (0.1%) | 16 (0.1%) |  |  |
| **Race and ethnicity** | | | | |
| Hispanic | 56331 (10.9%) | 2089 (6.6%) | 1939 (6.2%) | 2047 (6.6%) |
| Non-Hispanic Black | 55626 (10.8%) | 3530 (11.1%) | 3327 (10.7%) | 3447 (11.0%) |
| Non-Hispanic White | 305755 (59.4%) | 23004 (72.2%) | 22851 (73.2%) | 22538 (72.2%) |
| Others/Unknown | 96826 (18.8%) | 3257 (10.2%) | 3102 (9.9%) | 3187 (10.2%) |
| **Region of residence** | | | | |
| Midwest | 238679 (46.4%) | 14137 (44.3%) | 14140 (45.3%) | 14129 (45.3%) |
| Northeast | 128132 (24.9%) | 9869 (31.0%) | 9860 (31.6%) | 9867 (31.6%) |
| South | 94738 (18.4%) | 5355 (16.8%) | 5425 (17.4%) | 5352 (17.1%) |
| West | 34601 (6.7%) | 1872 (5.9%) | 1794 (5.7%) | 1871 (6.0%) |
| Other/Unknown | 18388 (3.6%) | 647 (2.0%) |  |  |
| **Number of established comorbidities** | | | | |
| 0 | 187675 (36.5%) | 5061 (15.9%) | 5033 (16.1%) | 4946 (15.8%) |
| 1 | 166816 (32.4%) | 8134 (25.5%) | 8066 (25.8%) | 7994 (25.6%) |
| 2 | 90417 (17.6%) | 7152 (22.4%) | 7042 (22.6%) | 7008 (22.4%) |
| 3 | 36223 (7.0%) | 4770 (15.0%) | 4640 (14.9%) | 4672 (15.0%) |
| ≥4 | 33407 (6.5%) | 6763 (21.2%) | 6438 (20.6%) | 6599 (21.1%) |
| **Number of possible comorbidities** | | | | |
| 0 | 120124 (23.3%) | 2224 (7.0%) | 2158 (6.9%) | 2182 (7.0%) |
| 1 | 187712 (36.5%) | 7395 (23.2%) | 7271 (23.3%) | 7242 (23.2%) |
| 2 | 141149 (27.4%) | 12370 (38.8%) | 12330 (39.5%) | 12138 (38.9%) |
| 3 | 51046 (9.9%) | 6916 (21.7%) | 6878 (22.0%) | 6768 (21.7%) |
| ≥4 | 14507 (2.8%) | 2975 (9.3%) | 2582 (8.3%) | 2889 (9.3%) |

**Table 2. Primary and secondary outcomes.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Before matching** | | **After matching** | |
| **Outcome** | **No cancer**  **(n=514538)** | **Cancer**  **(n=31880)** | **No cancer**  **(n= 31219)** | **Cancer**  **(n= 31219)** |
| Mortality | 10631 (2.1%) | 2052 (6.4%) | 1935 (6.2%) | 2015 (6.5%) |
| Admitted to a hospital | 85990 (16.7%) | 11498 (36.1%) | 9782 (31.3%) | 11244 (36.0%) |
| Admitted to an ICU | 15644 (3.0%) | 2099 (6.6%) | 2029 (6.5%) | 2057 (6.6%) |
| Required mechanical ventilation | 8049 (1.6%) | 1053 (3.3%) | 1124 (3.6%) | 1030 (3.3%) |
| Composite endpoint (yes/no) | 90511 (17.6%) | 12202 (38.3%) | 10492 (33.6%) | 11929 (38.2%) |
| Level of severity |  |  |  |  |
| None | 425030 (82.6%) | 19829 (62.2%) | 20877 (66.9%) | 19434 (62.3%) |
| Admitted to a hospital alone | 65039 (12.6%) | 8247 (25.9%) | 6750 (21.6%) | 8056 (25.8%) |
| Admitted to hospital + ICU/ventilation | 13838 (2.7%) | 1752 (5.5%) | 1657 (5.3%) | 1714 (5.5%) |
| Mortality | 10631 (2.1%) | 2052 (6.4%) | 1935 (6.2%) | 2015 (6.5%) |

