eLife

Epidemiology and Global Health
Cancer Biology

Disentangling the relationship between cancer mortality and COVID-19 in the US

Reviewed Preprint
v2 • July 16, 2024
Revised by authors

Reviewed Preprint
v1 • February 28, 2024

Chelsea L Hansen ✉, **Cécile Viboud, Lone Simonsen**

Division of International Epidemiology and Population Studies, Fogarty International Center, National Institutes of Health, Bethesda, MD, USA. 20892 • PandemiX Center, Dept of Science & Environment, Roskilde University, Denmark • Brotman Baty Institute, University of Washington, Seattle, WA

Abstract

Several countries have reported that deaths with a primary code of cancer did not rise during COVID-19 pandemic waves compared to baseline pre-pandemic levels. This is in apparent conflict with findings from cohort studies where cancer has been identified as a risk factor for COVID-19 mortality. Here we further elucidate the relationship between cancer mortality and COVID-19 on a population level in the US by testing the impact of death certificate coding changes during the pandemic and leveraging heterogeneity in pandemic intensity across US states. We computed excess mortality from weekly deaths during 2014-2020 nationally and for three states with distinct COVID-19 wave timing (NY, TX, and CA). We compared pandemic-related mortality patterns from underlying and multiple cause (MC) death data for six types of cancer and compared to that seen for chronic conditions such as diabetes and Alzheimer's. Any death certificate coding changes should be eliminated by study of MC data.

Nationally in 2020, we found only modest excess MC cancer mortality (~13,600 deaths), representing a 3% elevation over baseline level. Mortality elevation was measurably higher for less deadly cancers (breast, colorectal, and hematologic, 2-7%) than cancers with a poor 5-year survival (lung and pancreatic, 0-1%). In comparison, there was substantial elevation in MC deaths from diabetes (37%) and Alzheimer's (19%). Homing in on the intense spring 2020 COVID-19 wave in NY, mortality elevation was 1-16% for different types of cancer and 128% and 49% for diabetes and Alzheimer's, respectively. To investigate the peculiar absence of excess mortality on deadly cancers, we implemented a demographic model and simulated the expected covid-related mortality using COVID-19 attack rates, life expectancy, population size and mean age for each chronic condition. This model indicates that these factors largely explain the considerable differences in observed excess mortality between these chronic conditions during the COVID-19 pandemic, even if cancer had increased the relative risk of mortality by a factor of 2 or 5.

In conclusion, we found limited elevation in cancer mortality during COVID-19 waves, even after considering MC mortality, and this was especially pronounced for the deadliest cancers. Our demographic model predicted low expected excess mortality in populations living with certain types of cancer, even if cancer is a risk factor for COVID-19 fatality, due to competing mortality risk. We also find a moderate increase in excess mortality from hematological cancers, aligned with other types of observational studies. While our study concentrates on the immediate consequences of the COVID-19 pandemic on cancer mortality in 2020, further

research should consider excess mortality in the complete pandemic period. Also, a study of the delayed impact of the pandemic on cancer mortality due to delayed diagnosis and treatment during the pandemic period is warranted.

eLife assessment

This **valuable** work explores death coding data to understand the impact of COVID-19 on cancer mortality. The work provides **solid** evidence that deaths with cancer as a contributing cause were not above what would be expected during pandemic waves, suggesting that cancer did not strongly increase the risk of dying of COVID-19. These results are an interesting exploration into the coding of causes of death that can be used to make sense of how deaths are coded during a pandemic in the presence of other underlying diseases, such as cancer.

https://doi.org/10.7554/eLife.93758.2.sa2

Introduction

The dominant risk factors for COVID-19 mortality have consistently been shown to be advanced age, male gender and certain chronic diseases such as diabetes, obesity and heart disease (Chavez-MacGregor et al., 2022; Rüthrich et al., 2021; Williamson et al., 2020). Cancer has also been identified as a high-risk condition based on case-control and cohort studies, although these studies have provided conflicting results. In a large cohort study of ~500,000 COVID-19 inpatients, only cancer patients under recent treatment were at increased risk of COVID-19 related deaths (OR=1.7) relative to non-cancer patients (Chavez-MacGregor et al., 2022). Conversely, a smaller European study of 3,000 COVID-19 inpatients found that cancer was not a risk factor (Rüthrich et al., 2021), as did an international, multicenter study of 4,000 confirmed COVID-19 inpatients (Raad et al., 2023). More recently a meta-analysis of 35 studies from Europe, North America, and Asia found a 2-fold increased risk of COVID-19 mortality among cancer patients (Di Felice et al., 2022). Similarly, a large analysis from the UK found that the risk of COVID-19 mortality for cancer patients had declined over the course of the pandemic but remained 2.5 times higher than for non-cancer patients into 2022 (Starkey et al., 2023). Taken together, such observational studies provide a mixed picture of cancer as a COVID-19 mortality risk factor, with several studies reporting that controlling for important factors such as age is a challenge. Furthermore, cancer is often considered as a single disease category despite the diversity of conditions and patients represented.

Further evidence for the relationship between cancer and COVID-19 comes from population-level analysis of vital statistics. A recent US study showed no elevation in underlying cancer deaths concomitant with COVID-19 waves, in stark contrast to the sharp rise in mortality from other chronic diseases (W.-E. Lee et al., 2023). In several other countries, including Sweden, Italy, Latvia, Brazil, England and Wales, underlying cancer mortality was found to be stable or decreasing during the first year of the pandemic (Alicandro et al., 2023; Fernandes et al., 2021; Gobina et al., 2022; Grande et al., 2022; Kontopantelis et al., 2022; Lundberg et al., 2023). Further, an excess mortality study of 240,000 cancer patients in Belgium found a 33% rise in mortality in April 2020, but concluded that this was no different from the rise observed in the general population (Silversmit et al., 2021). The apparent lack of association between cancer mortality and COVID-19 on a population level raises the question of the true relationship between cancer and COVID-19.

The relationship between these two diseases could occur via multiple biological mechanisms. First, immunosuppression in cancer patients could increase susceptibility to SARS-CoV-2 infection and/or risk of severe clinical outcome upon infection. Conversely, immunosuppression could act as a protective factor in the face of a severe respiratory infection that kills by over-stimulating the immune system – the immune incompetence rescue hypothesis (Reichert 2004). This hypothesis was put forward to explain the observed absence in excess cancer mortality during the 1968 influenza pandemic, a departure from elevated mortality seen for other high-risk conditions such as heart disease and diabetes (Reichert 2004). A further mechanism that could affect the observed relationship between cancer deaths and COVID-19 is changing guidelines for establishing the primary cause of death. Coding guidelines evolved throughout the pandemic as testing for SARS-CoV-2 infection became more widespread, which presumably affected vital statistics studies.

To further elucidate the relationship between cancer mortality and COVID-19 on a population level, we analyzed US vital statistics in detail to understand the potential role of death certificate coding changes during the pandemic and explored putative differences in mortality patterns between different types of cancer. We considered death certificates listing cancer as the underlying cause of death (UC) or anywhere on the death certificate (multiple-cause (MC)). Assuming there is a high propensity to attribute a primary code of COVID-19 during the pandemic in any patient with COVID-19, deaths among individuals with both cancer and COVID-19 near the time of death would be coded as UC COVID-19. However, cancer should still be captured in the MC data, and thus, analysis of MC death data should control for any changes in death certificate coding practices during the pandemic (Fedeli et al., 2024). The US provides a particularly useful case study as the timing of COVID-19 waves varied considerably between states, so that elevations in cancer deaths, should they exist, should also be heterogeneous. For comparison, we also assessed population-level excess mortality patterns for other chronic conditions such as diabetes, ischemic heart disease (IHD), kidney disease, and Alzheimer's, for which the association with COVID-19 is less debated.

Results

Establishing patterns and timing of COVID-19 related deaths

We obtained individual ICD-10 coded death certificate data from the US for the period January 1, 2014, to December 31, 2020. We compiled time series by week and cause of death, for underlying cause (UC) and for multiple-cause (MC) mortality. We considered 10 causes of death, including diabetes, Alzheimer's disease, ischemic heart disease (IHD), kidney disease, and 6 types of cancer (all-cause cancer, colorectal, breast, pancreatic, lung, and hematological; see **Table 1** and **Appendix 1 - Table 1** for a list of disease codes). We chose these specific cancers to illustrate conditions for which the 5-year survival rate is low (13% and 25%, respectively, for pancreatic and lung cancers) and high (65% and 91%, respectively, for colorectal and breast cancers) (National Cancer Institute, n.d.). Hematological cancer (67% 5-year survival rate) was included because it has been singled out as a risk factor in several previous studies due to the immune suppression associated with both its malignancy and treatment. (Chavez-MacGregor et al., 2022; X. Han et al., 2022; Rüthrich et al., 2021; Williamson et al., 2020). To compare mortality patterns with the timing of COVID-19 pandemic waves, we accessed national and state-level counts of reported COVID-19 cases from the Centers for Disease Control and Prevention (CDC)(Centers for Disease Control and Prevention, 2022).

In national data, time series of COVID-19-coded death certificates (both UC and MC) tracked with the temporal patterns of laboratory-confirmed COVID-19 cases (**Figure 1**), revealing three distinct COVID-19 waves: a spring wave peaking on April 12, 2020, a smaller summer wave

Year	Diagnosis Group	ICD-10 Codes	No. Deaths	Mean age, years (IQR)	%Home/ER	%Nursing Home
2019	Cancer	C00-C99	493,397	72 (64-81)	45	12
	Pancreatic Cancer	C25	37,864	72 (64-80)	51	9
	Colorectal Cancer	C18-C20	42,484	71 (61-82)	46	13
	Hematologic Cancers	C81-C96	47,174	74 (67-84)	35	11
	Diabetes	E10-E14	70,763	72 (63-82)	53	17
	Alzheimer's	G30	98,675	87 (82-92)	29	50
2020	Cancer	C00-C99	513,275	72 (64-81)	55	8
	Pancreatic Cancer	C25	39,893	72 (65-80)	61	6
	Colorectal Cancer	C18-C20	43,990	71 (61-82)	56	9
	Hematologic Cancers	C81-C96	49,161	74 (67-84)	46	8
	Diabetes	E10-E14	88,124	71 (62-82)	58	15
	Alzheimer's	G30	115,256	86 (82-92)	33	46

Table 1.

Each diagnosis group and its corresponding ICD-10 codes, number of underlying deaths, mean age in years at time of death, the percentage of deaths occurring at home, and the percentage of deaths occurring in nursing homes for 2019 and 2020

peaking on July 26, 2020, and a large winter wave that had not yet peaked by the end of the study in December 2020. This correspondence between COVID-19 case and death activity represents a "signature" mortality pattern of COVID-19.

Analysis of state-level data reveals variable timing, intensity, and number of COVID-19 waves across the US during 2020. To focus on periods with substantial COVID-19 activity and explore the association with cancer, we identified three large US states with unique, well-defined waves (**Figure 1**). New York (NY) state experienced a large, early wave in March-May 2020, based on recorded COVID-19 cases and deaths and high seroprevalence of SARS-CoV-2 antibodies in this period (over 20% (Stadlbauer et al., 2021)). Meanwhile, California (CA) experienced a large COVID-19 wave at the end of the year and had little activity during the spring and summer. Finally, Texas (TX) had two large waves; one during late summer, followed by one in winter 2020.

National patterns in excess mortality from cancer

Similar to other influenza and COVID-19 population-level mortality studies (Islam et al., 2021; Karlinsky and Kobak, 2021; W.-E. Lee et al., 2023; Msemburi et al., 2023), we established a weekly baseline model for expected mortality in the absence of pandemic activity by modeling time trends and seasonality in pre-pandemic data and letting the model run forward during the pandemic (see Methods). Each cause of death (UC and MC) and geography (aggregated National, NY, TX, and CA) was modeled separately. We then computed excess mortality as the difference between observed deaths and the model-predicted baseline. We summed weekly estimates to calculate excess mortality for the full pandemic period and during each of the 3 waves (see Methods). In addition to these absolute effects of the pandemic on mortality, we also calculated the relative effects by dividing excess mortality by baseline mortality. This approach has been used in the past to standardize mortality effects in strata with very different underlying risks (e.g., age groups, geographies, or causes of death, see Methods).

Nationally, we found a drop in UC cancer deaths during spring 2020 (**Figure 2**, panel **a**; **Table 2**), although the drop was not statistically significant. A similar non-significant decline was also seen for specific cancer types (**Figure 2**, panels **d-f**; **Appendix 1 - Figure 1**, panels **a,f-j**). Further, pre-pandemic mortality trends for each cancer type continued unabated during the first pandemic year. We reasoned that this early drop in UC cancer deaths may be explained by changes in coding practices, so we next turned to MC mortality to resolve this question.

Time series of MC cancer mortality showed a significant increase in all three waves (**Figure 2**, panel **a**; **Appendix 1 - Table 2**). A similar pattern was seen in MC time series for colorectal (**Figure 2**, panel **h**), breast (**Appendix 1 - Figure 1**, panel **i**), and hematological cancer (**Appendix 1 - Figure 1**, panel **j**). However, the total excess mortality was modest with 13,600 excess cancer deaths in 2020, representing a statistically significant 3% elevation over baseline (**Table 2**). The largest relative increase in MC mortality was observed in hematological cancer at 7% (statistically significant, 3,600 excess deaths). No excess in MC mortality was seen for the two deadliest cancers, pancreatic cancer (**Figure 2**, panel **f**) and lung cancer (**Appendix 1 - Figure 1**, panel **g**).