**Table 3. COVID-19 outcome matched control analysis: mortality and severity for individuals with cancer history.**

|  | **Mortality** | | **Severity** | |
| --- | --- | --- | --- | --- |
|  | **Odds ratio** | **p-value** | **Odds ratio** | **p-value** |
| **Age** | | | | |
| 50-65 | 1.88 (1.35,2.61) | <0.0001 | 1.00 (0.92,1.10) | 0.971 |
| 65-75 | 4.09 (2.98,5.63) | <0.0001 | 1.46 (1.33,1.60) | <0.0001 |
| ≥75 | 10.38 (7.59,14.18) | <0.0001 | 2.56 (2.34,2.81) | <0.0001 |
| 18-50 (ref) |  |  |  |  |
| **Gender** | | | | |
| Male | 1.32 (1.20,1.45) | <0.0001 | 1.20 (1.14,1.27) | <0.0001 |
| Female (ref) |  |  |  |  |
| **Race and ethnicity** | | | | |
| Hispanic | 1.05 (0.84,1.30) | 0.691 | 1.65 (1.49,1.82) | <0.0001 |
| Non-Hispanic Black | 1.32 (1.14,1.52) | <0.0001 | 1.89 (1.74,2.04) | <0.0001 |
| Others/Unknown | 1.24 (1.06,1.45) | 0.008 | 1.25 (1.15,1.36) | <0.0001 |
| Non-Hispanic White (ref) |  |  |  |  |
| Established comorbidities | 1.21 (1.18,1.25) | <0.0001 | 1.37 (1.35,1.39) | <0.0001 |
| Possible comorbidities | 1.06 (1.02,1.11) | 0.008 | 1.17 (1.14,1.20) | <0.0001 |
| **Type of malignancy** | | | | |
| Hematologic malignancy | 1.52 (1.35,1.71) | <0.0001 | 1.56 (1.45,1.67) | <0.0001 |
| Multiple cancers | 1.20 (0.76,1.89) | 0.433 | 1.10 (0.85,1.41) | 0.478 |
| Solid tumor (ref) |  |  |  |  |
| **Recent systemic therapy** | | | | |
| Yes | 1.35 (1.17,1.55) | <0.0001 | 2.67 (2.45,2.90) | <0.0001 |
| No (ref) |  |  |  |  |
| **Recent radiation therapy** | | | | |
| Yes | 3.09 (2.29,4.17) | <0.0001 | 2.74 (2.19,3.43) | <0.0001 |
| No (ref) |  |  |  |  |
| **Survival years ≥ 5** | | | | |
| Yes | 0.71 (0.63,0.81) | <0.0001 | 0.63 (0.59,0.67) | <0.0001 |
| No (ref) |  |  |  |  |

**Table 4. Outcome by risk factors: 4a. before matching; 4b. after matching.**

| **4a. Before Matching** | **No cancer (n=514538)** | | | | **Cancer (n=31880)** | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Hospitalization**  **(n=85990)**  **(%: rate)** | **Ventilation**  **(n=8049)**  **(%: rate)** | **ICU**  **(n=15644)**  **(%: rate)** | **Mortality**  **(n=10631)**  **(%: rate)** | **Hospitalization**  **(n=11498)**  **(%: rate)** | **Ventilation**  **(n=1053)**  **(%: rate)** | **ICU**  **(n=2099)**  **(%: rate)** | **Mortality**  **(n=2052)**  **(%: rate)** |
| **Age** | | | | | | | | |
| 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%) |
| **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%) |
| **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%) |

| **4b. After Matching** | **No cancer (n=31219)** | | | | **Cancer (n=31219)** | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Hospitalization**  **(n=9782)**  **(%: rate)** | **Ventilation**  **(n=1124)**  **(%: rate)** | **ICU**  **(n=2029)**  **(%: rate)** | **Mortality**  **(n=1935)**  **(%: rate)** | **Hospitalization**  **(n=11244)**  **(%: rate)** | **Ventilation**  **(n=1030)**  **(%: rate)** | **ICU**  **(n=2057)**  **(%: rate)** | **Mortality**  **(n=2015)**  **(%: rate)** |
| **Age** | | | | | | | | |
| 18-50 | 509 (13.0%) | 25 (0.6%) | 56 (1.4%) | 9 (0.2%) | 922 (23.3%) | 45 (1.1%) | 117 (3.0%) | 40 (1.0%) |
| 50-65 | 1983 (20.4%) | 236 (2.4%) | 429 (4.4%) | 132 (1.4%) | 2578 (26.4%) | 238 (2.4%) | 456 (4.7%) | 225 (2.3%) |
| 65-75 | 2696 (33.1%) | 412 (5.1%) | 659 (8.1%) | 422 (5.2%) | 3128 (37.9%) | 353 (4.3%) | 634 (7.7%) | 472 (5.7%) |
| ≥75 | 4594 (48.9%) | 451 (4.8%) | 885 (9.4%) | 1372 (14.6%) | 4616 (49.8%) | 394 (4.3%) | 850 (9.2%) | 1278 (13.8%) |
| **Gender** | | | | | | | | |
| Female | 4664 (27.4%) | 456 (2.7%) | 847 (5.0%) | 750 (4.4%) | 5410 (31.8%) | 393 (2.3%) | 879 (5.2%) | 826 (4.9%) |
| Male | 5118 (36.1%) | 668 (4.7%) | 1182 (8.3%) | 1185 (8.4%) | 5834 (41.0%) | 637 (4.5%) | 1178 (8.3%) | 1189 (8.4%) |
| **Race and ethnicity** | | | | | | | | |
| Hispanic | 674 (34.8%) | 82 (4.2%) | 153 (7.9%) | 88 (4.5%) | 829 (40.5%) | 81 (4.0%) | 173 (8.5%) | 97 (4.7%) |
| Non-Hispanic Black | 1485 (44.6%) | 176 (5.3%) | 380 (11.4%) | 225 (6.8%) | 1693 (49.1%) | 167 (4.8%) | 387 (11.2%) | 256 (7.4%) |
| Non-Hispanic White | 6794 (29.7%) | 720 (3.2%) | 1314 (5.8%) | 1452 (6.4%) | 7665 (34.0%) | 656 (2.9%) | 1288 (5.7%) | 1466 (6.5%) |
| Others/Unknown | 829 (26.7%) | 146 (4.7%) | 182 (5.9%) | 170 (5.5%) | 1057 (33.2%) | 126 (4.0%) | 209 (6.6%) | 196 (6.1%) |
| **Established comorbidities** | | | | | | | | |
| 0 | 594 (11.8%) | 37 (0.7%) | 87 (1.7%) | 99 (2.0%) | 910 (18.4%) | 37 (0.7%) | 109 (2.2%) | 111 (2.2%) |
| 1 | 1607 (19.9%) | 145 (1.8%) | 260 (3.2%) | 258 (3.2%) | 1936 (24.2%) | 116 (1.5%) | 258 (3.2%) | 296 (3.7%) |
| 2 | 1971 (28.0%) | 191 (2.7%) | 361 (5.1%) | 351 (5.0%) | 2335 (33.3%) | 167 (2.4%) | 371 (5.3%) | 353 (5.0%) |
| 3 | 1824 (39.3%) | 203 (4.4%) | 397 (8.6%) | 354 (7.6%) | 2045 (43.8%) | 218 (4.7%) | 386 (8.3%) | 372 (8.0%) |
| ≥4 | 3786 (58.8%) | 548 (8.5%) | 924 (14.4%) | 873 (13.6%) | 4018 (60.9%) | 492 (7.5%) | 933 (14.1%) | 883 (13.4%) |
| **Possible comorbidities** | | | | | | | | |
| 0 | 201 (9.3%) | 17 (0.8%) | 32 (1.5%) | 49 (2.3%) | 434 (19.9%) | 35 (1.6%) | 70 (3.2%) | 61 (2.8%) |
| 1 | 1478 (20.3%) | 151 (2.1%) | 252 (3.5%) | 265 (3.6%) | 2006 (27.7%) | 141 (1.9%) | 318 (4.4%) | 342 (4.7%) |
| 2 | 3938 (31.9%) | 488 (4.0%) | 867 (7.0%) | 805 (6.5%) | 4190 (34.5%) | 371 (3.1%) | 765 (6.3%) | 752 (6.2%) |
| 3 | 2800 (40.7%) | 318 (4.6%) | 564 (8.2%) | 564 (8.2%) | 2969 (43.9%) | 306 (4.5%) | 557 (8.2%) | 569 (8.4%) |
| ≥4 | 1365 (52.9%) | 150 (5.8%) | 314 (12.2%) | 252 (9.8%) | 1645 (56.9%) | 177 (6.1%) | 347 (12.0%) | 291 (10.1%) |