National patterns in deaths due to other chronic conditions

We considered diabetes and Alzheimer's as "positive controls" as they are also considered COVID-19 risk factors and can illustrate how positive associations between chronic conditions and COVID-19 manifest in population level excess mortality studies. Diabetes provides a particularly useful comparator for cancer as the mean age at death is similar (~72 years, **Table 1**) and because few individuals live in a nursing home (Appendix 1 - Supplemental Methods). Mortality time series from UC and MC diabetes and Alzheimer's were highly correlated with COVID-19 activity, with statistically significant mortality elevation synchronous with pandemic waves (**Figure 2 b-c**; **Appendix 1 - Figures 2-5**). For diabetes, we measured an excess of 10,800 and 82,300 deaths

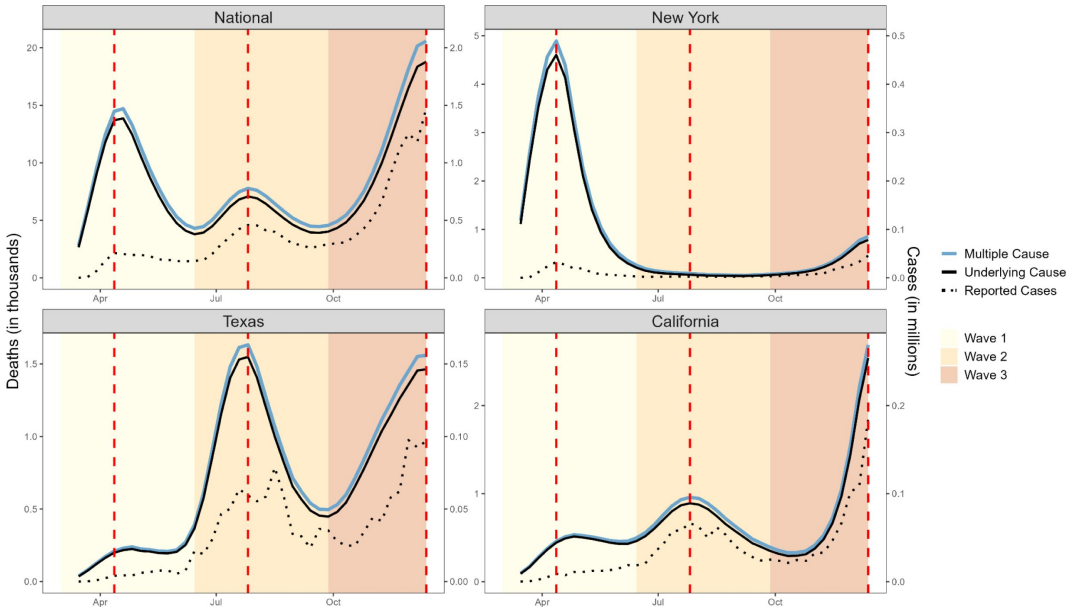


Figure 1.

Weekly counts of death certificates listing COVID-19 as either the underlying or a multiple cause

When included on a death certificate, COVID-19 was most often listed as the underlying cause of death rather than a contributing cause. National-level data reveal three distinct waves: Wave 1 (spring, March 1 - June 27, 2020), Wave 2 (summer, June 28 - October 3, 2020), and Wave 3 (winter, October 4 - December 6, 2020, incomplete). Vertical dashed lines represent the peak of each wave, dotted lines represent the number of reported cases (y-axis on the right). New York experienced its first large COVID-19 wave in Wave 1, while Texas had its first large wave in Wave 2 and California did not experience a large wave until Wave 3 which had not yet peaked at the end of 2020.

Figure 2.

National-level weekly observed and estimated baseline mortality for each diagnosis group as both the underlying cause or anywhere on the death certificate (multiple cause) from 2014 to 2020. Baselines during the pandemic are projected based on the previous years of data.

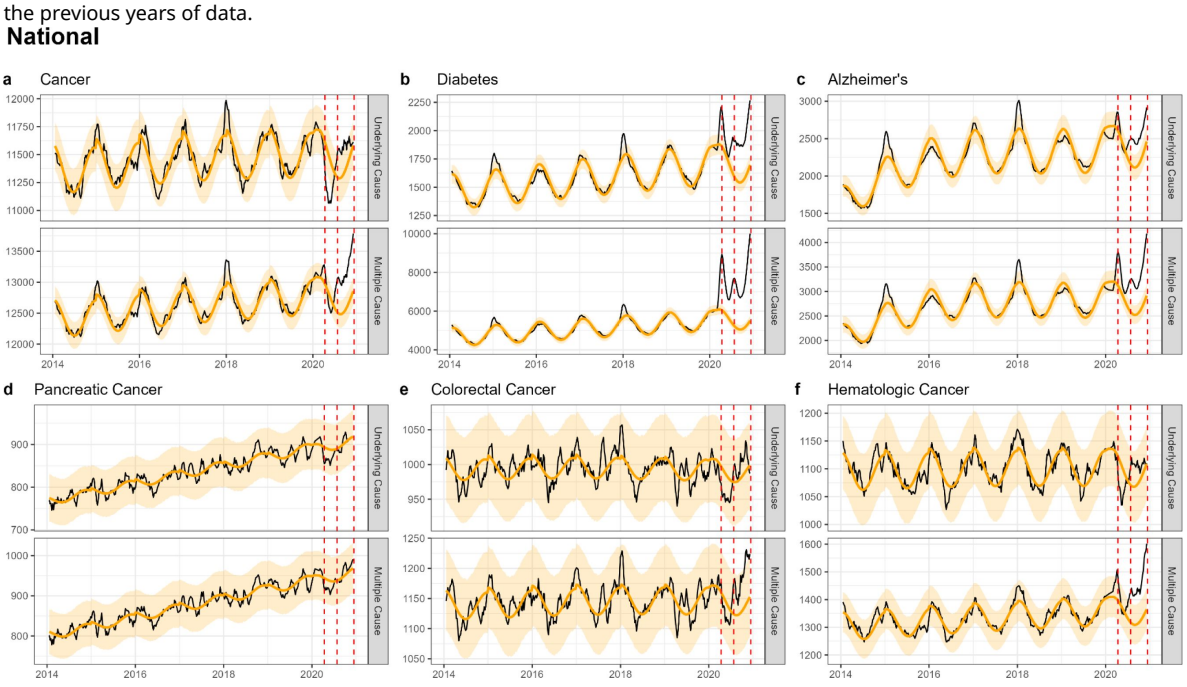


Table 2.

The estimated number of excess deaths and the percentage over baseline for each diagnosis group when listed as both the underlying cause or anywhere on the death certificate (multiple cause). Estimates for the national-level data are provided for the full pandemic period and for each state based on when the first large wave was experienced.

			Multiple Cause		Underlying Cause	
Cause of Death	State	Wave	Excess Deaths	% Over Baseline	Excess Deaths	% Over Baseline
Cancer	National	Overall	13601*	3.0	11	0.0
	New York	1	747	6.0	-474	-5.0
	Texas	2	467	4.0	39	0.0
	California	3	529	4.0	82	1.0
Pancreatic Cancer	National	Overall	-25	-0.0	-282	-1.0
	New York	1	8	1.0	-16	-2.0
	Texas	2	17	2.0	24	3.0
	California	3	0	0.0	-18	-2.0
Colorectal Cancer	National	Overall	988	2.0	-168	-0.0
	New York	1	91	9.0	-16	-2.0
	Texas	2	4	0.0	-34	-3.0
	California	3	27	2.0	-1	-0.0
Hematologic Cancers	National	Overall	3615*	7.0	111	0.0
	New York	1	121	10.0	-107	-11.0
	Texas	2	136	11.0	21	2.0
	California	3	114	8.0	20	2.0
Diabetes	National	Overall	82318*	37.0	10784*	16.0
	New York	1	5945*	128.0	568*	40.0
	Texas	2	4612*	77.0	420*	23.0
	California	3	3474*	59.0	575*	33.0
Alzheimer's	National	Overall	21712*	19.0	8528*	9.0
	New York	1	734*	49.0	188	16.0
	Texas	2	1398*	45.0	805*	31.0
	California	3	726*	18.0	259	8.0

*Confidence interval does not include zero

(UC and MC, respectively), corresponding to statistically significant elevations of 16% and 37% over baseline level mortality (**Table 2**). For Alzheimer's, we estimated 8,500 and 21,700 excess deaths, corresponding to statistically significant elevations of 9% and 19% elevation over baseline, respectively. Pandemic-related excess mortality was also seen for IHD and kidney disease (see supplement for estimates, **Appendix 1 - Table 2**).

State-level patterns in excess mortality

Similar to patterns seen in national level data, none of the state-level analyses revealed notable increases in UC cancer mortality, while there was a modest, non-significant increase in MC cancer mortality (**Figures 3**-**5**; **Appendix 1 - Figures 6**-**8**). The largest mortality increase was seen in NY during the spring wave, with a 6% rise in MC cancer mortality above the model baseline (**Table 2**; **Appendix 1 - Table 3**). The magnitude of the increase seen during the spring wave varied by cancer type, with minimal increases seen in pancreatic and lung cancers (1%) and higher increases in colorectal, hematological, and breast cancers (9%, 10%, and 16%, respectively). For comparison, there was a statistically significant rise in Alzheimer's and diabetes deaths during this wave of 49% and 128%.

In CA and TX, mortality fluctuations were less pronounced than in NY, coinciding with less intense COVID-19 waves, and this was seen across all conditions. MC excess mortality estimates remained within +/-4% of baseline levels for cancers, irrespective of the type of cancer and pandemic wave, except for hematological cancers which saw an 11% rise in Texas during the summer wave and an 8% rise in California during the winter wave. None of these elevations were statistically significant. In comparison, there was statistically significant excess mortality elevation for both Alzheimer's and diabetes deaths (range, 18-59% in the CA winter wave, and 45-77% in the TX summer wave, **Table 2**, **Appendix 1-Tables 4**-**5**).

Demographic mortality projections under the null hypothesis that cancer in and of itself is not a risk factor for COVID-19 mortality

Next, to get a sense of the expected mortality elevation, we ran simulations to gauge the level of individual-level association (traditionally measured as relative risk, RR) between COVID-19 and the studied chronic conditions that is consistent with the population-level excess mortality patterns observed during the pandemic. Using cancer as an example, two main factors could drive cancer mortality patterns during COVID-19, namely the size and age of the population living with cancer (since age is such a pronounced risk factor for COVID-19), and the life expectancy under cancer diagnosis. These factors would operate irrespective of the true biological relationship between COVID-19 severity and cancer. The same logic applies to mortality from other chronic conditions, such as diabetes or Alzheimer's.

To test the hypothesis that these population factors alone could explain differences in excess mortality between chronic conditions, we designed a simple model of COVID-19 mortality for individuals with chronic conditions (see methods for details). The model projected excess mortality during the pandemic under the null hypothesis that the chronic condition was not in and of itself a risk factor for COVID-19 mortality, with only the demography of the population living with the disease (namely, the age and size of the at-risk populations and baseline risk of death from each condition) affecting excess mortality. In the demographic model, we first estimated the number of expected SARS-CoV-2 infections among persons with a certain condition, by multiplying the estimated number of US individuals living with the condition (CDC, Division of Population Health, n.d.; Dhana et al., 2023; Rajan et al., 2021; U.S. Cancer Statistics Working Group, released in November 2023) by the reported SARS-CoV-2 seroprevalence at the end of our study period (December 2020 for the national, or after each wave for the state data)(Centers for Disease Control and Prevention, 2023). We focused on seroprevalence among individuals ≥65 years, the most relevant age group for the conditions we considered (we also run a sensitivity analysis considering seroprevalence in adults 50-64 years, see discussion). We then multiplied the

Figure 3.

The same as figure 1, but for New York. New York experienced its first large wave of COVID-19 in spring 2020 (Wave 1)

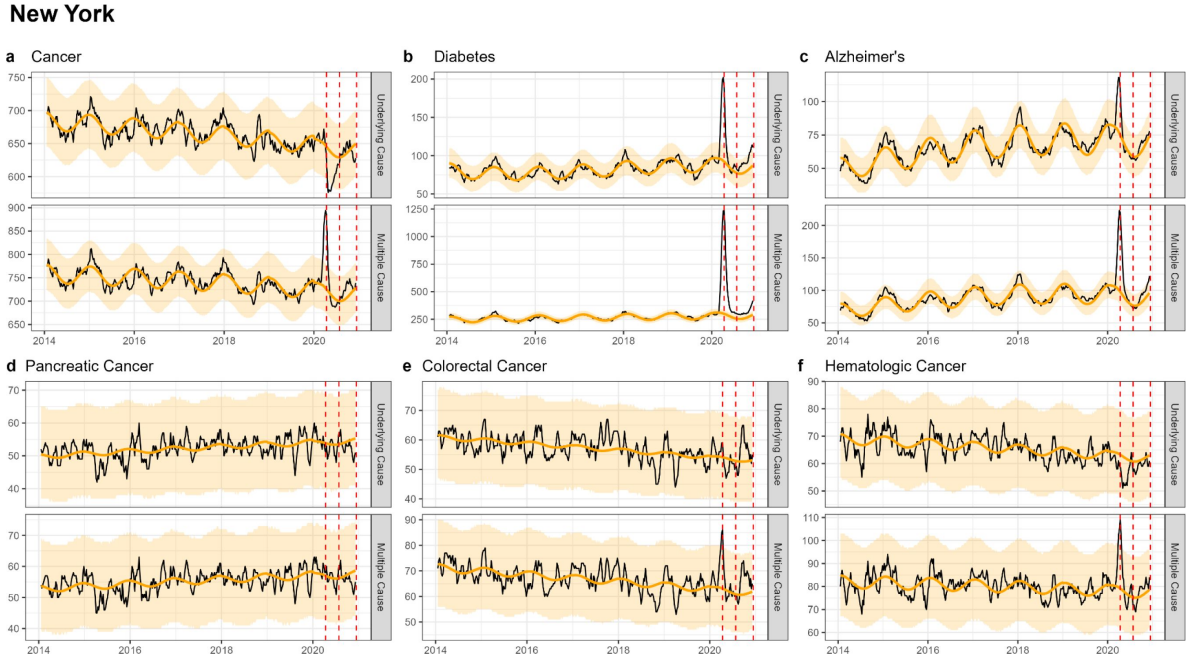
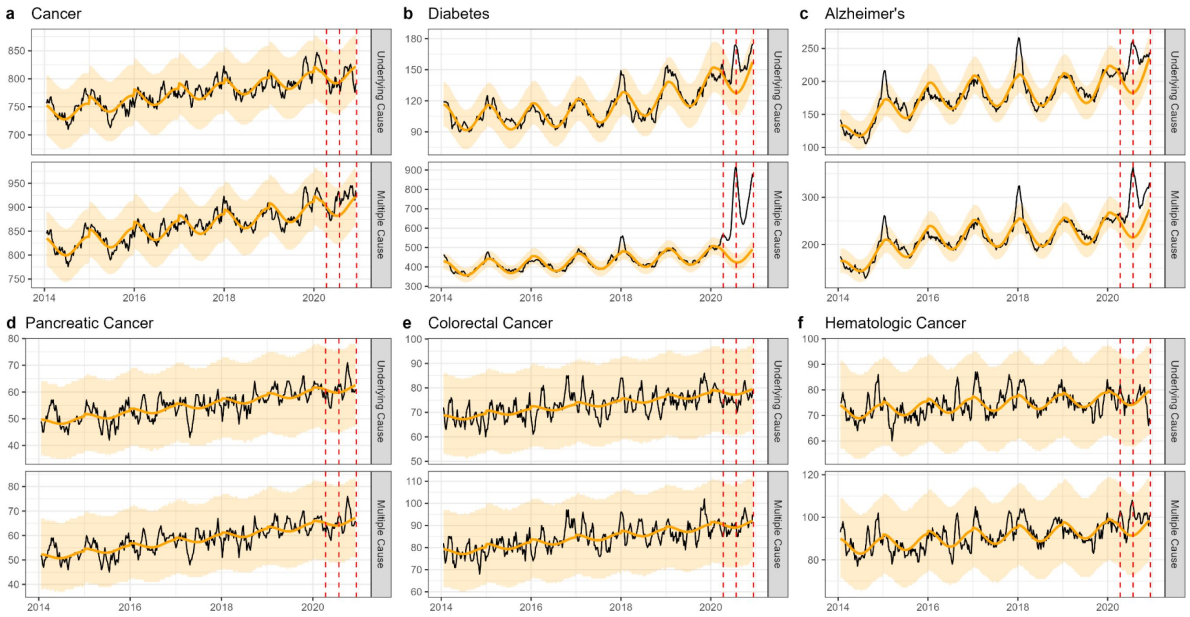


Figure 4.

The same as figure 1, but for Texas. Texas experienced its first large wave of COVID-19 in the summer of 2020 (Wave 2)



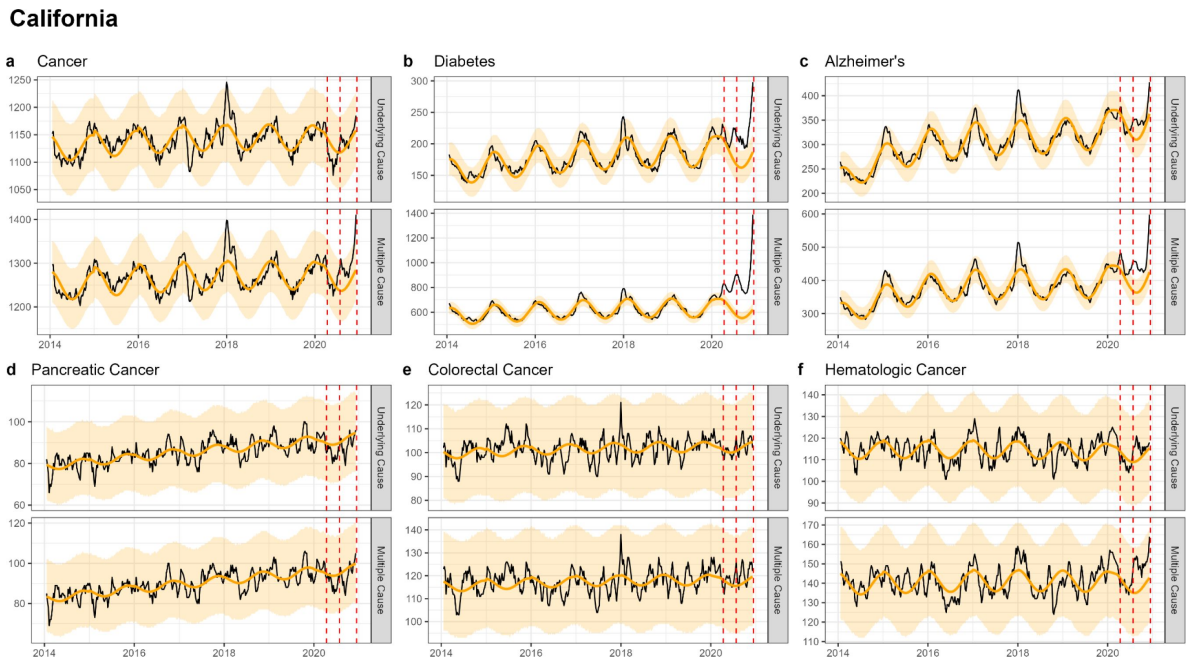


Figure 5.

The same as figure 1, but for California. California did not experience a large wave of COVID-19 until the winter of 2020-2021 (Wave 3), only the first half of which is captured here

Chronic condition	State	Population at risk	Mean age	Wave	Observed MC deaths over same period in 2019	Observed excess (% over baseline) in 2020	Expected excess (null)	Expected excess (RR=2)	Expected excess (RR=5)
All cancers	National	5718925	65	Overall	546453	3 (1-4)	1 (1-2)	2 (1-4)	6 (4-10)
	New York	400891	65	Wave 1	12244	6 (-1-15)	4 (2-10)	9 (3-20)	22 (8-51)
	Texas	397993	63	Wave 2	12187	4 (-3-11)	2 (1-6)	5 (2-12)	11 (4-29)
	California	599552	64	Wave 3	16713	4 (-1-10)	2 (0-5)	4 (1-9)	9 (2-23)
Pancreatic	National	66319	67	Overall	39798	0 (-6-7)	0 (0-0)	0 (0-1)	1 (1-2)
	New York	2584	67	Wave 1	963	1 (-21-35)	0 (0-1)	1 (0-2)	2 (1-5)
	Texas	2264	66	Wave 2	882	2 (-19-34)	0 (0-1)	0 (0-1)	1 (0-3)
	California	3482	67	Wave 3	1277	0 (-17-24)	0 (0-0)	0 (0-1)	1 (0-2)
Lung cancer	National	425015	70	Overall	123622	1 (-3-5)	1 (0-1)	1 (1-2)	3 (2-5)
	New York	17709	71	Wave 1	2643	1 (-13-20)	2 (1-4)	3 (1-8)	8 (3-20)
	Texas	12700	70	Wave 2	2513	2 (-11-20)	1 (0-2)	1 (1-4)	4 (1-9)

Table 3.

Projections of COVID-19-related excess mortality patterns for different cancers and chronic conditions in the US, under different hypotheses for the association between the condition and COVID-19.

Projections are provided for the null hypothesis of no biological interaction between the condition and COVID-19; these projection are solely driven by the size and mean age of the population living with each condition (where age determines the infection-fatality ratio from COVID-19), and the baseline risk of death from the condition over a similar time period (March to December 2019 for the national data, and for the states comparable dates in 2019 corresponding to the relevant COVID-19 wave). Additional projections are provided under alternative hypotheses, where each condition is associated with a relative risk (RR) of 2 or 5 for COVID-19 related death (infection-fatality ratio multiplied by 2 or 5).

Table 3. (continued)

	California	19079	70	Wave 3	2861	3 (-10-18)	1 (0-2)	1 (0-3)	3 (1-8)
Hematological	National	459463	62	Overall	57892	7 (1-13)	1 (0-1)	1 (1-2)	3 (2-5)
	New York	15577	62	Wave 1	1305	10 (-11-40)	1 (0-3)	2 (1-5)	6 (2-13)
	Texas	14927	59	Wave 2	1231	11 (-9-38)	1 (0-1)	1 (0-3)	3 (1-7)
	California	21290	61	Wave 3	1916	8 (-8-29)	0 (0-1)	1 (0-2)	2 (1-5)
Colorectal	National	473264	66	Overall	49053	2 (-4-8)	1 (1-2)	2 (1-4)	6 (4-10)
	New York	30859	66	Wave 1	1048	9 (-13-44)	4 (2-10)	9 (3-20)	22 (8-51)
	Texas	36641	65	Wave 2	1224	0 (-18-26)	3 (1-7)	5 (2-13)	13 (4-33)
	California	51863	65	Wave 3	1575	2 (-14-24)	2 (1-5)	4 (1-10)	9 (3-24)
Breast	National	1097917	64	Overall	43519	2 (-4-9)	2 (2-4)	5 (3-8)	12 (8-21)
	New York	74459	64	Wave 1	981	16 (-8-53)	9 (3-21)	18 (7-42)	46 (17-106)
	Texas	77860	62	Wave 2	1019	3 (-17-32)	5 (2-12)	10 (3-24)	24 (8-61)
	California	123433	63	Wave 3	1421	2 (-15-25)	4 (1-10)	8 (2-20)	20 (5-51)
Diabetes	National	29105146	60	Overall	229326	37 (31-43)	8 (5-14)	16 (10-28)	40 (26-69)
	New York	1792926	60	Wave 1	4804	128 (104-158)	30 (11-68)	59 (22-136)	148 (55-340)
	Texas	2450005	58	Wave 2	5898	77 (61-96)	17 (6-44)	35 (12-87)	86 (30-218)
	California	3514440	59	Wave 3	8399	59 (47-74)	12 (3-32)	25 (7-64)	62 (17-160)
Alzheimer's	National	6070000	81	Overall	118993	19 (11-28)	28 (18-48)	57 (36-96)	142 (90-240)
	New York	426500	81	Wave 1	1563	49 (23-87)	191 (70-432)	381 (140-863)	953 (350-2158)
	Texas	459300	80	Wave 2	2974	45 (27-69)	63 (21-158)	126 (43-315)	315 (107-788)
	California	719700	81	Wave 3	5394	18 (6-33)	39 (11-98)	78 (21-196)	195 (53-491)

estimated number of SARS-CoV-2 infections by an age-specific infection-fatality ratio (IFR) for SARS-CoV-2 (COVID-19 Forecasting Team, 2022). This gave an estimate of COVID-19-related deaths, or excess deaths, for a given condition. To estimate a percent elevation over baseline and compare with our vital statistics analysis, we divided the excess death estimate derived from the demographic model by the total deaths for that condition for a similar period of time in 2019 (see Methods). We repeated this analysis for each cancer type, diabetes, and Alzheimer's. In addition to the null hypothesis, we also projected alternative hypotheses of a biological association between chronic conditions and COVID-19, assuming that a given chronic condition would raise the risk of COVID-19 mortality (via the infection fatality ratio) by a factor of 2 or 5. We compared these modeled expectations for the null and alternative hypotheses with the observed excess mortality in 2020, using MC mortality as the outcome (**Table 2**).

Under the null hypothesis we projected a 0-2% elevation over the 2019 baseline in deaths for all cancer types in national data, and 0-9% elevations in state-level data (**Table 3**). Under the alternative hypothesis that cancer increases COVID-19 mortality risk by a factor of 2, the projected elevation is 0-5% in national data and 0-18% in state-level data. In general, the largest projected increases were found in NY state, driven by the higher attack rates. We also see systematic differences in the percent elevation over baseline by type of cancer, related to the lethality of different cancers. For instance, even if cancer increases COVID-19 mortality risk by a factor of 2, we expect to see only a 0-1% increase for particularly deadly cancers such as pancreatic and lung cancer, in part driven by the high competing risk of death from these cancers (short life expectancy) and the small size of the population-at-risk. The expected increases for less deadly cancers, such as colorectal and breast, were notably higher (2-5% in national data, and 9-18% during the large spring wave in New York), in part driven by the lower risk of death from these cancers (longer life expectancy). Based on the observations from our time series analysis of MC mortality in all states, non-hematological cancers are most consistent with a 1- to 2-fold increase in mortality, with the caveat that most of the confidence intervals include zero, and the differences in projected mortality under these hypotheses are minimal. In contrast, for hematological cancer, the observed rise in mortality exceeds the expected elevation even under the assumption of a 5-fold increase in mortality.

We repeated this analysis for diabetes and Alzheimer's (**Table 3**). For diabetes under the null hypothesis, we projected an 8% elevation over baseline in national data and 12-30% in state-level data based on the age distribution and substantial size of the population-at-risk alone. In fact, we observed in vital statistics analysis a 37% elevation over baseline in national US data and 59-128% in state-level data, with the largest increase seen in NY and lowest increase in CA. These observations are most consistent with a 5-fold increase in mortality based on our demographic model (projected elevation 40% nationally and 62-148% at the state level). For Alzheimer's under the null hypothesis, we projected a 28% increase over baseline nationally, and 30-191% increases at the state level, largely driven by the advanced age of the population-at-risk. In contrast, analysis of vital statistics data reveals a 19% increase nationally and 18-49% across states, which is in fact lower than the null hypothesis (we return to this surprising result in the discussion). Strikingly, our demographic model supports that COVID-19 will manifest differently in population-level excess mortality for each of these chronic conditions, even under the null hypothesis of no biological association between viral infection and these underlying comorbidities. Overall, these projections support the idea that demography alone (age, size, and baseline mortality of the population living with each of these conditions) can explain much of the differences in absolute and relative mortality elevations seen during the pandemic across conditions like cancer, diabetes, and Alzheimer's.