**Table 5. Outcome by cancer and treatment characteristics.**

|  | **Total**  **(n=31880)**  **(%: column percent)** | **Hospitalization**  **(n=11498)**  **(%: rate)** | **Ventilation**  **(n=1053)**  **(%: rate)** | **ICU**  **(n=2099)**  **(%: rate)** | **Mortality**  **(n=2052)**  **(%: rate)** |
| --- | --- | --- | --- | --- | --- |
| **Type of malignancy** | | | | | |
| Solid tumor | 27195 (85.3%) | 9441 (34.7%) | 817 (3.0%) | 1709 (6.3%) | 1642 (6.0%) |
| Hematologic | 4382 (13.7%) | 1944 (44.4%) | 221 (5.0%) | 363 (8.3%) | 388 (8.9%) |
| Multiple | 303 (1.0%) | 113 (37.3%) | 15 (5.0%) | 27 (8.9%) | 22 (7.3%) |
| **Recent systemic therapy** | | | | | |
| No | 28884 (90.6%) | 9731 (33.7%) | 896 (3.1%) | 1778 (6.2%) | 1770 (6.1%) |
| Yes | 2996 (9.4%) | 1767 (59.0%) | 157 (5.2%) | 321 (10.7%) | 282 (9.4%) |
| **Recent radiation** | | | | | |
| No | 31483 (98.8%) | 11257 (35.8%) | 1028 (3.3%) | 2058 (6.5%) | 1991 (6.3%) |
| Yes | 397 (1.2%) | 241 (60.7%) | 25 (6.3%) | 41 (10.3%) | 61 (15.4%) |
| **Age at cancer diagnosis** | | | | | |
| 0-20 | 149 (0.5%) | 41 (27.5%) | 1 (0.7%) | 2 (1.3%) | 0 (0.0%) |
| 20-40 | 2033 (6.4%) | 427 (21.0%) | 17 (0.8%) | 49 (2.4%) | 16 (0.8%) |
| 40-60 | 10103 (31.7%) | 2464 (24.4%) | 213 (2.1%) | 428 (4.2%) | 180 (1.8%) |
| 60-80 | 15696 (49.2%) | 6329 (40.3%) | 664 (4.2%) | 1278 (8.1%) | 1128 (7.2%) |
| ≥ 80 | 3899 (12.2%) | 2237 (57.4%) | 158 (4.1%) | 342 (8.8%) | 728 (18.7%) |
| **Survival years ≥ 5** | | | | | |
| No | 25799 (80.9%) | 9712 (37.6%) | 880 (3.4%) | 1777 (6.9%) | 1730 (6.7%) |
| Yes | 6081 (19.1%) | 1786 (29.4%) | 173 (2.8%) | 322 (5.3%) | 322 (5.3%) |

**Table 6. Results of logistic regression analysis indicate odds ratio of exposure variables.**

|  | **Mortality** | | **Severity** | |
| --- | --- | --- | --- | --- |
|  | **Odds ratio** | **p-value** | **Odds ratio** | **p-value** |
| **Age** | | | | |
| 50-65 | 1.88 (1.35,2.61) | <0.0001 | 1.00 (0.92,1.10) | 0.971 |
| 65-75 | 4.09 (2.98,5.63) | <0.0001 | 1.46 (1.33,1.60) | <0.0001 |
| ≥75 | 10.38 (7.59,14.18) | <0.0001 | 2.56 (2.34,2.81) | <0.0001 |
| 18-50 (ref) |  |  |  |  |
| **Gender** | | | | |
| Male | 1.32 (1.20,1.45) | <0.0001 | 1.20 (1.14,1.27) | <0.0001 |
| Female (ref) |  |  |  |  |
| **Race and ethnicity** | | | | |
| Hispanic | 1.05 (0.84,1.30) | 0.691 | 1.65 (1.49,1.82) | <0.0001 |
| Non-Hispanic Black | 1.32 (1.14,1.52) | <0.0001 | 1.89 (1.74,2.04) | <0.0001 |
| Others/Unknown | 1.24 (1.06,1.45) | 0.008 | 1.25 (1.15,1.36) | <0.0001 |
| Non-Hispanic White (ref) |  |  |  |  |
| Established comorbidities | 1.21 (1.18,1.25) | <0.0001 | 1.37 (1.35,1.39) | <0.0001 |
| Possible comorbidities | 1.06 (1.02,1.11) | 0.008 | 1.17 (1.14,1.20) | <0.0001 |
| **Type of malignancy** | | | | |
| Hematologic malignancy | 1.52 (1.35,1.71) | <0.0001 | 1.56 (1.45,1.67) | <0.0001 |
| Multiple cancers | 1.20 (0.76,1.89) | 0.433 | 1.10 (0.85,1.41) | 0.478 |
| Solid tumor (ref) |  |  |  |  |
| **Recent systemic therapy** | | | | |
| Yes | 1.35 (1.17,1.55) | <0.0001 | 2.67 (2.45,2.90) | <0.0001 |
| **Recent radiation therapy** | | | | |
| Yes | 3.09 (2.29,4.17) | <0.0001 | 2.74 (2.19,3.43) | <0.0001 |
| **Survival years ≥ 5** | | | | |
| Yes | 0.71 (0.63,0.81) | <0.0001 | 0.63 (0.59,0.67) | <0.0001 |