Discussion

Cancer is generally thought of as a risk factor for severe COVID-19 outcomes, yet observational studies have produced conflicting evidence. With recent availability of more detailed US vital statistics data, we used statistical time series approaches to generate excess mortality estimates for multiple cause of death data, different types of cancer, and several geographic locations during 2020. We accounted for potential changes in coding practices during the pandemic, for instance capturing a COVID-19 patient with cancer whose death may have been coded as an underlying COVID-19 death and not a cancer death. Based on multiple cause of death data, we estimated 13,600 national COVID-19-related excess cancer deaths, which aligns well with reporting on death certificate data, where 13,400 deaths are ascribed to COVID-19 in cancer patients (**Appendix 1 - Figure 9**) (Fedeli et al., 2024). Yet these deaths only represent a 3% elevation over the expected baseline cancer mortality. Percent mortality elevation was measurably higher for less deadly cancers (breast and colorectal) than cancers with a poor 5-year survival (lung and pancreatic). Consistent with other studies (Chavez-MacGregor et al., 2022; S. Han et al., 2022; Rüthrich et al., 2021; Williamson et al., 2020), we found that the largest mortality increase for specific cancer types was seen in hematological cancers with a 7% elevation over baseline in national data. Across the board, the largest elevations in cancer mortality were observed in the states most impacted by the first year of the COVID-19 pandemic (e.g., NY), lending support to the specificity of our excess mortality approach.

In contrast to cancer, we observed substantial COVID-19-related excess mortality for diabetes and Alzheimer's, temporally and geographically consistent with the three-wave "signature" pattern observed in reported COVID-19 cases and deaths across the US. To investigate whether demographic differences in underlying patient populations (age distribution, population size, and baseline risk of death due to chronic condition) could explain differences in excess mortality during the pandemic, we ran a simple demographic model for each condition – first assuming the condition in and of itself was not a risk factor for COVID-19-related mortality (null hypothesis). Doing so we found that the rise in cancer deaths during COVID-19 was expected to remain low compared to these other chronic conditions, largely driven by the higher risk of death from cancer itself compared to diabetes and Alzheimer's. These demographic projections illustrate the importance of competing risks (**Figure 6**), where the risk of cancer death predominates over the risk of COVID-19 death in 2020. This is exacerbated in cancers with high mortality rates. For instance, even if pancreatic cancer had in fact doubled the risk of dying of COVID-19 (IFR = 4.2% vs. 2.1%), we would only expect a rise in excess mortality around 0.4% during the pandemic (**Table 3**), while the 2019 baseline risk of death for pancreatic cancer itself is over 60% (**Figure 6**). On the other hand, for conditions with a lower baseline level mortality, such as diabetes, we expect substantial COVID-19 driven elevations in mortality.

Our analysis revealed interesting differences between types of cancers. Both nationally and at the state-level, the observed excess mortality for non-hematological cancers was consistent with a 1-to 2-fold increase in COVID-19 mortality risk in patients with these types of cancer. Importantly, our analysis ignores any behavioral effects associated with the pandemic. It is conceivable that cancer patients may have shielded themselves from COVID-19 more than the average person in 2020. Our projections assume an average risk of infection for a typical individual over 65 years as there is no serologic data on infection attack rates for specific clinical population subgroups (of any age). If shielding from exposure to SARS-CoV-2 was high among cancer patients, our projections of cancer excess mortality during the pandemic would be inflated. In other words, if shielding was particularly pronounced, cancer may conceivably be a higher risk factor than shown here. Retrospective serologic analysis of banked sera from the first year of the pandemic, broken down by underlying comorbidities, may shed light on whether infection risk may have varied by chronic condition.

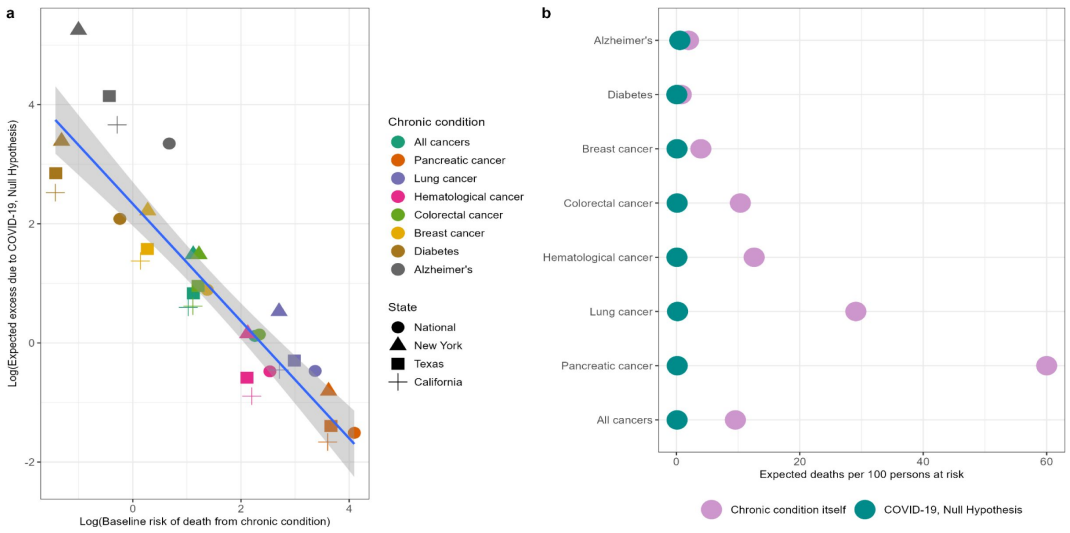


Figure 6.

Illustration of competing risks

Based on our demographic model, we expect a small increase in cancer mortality relative to diabetes and Alzheimer's due to the higher competing risk of death from cancer compared to COVID-19. Panel a) shows the log of the baseline mortality rate (based on observed mortality in 2019) from each condition on the x-axis and the log of the expected excess mortality (elevation over baseline) on the y-axis. Chronic conditions are shown in colors while states are shown in different shapes. Pancreatic cancer, the deadliest cancer considered, is on the bottom right (highest baseline mortality, lowest expected excess) while diabetes and Alzheimer's are on the top left (lowest baseline mortality, highest expected excess). Panel b) shows the baseline number of deaths per 100 persons at risk for each condition expected from March - December 2020 (based on deaths over this same period in 2019, purple dots) compared to the expected number of deaths due to COVID-19 under the null hypothesis (green dots). The null hypothesis stipulates that there is no biological association between any of these chronic diseases and COVID-19. For diabetes and Alzheimer's, the risks of baseline death and COVID-19 death are similar, while risk of death from cancer out competes risk of COVID-19 death for all types of cancer.

State-level mortality patterns can potentially provide complementary insights on the question of shielding. Because NY state experienced the earliest and most intense COVID-19 wave of the US, with over 20% of the population infected in Spring 2020 (Stadlbauer et al., 2021), and because social distancing did not come into effect until March 2020, shielding would have had a more limited impact there than in other states. Thus, a biological relationship between cancer and COVID-19 would have been most dramatic in NY in spring 2020. Indeed, cancer excess mortality was exacerbated in NY, including an 9-16% increase in colorectal and breast cancer mortality, consistent with a 2-fold increase in COVID-19 mortality risk from these cancers, and a 10% increase in hematological cancers, consistent with a 5-fold increase in COVID-19 mortality risk. In NY, the absence of excess mortality in lethal cancers, such as pancreatic and lung cancers (1% over baseline) are, as discussed above, still consistent with what would be expected under a high competing risk situation.

We used diabetes and Alzheimer as positive controls for a known biological association between COVID-19 and chronic conditions. Diabetes stood out in our analyses with the highest absolute and relative increases in excess mortality during the pandemic. The magnitude of the mortality increases, both nationally and at the state level, were close to what would be expected if diabetes increased COVID-19 mortality by 5-fold. Many studies have shown that diabetes increases the risk of COVID-19 mortality, with an effect size around 2 (Williamson et al. 2020; Huang et al. 2020; Kastora et al. 2022). Impaired immune function and chronic inflammation have been identified as mechanisms driving poor outcomes for diabetes patients (Figueroa-Pizano et al. 2021). The discrepancy between the observed excess and our expectations may come down to uncertainty in the SARS-CoV-2 infection rates assumed in our demographic model. The population living with diabetes is slightly younger than that of the other conditions (mean age, 58-60 years), while we used serologic infection rates reported for individuals over 65 years in our main analysis. The SARS-CoV-2 attack-rate among those 50-64 years was 10.1% at the end of 2020, compared to 6.3% in individuals over 65 (Centers for Disease Control and Prevention, 2023). A sensitivity analysis using this higher attack rate in our demographic model lends more support to the hypothesis that diabetes increases COVID-19 mortality by 2-fold, rather than 5-fold as found in our main analysis.

Our second positive control, Alzheimer's, revealed surprising results. Although we observed significant excess mortality in MC Alzheimer's data, it was still less than expected under the null hypothesis that Alzheimer's was not a risk factor for COVID-19 mortality. This is unexpected in light of several observational studies that have shown Alzheimer's to be a risk factor (Tahira et al., 2021; Wang et al., 2021; Zhang et al., 2021). As with cancer and diabetes, there is uncertainty in the SARS-CoV-2 infection rates used in the demographic model, due to the potential effect of shielding and the age-specific SARS-CoV-2 infection risk of the Alzheimer's population. We estimated that the average age of the population living with Alzheimer's disease was 80-81 years, and the infection rates for the general population over 65 years may not accurately reflect exposure in this subpopulation. Decreasing the attack rates by 20-30% (down to 4.4-5.0%) puts the observed estimates in the range of the expectations under the null hypothesis. Overall, given uncertainty in SARS-CoV-2 attack-rates and the age and size of the population-at-risk for all studied conditions, our demographic model projections are not an exact tool to titrate excess mortality nor the relative risk associated with each condition. Our model merely serves as an illustration of the role of demography and competing risks.

Most vital statistics studies of the COVID-19 pandemic have relied on underlying cause-specific deaths, which are prone to changes in coding practices. Our initial hypothesis going into this work was that coding changes associated with a better recognition of the impact of SARS-CoV-2 led to an underestimation of excess mortality from cancer, affecting our perception of the relationship between cancer and COVID-19. We certainly found an effect of coding changes, where for instance a drop in excess mortality in underlying cancer deaths turned into an increase in multiple-cause (any-listed) cancer deaths, particularly in the first COVID-19 pandemic wave. A similar observation was made by Fedeli et al. The impact of coding changes was also seen in mortality from other

chronic conditions but was particularly important for cancer. Yet both the absolute and relative excess mortality elevation remained modest for cancer, even after adjustment for coding changes, highlighting the importance of additional mechanisms such as competing mortality risks between COVID-19 and cancer.

An interesting hypothesis was put forward 20 years ago proposing that immunosuppression from cancer may explain the lack of excess cancer mortality in the 1968 influenza pandemic – the immune incompetence rescue hypothesis (Reichert et al 2004). This hypothesis contends that it is a detrimental immune response that leads to influenza death. A similar hypothesis was put forward to explain the extreme mortality in young healthy adults in the 1918 pandemic (Short et al., 2018). However, observational studies have found that patients with hematological cancers have twice the risk of dying compared to patients without cancer, likely due to the immunosuppression associated with their malignancy and treatment (X. Han et al., 2022; Starkey et al., 2023; Williamson et al., 2020). Under the immune incompetence rescue hypothesis, hematological cancers would be expected to have the lowest excess mortality of all types of cancers. Our excess mortality analysis reveals instead that hematologic cancers were the most impacted by the pandemic, relative to other types of cancer, with observed mortality patterns consistent with a 5-fold increase in risk of COVID-19 death in patients with hematological cancers. Overall, we do not find any support for the immune competence rescue hypothesis.

Our study is subject to limitations. First, we did not study the potential long-term consequences of the pandemic on cancer care, which may have resulted in avoidance of the health care system for diagnosis or treatment. We did not see any delayed pandemic effect on mortality from pancreatic cancer, which may have manifested in 2020 given the very low survival rate of this cancer (Lemanska et al., 2023), but we cannot rule out longer-term effects on breast or colorectal cancers that would not be seen until 2021 or later (Doan et al., 2023; Han et al., 2023; Haribhai et al., 2023; R. Lee et al., 2023; Nascimento de Lima et al., 2023; Nickson et al., 2023; Nonboe et al., 2023; Tope et al., 2023). Interestingly, in the US, all-cause underlying cancer mortality rates do not appear to rise between 2020 and 2023 (**Appendix 1 - Figure 10**), but data prior to the pandemic show a rise in cancer incidence, largely driven by increasing cancer rates in younger adults (Zhao et al. 2023; Siegel et al. 2024). Additional years of data will be important to evaluate the long-term impacts of the COVID-19 pandemic and these changing demographics on cancer mortality rates. Additional years of data will also be important for assessing the impact of vaccination on the relationship between cancer and COVID-19; there is evidence that vaccines may be less immunogenic in patients with cancer compared to those without (Seneviratne et al., 2022). Another limitation of our study is the reliance on mortality as an outcome, and not the risk of COVID-19-related hospitalization and morbidity, and Long COVID in cancer patients. A small US study reported that 60% of cancer patients suffered Long COVID symptoms (Dagher et al., 2023). Future analyses using hospitalization data and electronic medical records may provide additional insights on how different cancer stages or other comorbidities may contribute to increased risk of severe COVID-19 outcomes. Lastly, a few methodological limitations are worth raising. Though it was important to assess excess mortality in state level data because of asynchrony in pandemic waves, confidence intervals in state-level estimates were large, particularly for specific types of cancers, affecting significance levels. Additional methodological limitations relate to our demographic model, especially as regards assumptions about SARS-CoV-2 infection rates in populations of different ages and with different chronic conditions. Importantly, our conclusions regarding the importance of competing risks are robust to these assumptions. Lastly, our study is a time-trend analysis and – like cohort and case-control studies – correlation does not necessarily imply causation. However, the intensity and brevity of COVID-19 pandemic waves in space and time lends support to our analyses.

Conclusion

Our detailed excess mortality study considered six cancer types and found that there is at most a modest elevation in cancer mortality during the COVID-19 pandemic in the US. Our results demonstrate the importance of considering multiple-causes-of-death records to accurately reflect changes in coding practices associated with the emergence of a new pathogen. In contrast to earlier studies, we propose that lack of excess cancer mortality during the COVID-19 pandemic reflects the competing mortality risk from cancer (especially for deadly types like pancreatic and lung cancers) itself rather than protection conferred from immunosuppression. We note the more pronounced elevation in mortality from hematological cancers during the pandemic, compared to other cancers and to expectations from a demographic model, which aligns with a particular group of cancer patients singled out in several cohort studies. Future research on the relationship between COVID-19 and cancer should concentrate on additional outcomes, such as excess hospitalizations, Long COVID, changes in screening practices during COVID-19, and longer-term patterns in cancer mortality.

Materials and Methods

Data sources

US National vital statistics

We obtained individual ICD-10 coded death certificate data with exact date of death from the United States for the period January 1, 2014, to December 31, 2020. Each death certificate has one underlying cause (UC) of death, defined as the disease or injury that initiated the train of events leading directly to death, and up to twenty causes of death in total, referred to here as multiple cause mortality (MC). We considered 10 conditions, including diabetes, Alzheimer's disease, ischemic heart disease (IHD), kidney disease, and 6 types of cancer (all-cause cancer, colorectal, breast, pancreatic, lung, and hematological; see **Table 1** and **Appendix 1 – Table 1** for a list of disease codes). We chose these types of cancer to illustrate conditions for which the 5-year survival rate is low (13% and 25%, respectively, for pancreatic and lung cancers) and high (65% and 91%, respectively, for colorectal and breast cancers) (National Cancer Institute, n.d.). Hematological cancer (67% 5-year survival) was included because it was singled out as a risk factor by previous studies. We compiled time series by week, geography (aggregated National, NY, TX, and CA) and cause of death, separately for underlying and multiple cause mortality.

To observe longer-term trends in later years of the COVID-19 pandemic, we downloaded aggregated weekly-level data from 2021 to 2023 for all-cause cancer, diabetes, and Alzheimer's disease from CDC Wonder.

Estimated populations living with each chronic condition

We estimated the size of the population-at-risk for all-cause and specific cancers using the 5-year limited duration prevalence estimates provided by the U.S. Cancer Statistics webpage (U.S. Cancer Statistics Working Group …). Estimates for diabetes were drawn from CDC's Behavioral Risk Factor Surveillance System Chronic Disease Indicators (CDC, Division of Population Health). Estimates for Alzheimer's disease were taken from publications from the Alzheimer's Association (Rajan et al. 2021; Dhana et al. 2023).

For each condition, age-specific prevalence data were tabulated for the US and for each state separately. For cancer, age-level data were only available at the national level so these age-specific prevalence estimates were applied to the populations for each of the three states considered (NY, CA, TX). Age-level data were provided for all ages for cancer (<20 years, 20-80 years in 10-year groupings, ≥80 years), for adults ≥ 18 for diabetes (18-44years, 45-64years, ≥65 years) and for adults ≥65 for Alzheimer's disease (65-74years, 75-84yrs, ≥85years). A weighted mean age for the population-at-risk for each condition was calculated using the mid-point for each age group.

Other data sources

To compare vital statistics patterns with COVID-19 surveillance data, we accessed national and state counts of laboratory-confirmed COVID-19 cases in 2020, from the CDC (Centers for Disease Control and Prevention, 2022).

To clarify the expected role of COVID-19 on excess mortality, we compiled data on the proportion of the population with serologic evidence of SARS-CoV-2 infection from the CDC dashboard (Centers for Disease Control and Prevention, 2023). We further compiled data on estimated age-specific infection-fatality ratios from COVID-19, provided by single year of age (COVID-19 Forecasting Team, 2022).

Statistical approach

Weekly excess mortality models

Similar to other influenza and COVID-19 excess mortality studies (Islam et al., 2021; Karlinsky and Kobak, 2021; W.-E. Lee et al., 2023; Msemburi et al., 2023), we established a predicted baseline of expected mortality for each time series, and computed the excess mortality as the excess in observed deaths over this baseline. To establish baselines for each disease nationally and in each state, we applied negative binomial regression models to weekly mortality counts for each cause of death, smoothed with a 5-week moving average and rounded to the nearest integer. Models included harmonic terms for seasonality, time trends, and an offset for population size. For each condition and location, we used Akaike Information Criterion (AIC) to select between three models with different time trends (see Appendix 1 - Supplemental Methods, **Appendix 1- Figure 11**, for the final model selection for each location and condition), following:

Model 1:

Weekly_mortality = t + cos(2πt/52.17) + sin(2πt/52.17) + offset(log(population)), where t represents week.

Model 2:

Weekly_mortality = t + t^2 + cos(2πt/52.17) + sin(2πt/52.17) + offset(log(population)), where t represents week.

Model 3:

Weekly_mortality = t + t2 + t3 + cos(2πt/52.17) + sin(2πt/52.17) + offset(log(population)), where t represents week.

We fitted national and state-level models for each mortality outcome from January 19, 2014, to March 1, 2020, and projected the baseline forward until December 6, 2020, the last complete week of smoothed mortality data. Models were fitted using the MASS package in R version 4.3.

Using COVID-19 coded death certificates from March 1, 2020, to December 6, 2020, we established the timing of each pandemic wave from trough to trough. We found that nationally, the first wave occurred from March 1, 2020, to June 27, 2020; the second wave from June 28, 2020, to October 3, 2020, and the third from October 4, 2020, to December 6, 2020 (the 3rd wave was not completed by the last week of available smoothed data on December 6, 2020). For NY, the pandemic pattern was characterized by an intense first wave in Spring 2020, while TX had its major wave in summer 2020 and CA in late 2020. Comparison of mortality patterns from these three states provides an opportunity to separate the effect of SARS-CoV-2 infection from that of behavioral changes later in the pandemic. For instance, the effects of healthcare avoidance would predominate in CA or TX in Spring 2020, as there was little SARS-CoV-2 activity but much media attention on COVID-19, with cancer patients potentially avoiding medical care out of fear of getting infected. In contrast, risk of infection would dominate in NY in Spring 2020, and behavioral factors may only play a role as SARS-CoV-2 awareness increased and the wave was brought under control by social distancing.

We estimated weekly excess mortality by subtracting the predicted baseline from the observed mortality. We summed weekly estimates to calculate excess mortality for the full pandemic period and for each of the 3 waves within the first year of the pandemic. In addition to estimating the absolute effects of the pandemic on mortality, we also calculated relative effects by dividing excess deaths in each diagnosis group by the model baseline. Confidence intervals on excess mortality estimates were calculated by resampling the estimated model coefficients 10,000 times using a multivariate normal distribution and accounting for negative binomial errors in weekly mortality counts.

We used Pearson correlation to test synchronicity patterns in weekly excess mortality from different cancers and chronic conditions to underlying COVID-19 deaths. Correlation analysis assumes a direct and immediate effect of COVID-19 on cancer mortality. We also investigated the possibility of delayed effects or harvesting by inspecting the time series for evidence of such effects and by comparing total excess deaths for distinct pandemic waves and the whole of 2020.

Projections of excess mortality under the null hypothesis of no specific COVID-19 mortality risk of each condition

To further test the impact of age on the association between chronic conditions and COVID-19 and clarify the additional risk due to each chronic condition, we projected the number of COVID-19 deaths under the null hypothesis that demographic characteristics alone (size, age, and baseline mortality risk for each condition) are driving excess mortality, and that there is no interaction between the condition and SARS-CoV-2 infection. Excess mortality projections were then compared with observed excess mortality. We only used MC deaths for this approach to account for the possibility that some individuals may suffer from multiple conditions. For example, an estimated 11.5% of US adults with type 2 diabetes also have a history of cancer (Yeh et al., 2018).

We first calculated the number of expected COVID-19 infections among persons living with a certain chronic condition, by multiplying the estimated number of individuals living with the condition by the reported SARS-CoV-2 seroprevalence among individuals ≥65 years at specific time points during 2020. For the national data and California, we used results from the survey conducted from November 23 - December 12, 2020. For New York we used estimates from the survey conducted from July 27 - August 13, 2020 (the earliest data available). And for Texas we used the survey conducted from October 5-19, 2020 (following the large summer wave). (Centers for Disease Control and Prevention, 2023). We then multiplied this by the COVID-19 IFR based on the estimated mean age of individuals living with the condition (COVID-19 Forecasting Team, 2022) to arrive at the projected number of COVID-19-related excess deaths for a particular condition during 2020. We put uncertainty intervals around these estimates using the lower and upper bounds from the estimated attack-rates and COVID-19 IFRs.

To obtain a relative metric of expected COVID-19 burden, we divided projected COVID-19 excess deaths by total deaths in each diagnosis group in the 2019 baseline period (March to December 2019, for the national data. For the states we used the months in 2019 corresponding to their large waves in 2020), resulting in an expected percentage elevation over baseline in 2020. We compared this null expectation to the observed percentage elevation over baseline from our excess mortality models. We also generated the expected number of excess deaths under alternative hypotheses where each condition is associated with a 2- or 5-fold increased risk of COVID-19 related death given infection (i.e., the baseline age-adjusted infection fatality ratio used in the null hypothesis was increased 2- or 5-fold).

The equation for the expected percent increase in excess mortality over baseline deaths under the null hypothesis, for a specific risk condition (cancer, diabetes., Alzheimer) and time-period, can be written as:

Expected percent increase in excess mortality for a chronic condition and time period = (size of population-at-risk for the condition * SARS-CoV-2 infection rate for the period * age-specific IFR) / baseline mortality for the condition in comparable period in 2019

The expected mortality increases under the alternative hypothesis of a 2- or 5-fold increased risk of COVID-19 death from the condition under study is modeled by multiplying the right-hand side of the above equation by the increased risk (i.e., we assume that presence of the underlying condition will increase the IFR by 2- or 5-fold compared to the IFR for the general population).

Acknowledgements

This paper is dedicated to our colleague Robert J Taylor who succumbed to cancer in 2022 and who wanted to know if a cancer diagnosis was a COVID-19 mortality risk factor.

Additional Information

Funding

LS acknowledges funding from the Carlsberg Foundation, grant number CF20-0046. LS and CLH acknowledge funding from Danish National Research Foundation (grant number DNRF170) for the PandemiX Center of Excellence. CLH has received contract-based hourly consulting fees from Sanofi outside of the submitted work.

Author contributions

Chelsea Hansen, Data curation, Formal analysis, Visualization, Methodology, Writing – original draft, Writing – review and editing; Cécile Viboud, Data curation, Formal analysis, Visualization, Methodology, Writing – original draft, Writing – review and editing; Lone Simonsen, Conceptualization, Data curation, Formal analysis, Visualization, Methodology, Writing – original draft, Writing – review and editing

Data availability

Individual-level mortality data were obtained from the National Center for Healthcare Statistics. These data are not publicly available due to privacy concerns, but descriptive characteristics have been summarized in **Table 1** and **Appendix - Table 1**. The excess mortality models in this

paper use mortality data aggregated by week and US state. These data, along with the model code, have been posted to the following public GitHub repository: *https://github.com/chelsea-hansen/Disentangling-the-relationship-between-cancer-mortality-and-COVID-19*

Additional weekly, aggregated mortality data are publicly available through CDC Wonder.

Data used for the demographic model, along with the code have also been posted to the GitHub repository.

Weekly, state-level data on recorded COVID-19 cases and deaths are publicly available. Data were downloaded from here: *https://data.cdc.gov/Case-Surveillance/Weekly-United-States-COVID-19-Cases-and-Deaths-by-/pwn4-m3yp* and have also been posted as a .csv file to the GitHub repository referenced above.

Disclaimer

This article represents the views of the authors and not necessarily those of the National Institutes of Health or the US government.

Appendix 1

Supplemental Methods

Model selection and cross validation

Time series models included harmonic terms for seasonality, time trends, and an offset for population size. For each condition and location, we used Akaike Information Criterion (AIC) to select between three models with different time trends. The starting model (Model 1) included only a linear time trend. We then tested this against a model with linear and quadratic time trends (Model 2). If the AIC of Model 2 was not 2 less than Model 1, Model 1 was used as the final model. If the AIC of Model 2 was 2 less than Model 1, then Model 2 was tested against a model with linear, quadratic, and cubic time trends (Model 3). If the AIC of Model 3 was not 2 less than Model 2, then Model 2 was taken as the final model. If the AIC of Model 3 was 2 less than Model 2, Model 3 was taken as the final model. The final model for each condition and location was then applied to the data from 2014-2018 only and used to predict the 2019 data. The coverage probability was calculated as the proportion of weeks of observed data in 2019 that fell within the 95% prediction interval of the time series model. The final model selected for each condition and location is provided in the appendix (**Appendix 1-Figure 11**).