**APPENDIX**

| **Table A.1 Statistics of comorbidities.** | **Before matching** | | **After matching** | |
| --- | --- | --- | --- | --- |
|  | **No cancer (n=514538)** | **Cancer (n=31880)** | **No cancer(n=31219)** | **Cancer (n=31219)** |
| **Established comorbidities** | | | | |
| Chronic kidney disease | 38891 (7.6%) | 7798 (24.5%) | 6757 (21.6%) | 7612 (24.4%) |
| Chronic Obstructive pulmonary disease | 29462 (5.7%) | 6061 (19.0%) | 5307 (17.0%) | 5916 (18.9%) |
| Down syndrome | 395 (0.1%) | 9 (0.0%) | 19 (0.1%) | 8 (0.0%) |
| Immunocompromised state from  Solid organ transplant | 177 (0.0%) | 83 (0.3%) | 24 (0.1%) | 80 (0.3%) |
| Obesity | 212518 (47.2%) | 13796 (44.0%) | 15059 (50.7%) | 13519 (44.0%) |
| Pregnancy | 14868 (2.9%) | 188 (0.6%) | 208 (0.7%) | 182 (0.6%) |
| Heart failure | 30768 (6.0%) | 5820 (18.3%) | 5774 (18.5%) | 5680 (18.2%) |
| Coronary artery disease | 49553 (9.6%) | 8973 (28.1%) | 8706 (27.9%) | 8786 (28.1%) |
| Cardiomyopathies | 11044 (2.1%) | 2171 (6.8%) | 1760 (5.6%) | 2124 (6.8%) |
| Sickle cell disease | 1308 (0.3%) | 116 (0.4%) | 58 (0.2%) | 113 (0.4%) |
| Smoking | 140015 (27.2%) | 14140 (44.4%) | 12697 (40.7%) | 13834 (44.3%) |
| Type 2 diabetes mellitus | 87630 (17.0%) | 10731 (33.7%) | 11062 (35.4%) | 10493 (33.6%) |
| **Possible comorbidities** | | | | |
| Asthma | 69564 (13.5%) | 5491 (17.2%) | 5822 (18.6%) | 5360 (17.2%) |
| Cerebrovascular | 7757 (1.5%) | 1527 (4.8%) | 1443 (4.6%) | 1477 (4.7%) |
| Cystic fibrosis | 14780 (2.9%) | 1408 (4.4%) | 1367 (4.4%) | 1366 (4.4%) |
| Hypertension | 183861 (35.7%) | 22279 (69.9%) | 21713 (69.6%) | 21819 (69.9%) |
| Immunodeficiency, unspecified | 1757 (0.3%) | 710 (2.2%) | 187 (0.6%) | 693 (2.2%) |
| Liver disease | 29029 (5.6%) | 5198 (16.3%) | 3150 (10.1%) | 5096 (16.3%) |
| Dementia | 15695 (3.1%) | 2485 (7.8%) | 3230 (10.3%) | 2390 (7.7%) |
| Overweight | 346018 (76.8%) | 23786 (75.9%) | 24282 (81.7%) | 23296 (75.9%) |
| Pulmonary fibrosis | 6565 (1.3%) | 1679 (5.3%) | 1134 (3.6%) | 1643 (5.3%) |
| Thalassemia | 1096 (0.2%) | 139 (0.4%) | 80 (0.3%) | 136 (0.4%) |
| Type 1 diabetes mellitus | 7927 (1.5%) | 858 (2.7%) | 996 (3.2%) | 843 (2.7%) |

**Figure A.1 Frequency distribution of cancer types.**

**Table A.2 Characteristics of cancer types and treatment.**

|  |  |
| --- | --- |
| **Characteristics of Cancer (n=31880)** | |
| **Type of malignancy** | |
| Solid tumor | 27195 (85.3%) |
| Hematologic | 4382 (13.7%) |
| Multiple | 303 (1.0%) |
| **Recent systemic therapy** | |
| No | 28884 (90.6%) |
| Yes | 2996 (9.4%) |
| **Recent radiation therapy** | |
| No | 31483 (98.8%) |
| Yes | 397 (1.2%) |
| **Age at cancer diagnosis** | |
| 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%) |