Characteristics of cancer, diabetes, and Alzheimer's deaths in the pre-pandemic period

For each chronic condition studied (cancer, diabetes, Alzheimer's), we assessed potential changes in the characteristics of deaths during the pandemic period that are unrelated to timing but may signal an association with COVID-19. For instance, age is known to be a major risk factor for COVID-19 mortality. For each chronic condition, we computed the average age-at-death in the pre-pandemic year 2019, and compared this to the average age-at-death in 2020. The second potential confounder is living arrangement, as individuals living in nursing homes may be at increased risk of exposure (and death) to COVID-19 due to mixing, even if their underlying condition is not per se a risk factor. To test this hypothesis, we also compared the proportion of individuals in each disease group who died in nursing homes in 2019 and 2020. And finally, to illustrate the impact of coding practices we compared ICD-10 letter categories between 2020 and 2019 for the underlying cause of death when cancer or diabetes are included on the death certificate but are not listed as

the underlying cause of death (**Appendix 1 - Figure 9**). For 2020, we further compared death certificates listing both COVID-19 and cancer to those listing both COVID-19 and diabetes. For all comparisons between 2019 and 2020 data are limited to March to December to isolate the pandemic period.

Supplemental tables and figures

Appendix 1 - Table 1.

Diagnosis groups and corresponding ICD-10 codes, number of underlying and multiple cause deaths, mean age in years at time of death, the percentage of deaths occurring at home, and the percentage of deaths occurring in nursing homes for 2019 and 2020

			Underlying Cause				Multiple Cause			
Year	Diagnosis group	ICD-10 codes	No. Deaths	Mean age, years (IQR)	%Home/ER	%Nursing Home	No. Deaths	Mean age, years (IQR)	%Home/ER	%Nursing Home
2019	Cancer	C00-C99	493,397	72 (64-81)	45	12	546,453	72 (64-82)	44	13
	Pancreatic Cancer	C25	37,864	72 (64-80)	51	9	39,798	72 (64-80)	50	9
	Lung Cancer	C34	114,552	72 (65-80)	45	12	123,622	72 (65-80)	44	12
	Colorectal Cancer	C18-C20	42,484	71 (61-82)	46	13	49,053	72 (62-83)	45	14
	Breast Cancer	C50	35,115	69 (59-81)	44	13	43,519	71 (61-83)	43	15
	Hematological Cancer	C81-C96	47,174	74 (67-84)	35	11	57,892	74 (67-84)	35	12
	Diabetes	E10-E14	70,763	72 (63-82)	53	17	229,326	74 (65-84)	46	19
	Alzheimer's	G30	98,675	87 (82-92)	29	50	118,993	87 (82-92)	29	48
	Ischemic Heart Disease	I20-I25	292,659	77 (67-88)	50	18	440,225	77 (68-87)	47	18
	Kidney Disease	N00-07, 17-19,25-28	46,120	76 (68-87)	25	18	189,938	76 (67-87)	20	15
2020	Cancer	C00-C99	513,275	72 (64-81)	55	8	586,503	72 (64-82)	52	9
	Pancreatic Cancer	C25	39,893	72 (65-80)	61	6	42,383	72 (65-80)	60	6
	Lung Cancer	C34	115,554	72 (65-80)	54	8	127,671	72 (65-80)	53	8
	Colorectal Cancer	C18-C20	43,990	71 (61-82)	56	9	52,319	72 (62-83)	53	10
	Breast Cancer	C50	36,296	70 (60-81)	54	10	47,094	72 (62-83)	51	12
	Hematological Cancer	C81-C96	49,161	74 (67-84)	46	8	64,840	74 (68-84)	43	9
	Diabetes	E10-E14	88,124	71 (62-82)	58	15	343,061	73 (65-83)	45	16
	Alzheimer's	G30	115,256	86 (82-92)	33	46	151,206	86 (82-92)	31	47
	Ischemic Heart Disease	I20-I25	327,854	76 (67-88)	54	16	533,204	77 (68-87)	49	16
	Kidney Disease	N00-07, 17-19,25-28	49,796	76 (68-87)	30	15	255,708	75 (67-86)	21	12

Appendix 1 - Table 2.

Supplemental Table 2. Estimated number of excess deaths and the percentage over baseline for each diagnosis group (National). Estimates are aggregated over all of 2020 and for each COVID-19 wave during 2020

		Multiple Cause		Underlying Cause	
Cause of death	Wave	Excess Deaths	% Over Baseline	Excess Deaths	% Over Baseline
Cancer	Overall	13601*	3.0	11	0.0
	1	79	0.0	-3917*	-2.0
	2	6519*	4.0	2662	2.0
	3	7003*	6.0	1266	1.0
Pancreatic Cancer	Overall	-25	-0.0	-282	-1.0
	1	-213	-1.0	-281	-2.0
	2	44	0.0	-30	-0.0
	3	144	1.0	29	0.0
Lung Cancer	Overall	1102	1.0	-814	-1.0
	1	-729	-1.0	-1221	-3.0
	2	784	2.0	249	1.0
	3	1047	4.0	158	1.0
Breast Cancer	Overall	838	2.0	-438	-1.0
	1	-66	-0.0	-415	-3.0
	2	437	3.0	81	1.0
	3	467	5.0	-105	-1.0
Colorectal Cancer	Overall	988	2.0	-168	-0.0
	1	-169	-1.0	-463	-3.0
	2	454	3.0	112	1.0
	3	703*	6.0	183	2.0
Hematological Cancers	Overall	3615*	7.0	111	0.0
	1	546	2.0	-447	-2.0
	2	1412*	8.0	412	3.0
	3	1657*	12.0	146	1.0
Diabetes	Overall	82318*	37.0	10784*	16.0
	1	25306*	25.0	2305*	7.0
	2	27534*	38.0	4330*	20.0
	3	29477*	56.0	4148*	26.0
Alzheimer's	Overall	21712*	19.0	8528*	9.0
	1	4763*	9.0	547	1.0
	2	8054*	22.0	4257*	14.0
	3	8894*	33.0	3724*	16.0
Ischemic Heart Disease	Overall	58793*	14.0	17194*	6.0
	1	12042*	6.0	862	1.0
	2	21783*	16.0	7912*	9.0
	3	24967*	25.0	8419*	13.0
Kidney Disease	Overall	41907*	22.0	785	2.0
	1	8182*	10.0	-1048	-5.0
	2	14767*	25.0	777	5.0
	3	18958*	44.0	1056*	10.0

*Confidence interval does not include zero

Appendix 1 - Table 3.

Supplemental Table 2. Estimated number of excess deaths and the percentage over baseline for each diagnosis group (New York). Estimates are aggregated over all of 2020 and for each COVID-19 wave during 2020

		Multiple Cause		Underlying Cause	
Cause of death	Wave	Excess Deaths	% Over Baseline	Excess Deaths	% Over Baseline
Cancer	Overall	1012	4.0	-557	-2.0
	1	747	6.0	-474	-5.0
	2	120	1.0	-6	-0.0
	3	144	2.0	-77	-1.0
Pancreatic Cancer	Overall	-29	-1.0	-58	-3.0
	1	8	1.0	-16	-2.0
	2	-1	-0.0	-9	-1.0
	3	-37	-6.0	-33	-6.0
Lung Cancer	Overall	47	1.0	-163	-3.0
	1	27	1.0	-143	-7.0
	2	23	1.0	16	1.0
	3	-3	-0.0	-36	-3.0
Breast Cancer	Overall	205	9.0	-46	-2.0
	1	151	16.0	-34	-4.0
	2	31	4.0	3	0.0
	3	23	4.0	-15	-3.0
Colorectal Cancer	Overall	189	8.0	42	2.0
	1	91	9.0	-16	-2.0
	2	40	5.0	26	4.0
	3	58	9.0	33	6.0
Hematological Cancers	Overall	156	5.0	-149	-6.0
	1	121	10.0	-107	-11.0
	2	1	0.0	-25	-3.0
	3	35	5.0	-18	-3.0
Diabetes	Overall	7240*	66.0	866*	26.0
	1	5945*	128.0	568*	40.0
	2	631*	18.0	121	11.0
	3	664*	24.0	177	21.0
Alzheimer's	Overall	884*	26.0	233	9.0
	1	734*	49.0	188	16.0
	2	1	0.0	1	0.0
	3	150	17.0	44	6.0
Ischemic Heart Disease	Overall	7118*	25.0	3756*	17.0
	1	6607*	54.0	4092*	44.0
	2	179	2.0	-184	-3.0
	3	331	5.0	-152	-3.0
Kidney Disease	Overall	2438*	34.0	51	3.0
	1	1946*	63.0	22	3.0
	2	144	6.0	-13	-2.0
	3	349*	19.0	42	8.0

*Confidence interval does not include zero

Appendix 1 - Table 4.

Supplemental Table 2. Estimated number of excess deaths and the percentage over baseline for each diagnosis group (Texas). Estimates are aggregated over all of 2020 and for each COVID-19 wave during 2020

		Multiple Cause		Underlying Cause	
Cause of death	Wave	Excess Deaths	% Over Baseline	Excess Deaths	% Over Baseline
Cancer	Overall	602	2.0	-130	-0.0
	1	-48	-0.0	-62	-0.0
	2	467	4.0	39	0.0
	3	183	2.0	-107	-1.0
Pancreatic Cancer	Overall	1	0.0	5	0.0
	1	-36	-3.0	-36	-4.0
	2	17	2.0	24	3.0
	3	20	3.0	17	3.0
Lung Cancer	Overall	176	2.0	108	2.0
	1	33	1.0	31	1.0
	2	60	2.0	27	1.0
	3	84	5.0	49	3.0
Breast Cancer	Overall	-19	-1.0	-131	-5.0
	1	-54	-4.0	-54	-6.0
	2	29	3.0	-25	-3.0
	3	6	1.0	-51	-8.0
Colorectal Cancer	Overall	-12	-0.0	-92	-3.0
	1	-33	-2.0	-49	-4.0
	2	4	0.0	-34	-3.0
	3	17	2.0	-10	-1.0
Hematological Cancers	Overall	194	5.0	-12	-0.0
	1	24	2.0	1	0.0
	2	136	11.0	21	2.0
	3	33	3.0	-34	-4.0
Diabetes	Overall	8902*	49.0	618	11.0
	1	1411*	19.0	61	3.0
	2	4612*	77.0	420*	23.0
	3	2879*	62.0	138	9.0
Alzheimer's	Overall	2242*	24.0	1184	15.0
	1	309	8.0	197	6.0
	2	1398*	45.0	805*	31.0
	3	536*	21.0	181	8.0
Ischemic Heart Disease	Overall	6018*	20.0	1700	9.0
	1	736	6.0	99	1.0
	2	3376*	34.0	1228*	19.0
	3	1905*	24.0	374	7.0
Kidney Disease	Overall	6724*	47.0	579	19.0
	1	886*	15.0	115	9.0
	2	3535*	76.0	285*	28.0
	3	2303*	66.0	179	23.0

*Confidence interval does not include zero

Appendix 1 - Table 5.

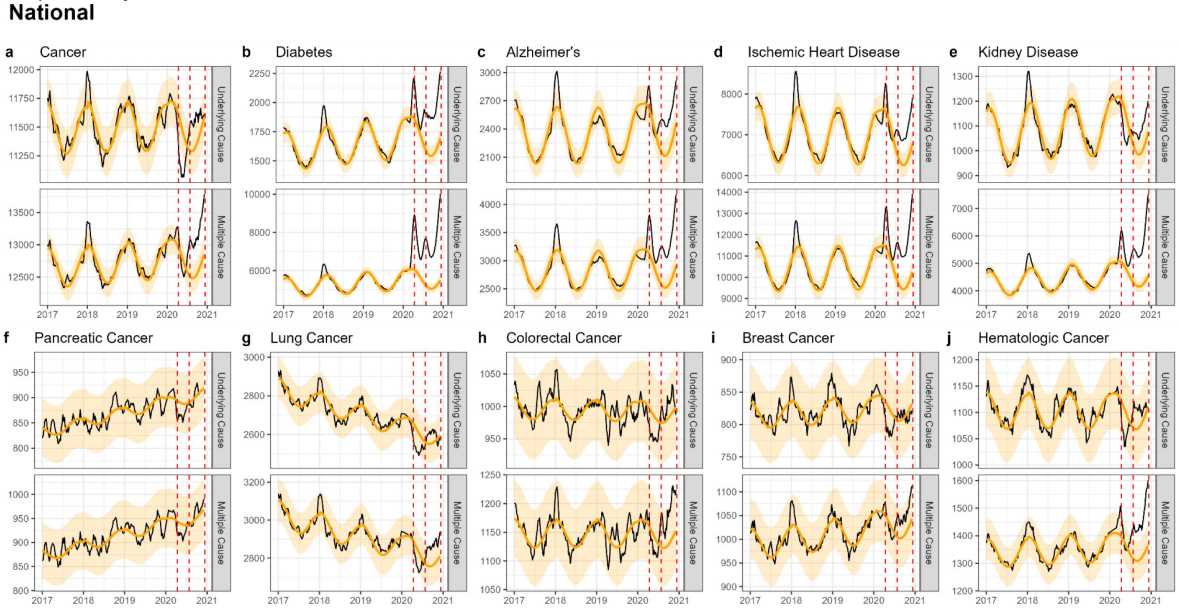
Supplemental Table 2. Estimated number of excess deaths and the percentage over baseline for each diagnosis group (California). Estimates are aggregated over all of 2020 and for each COVID-19 wave during 2020

		Multiple Cause		Underlying Cause	
Cause of death	Wave	Excess Deaths	% Over Baseline	Excess Deaths	% Over Baseline
Cancer	Overall	991	2.0	-29	-0.0
	1	-102	-1.0	-236	-1.0
	2	564	3.0	125	1.0
	3	529	4.0	82	1.0
Pancreatic Cancer	Overall	-97	-3.0	-126	-4.0
	1	-28	-2.0	-39	-3.0
	2	-69	-5.0	-70	-6.0
	3	0	0.0	-18	-2.0
Lung Cancer	Overall	-10	-0.0	-132	-2.0
	1	-82	-3.0	-96	-3.0
	2	18	1.0	-48	-2.0
	3	54	3.0	13	1.0
Breast Cancer	Overall	67	2.0	-22	-1.0
	1	-44	-3.0	-34	-3.0
	2	92	6.0	44	4.0
	3	20	2.0	-33	-4.0
Colorectal Cancer	Overall	100	2.0	20	1.0
	1	7	0.0	-4	-0.0
	2	66	4.0	25	2.0
	3	27	2.0	-1	-0.0
Hematological Cancers	Overall	279	5.0	52	1.0
	1	0	0.0	-33	-2.0
	2	164	9.0	64	4.0
	3	114	8.0	20	2.0
Diabetes	Overall	9163*	39.0	1408*	20.0
	1	1843*	18.0	213	7.0
	2	3846*	49.0	620*	27.0
	3	3474*	59.0	575*	33.0
Alzheimer's	Overall	2143*	14.0	594	5.0
	1	375	6.0	-76	-1.0
	2	1041*	20.0	410	9.0
	3	726*	18.0	259	8.0
Ischemic Heart Disease	Overall	5905*	16.0	2888*	11.0
	1	650	4.0	104	1.0
	2	2966*	24.0	1581*	19.0
	3	2289*	25.0	1204*	19.0
Kidney Disease	Overall	3858*	21.0	8	0.0
	1	301	4.0	-114	-8.0
	2	1967*	33.0	63	6.0
	3	1590*	36.0	59	7.0

*Confidence interval does not include zero

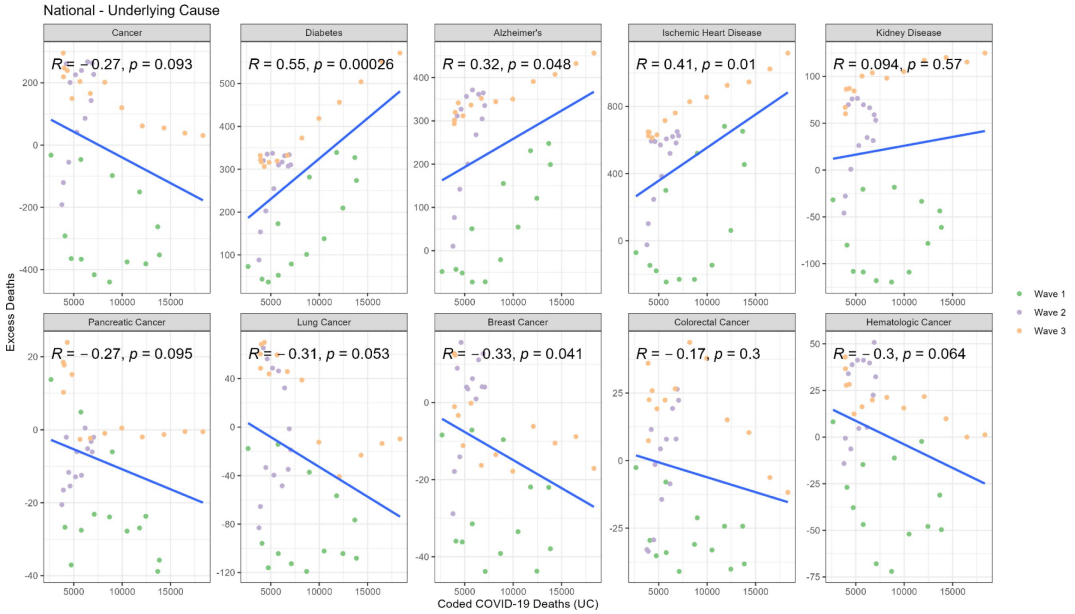
Appendix 1 - Figure 1.

National-level weekly observed and estimated baseline mortality for each diagnosis group as both the underlying cause or anywhere on the death certificate (multiple cause) from 2017 to 2020. Baselines during the pandemic are projected based on the previous years of data.



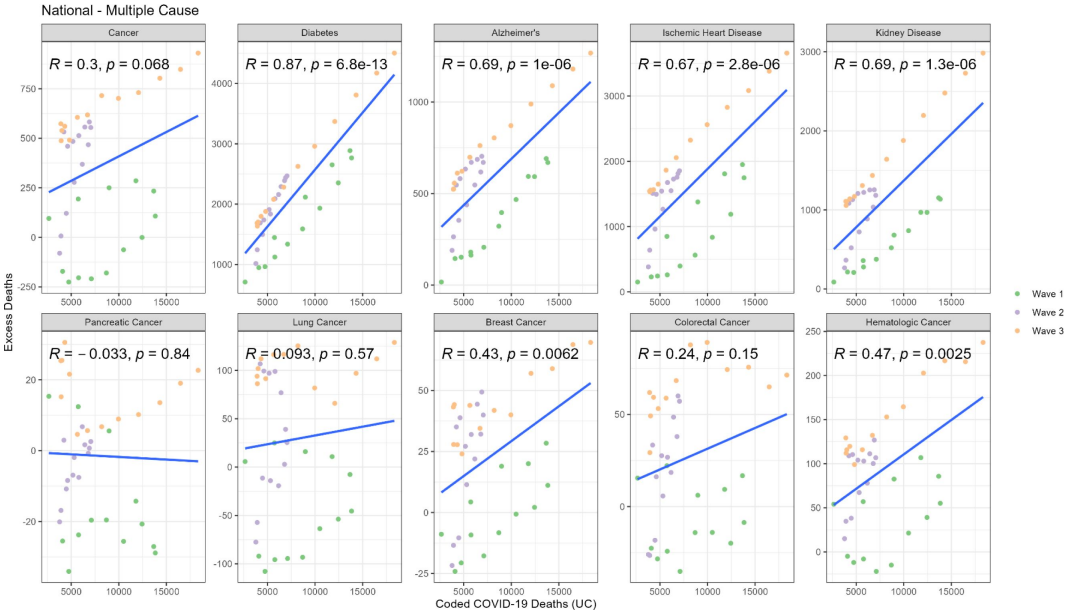
Appendix 1 - Figure 2.

Correlation between weekly number of COVID-19 coded deaths and excess underlying deaths for each diagnosis group (National).



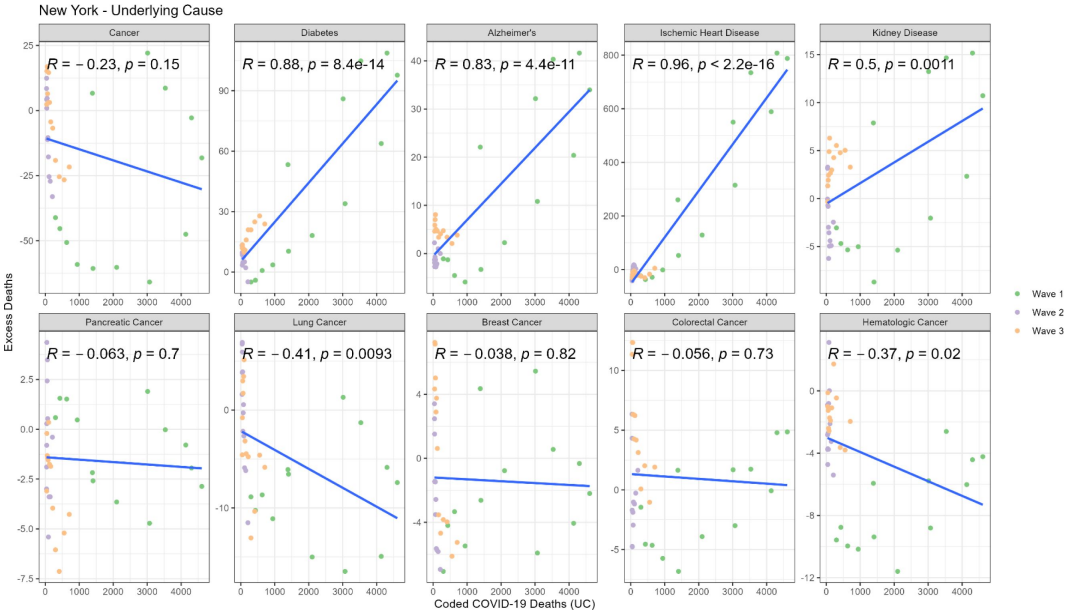
Appendix 1 - Figure 3.

Correlation between weekly number of COVID-19 coded deaths and excess multiple cause deaths for each diagnosis group (National).



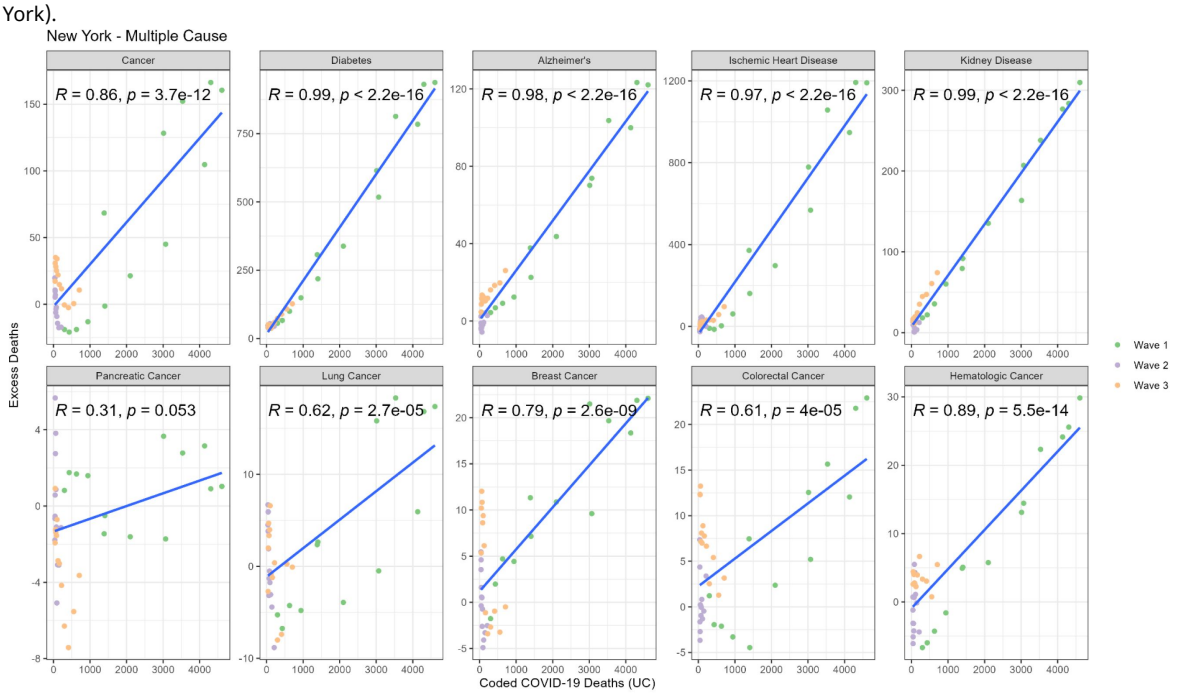
Appendix 1 - Figure 4.

Correlation between weekly number of COVID-19 coded deaths and excess underlying deaths for each diagnosis group (New York).



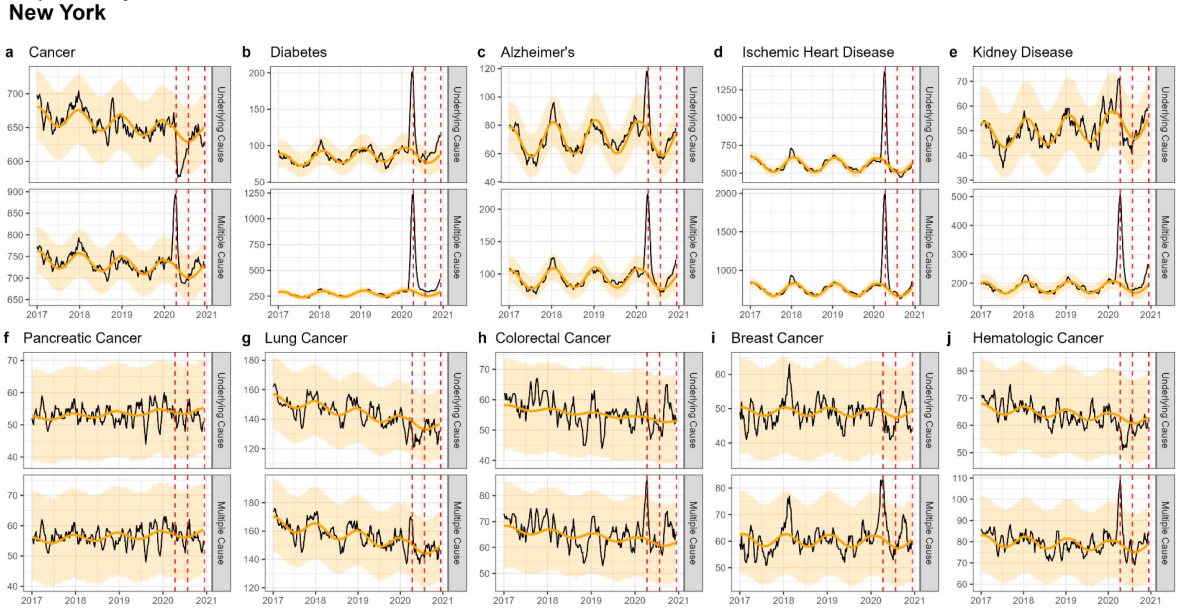
Appendix 1 - Figure 5.

Correlation between weekly number of COVID-19 coded deaths and excess underlying deaths for each diagnosis group (New York).



Appendix 1 - Figure 6.

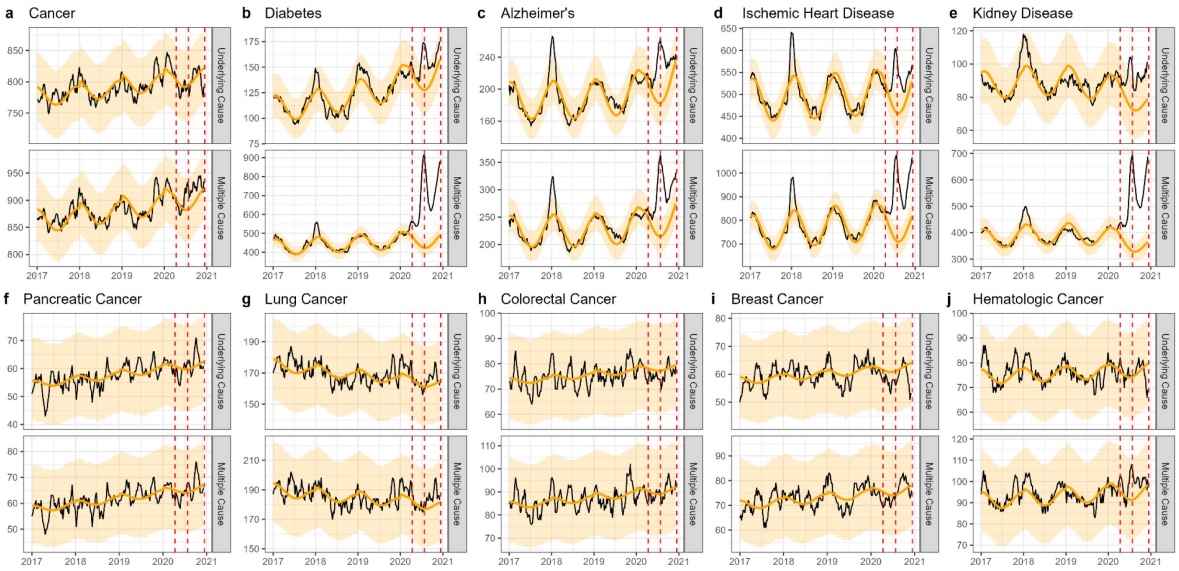
Weekly observed and estimated baseline mortality for each diagnosis group as both the underlying cause or anywhere on the death certificate (multiple cause) from 2017 to 2020 in New York. Baselines during the pandemic are projected based on the previous years of data.



Appendix 1 - Figure 7.

Weekly observed and estimated baseline mortality for each diagnosis group as both the underlying cause or anywhere on the death certificate (multiple cause) from 2017 to 2020 in Texas. Baselines during the pandemic are projected based on the previous years of data.

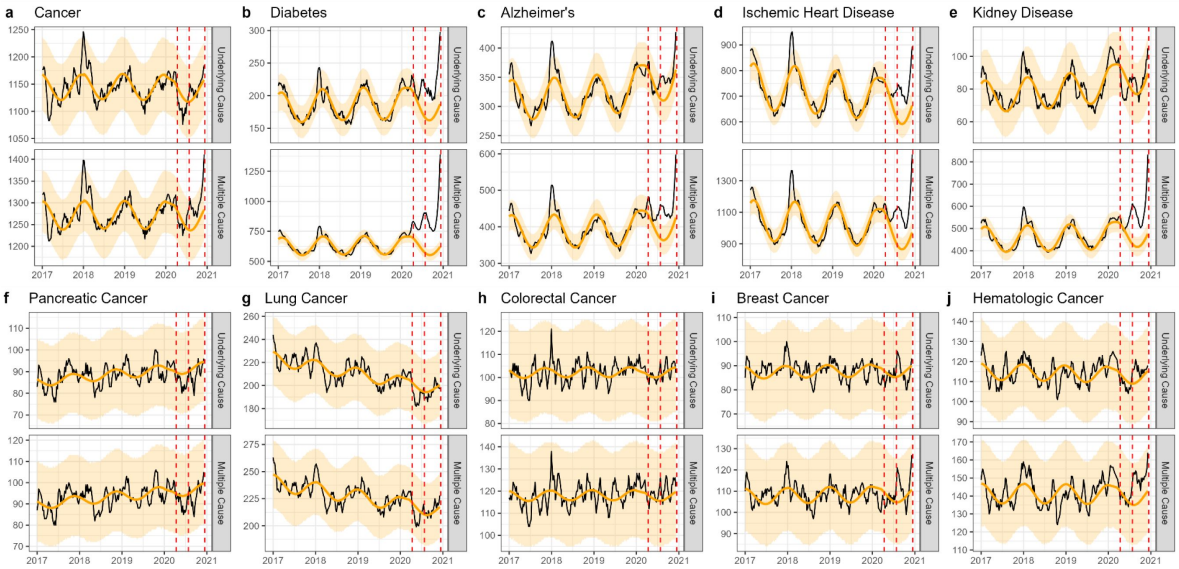
Texas

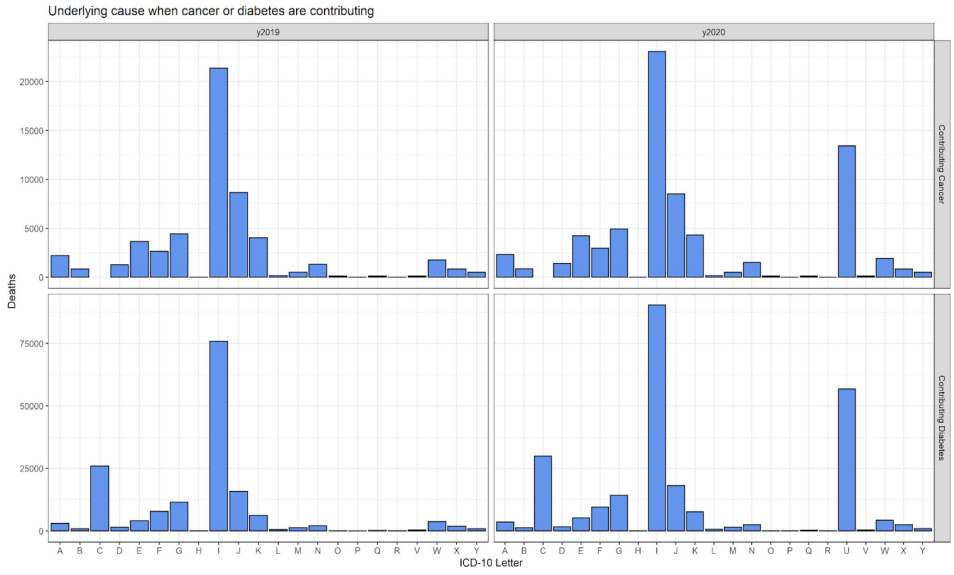


Appendix 1 - Figure 8.

Weekly observed and estimated baseline mortality for each diagnosis group as both the underlying cause or anywhere on the death certificate (multiple cause) from 2017 to 2020 in New York. Baselines during the pandemic are projected based on the previous years of data.

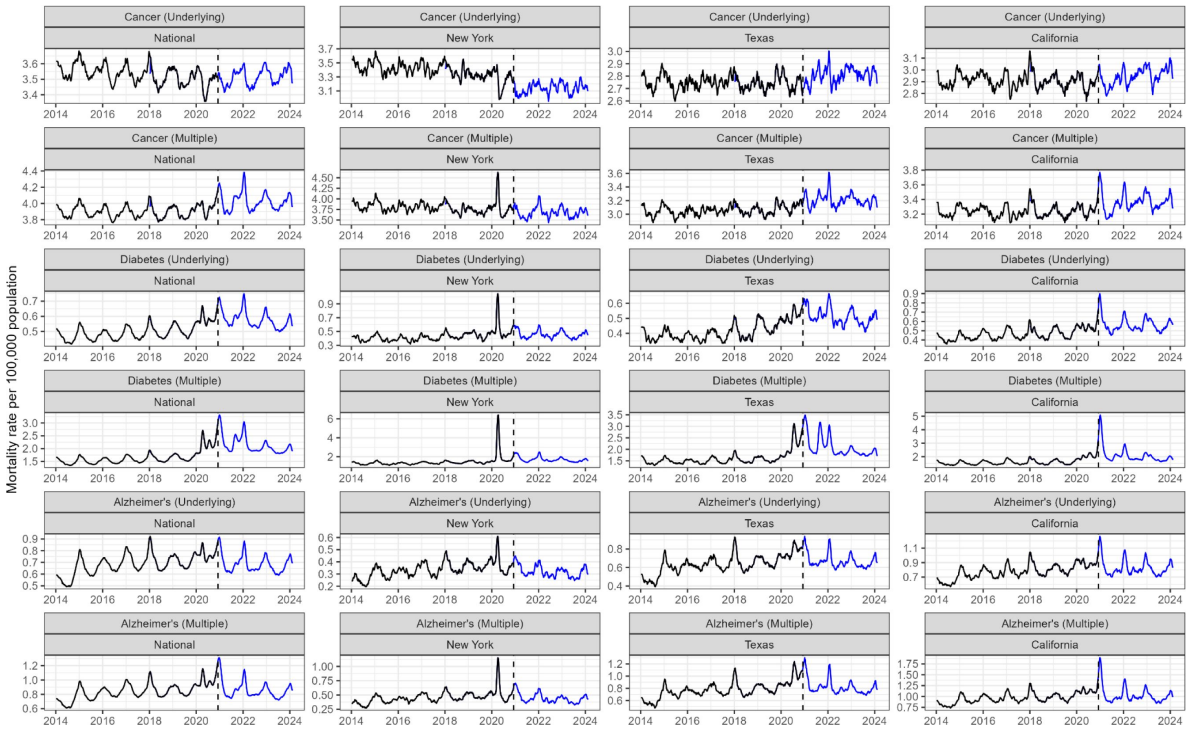
California





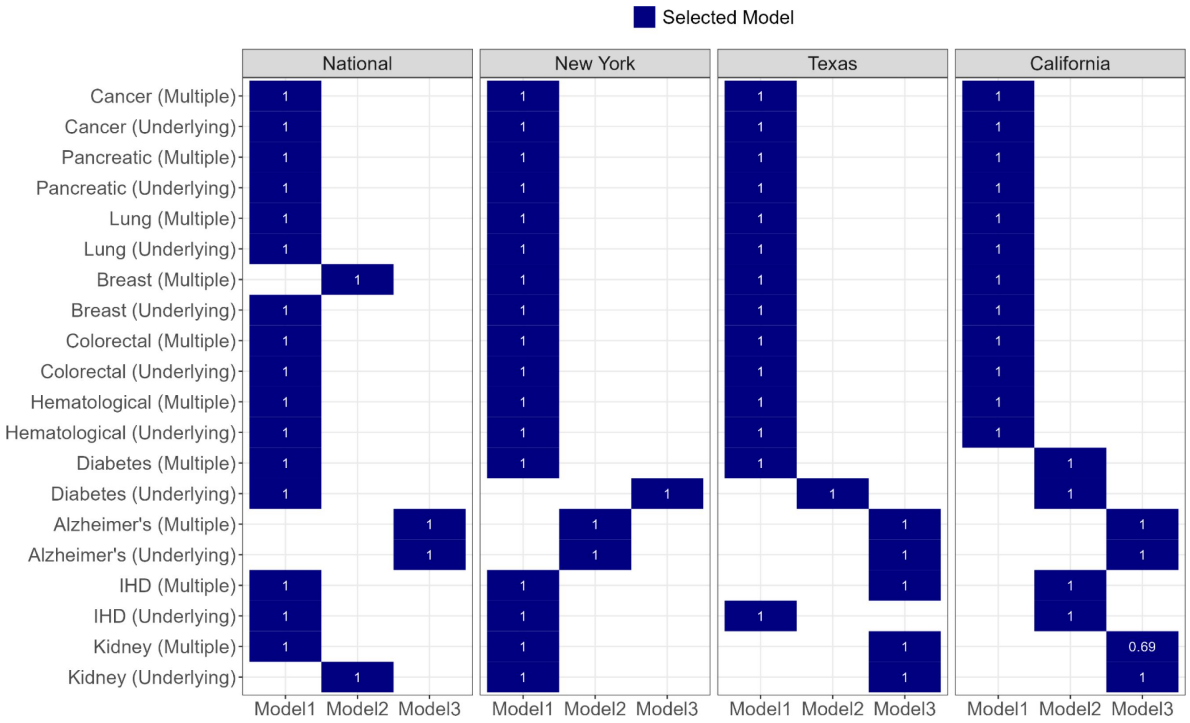
Appendix 1 - Figure 9.

Comparison of ICD-10 letter categories between 2020 and 2019 for the underlying cause of death when cancer or diabetes are included on the death certificate, but are not listed as the underlying cause of death. For both cancer and diabetes, I codes (diseases of the circulatory system) make up the majority of underlying deaths. The most notable difference between 2019 and 2020 is the increase in U codes, which includes COVID-19 (U071). In total there were 13,434 deaths ascribed to COVID-19 (UC deaths) among cancer MC deaths. COVID-19 was included in <3% of all cancer deaths and 17% of diabetes deaths. In both cases it was listed as the UC on the majority of death certificates where it was included (81% and 97% for cancer and diabetes, respectively).



Appendix 1 - Figure 10.

Post-2020 trends in cancer, diabetes, and Alzheimer's mortality. Aggregated weekly data was downloaded from CDC Wonder. Trends in cancer mortality rate appear stable in the national data and in Texas and California, but decreasing in New York. The diabetes mortality rate is higher post-2020 compared to earlier years across all states. Alzheimer's appears stable and slowly decreasing.



Appendix 1 - Figure 11.

For each condition three time series models with different time trends were considered (see Methods). The final model for each condition and location is indicated in blue. The final model was fit to 2014-2018 data only and used to predict the 2019 data. A coverage proportion (shown in white) was calculated as the proportion of observed 2019 data that fell within the projection intervals of the model. For all causes of death and states (except MC Kidney disease in California) the coverage proportion was 1, indicating that all data points fell within the prediction intervals.

References

Alicandro G, La Vecchia C, Islam N, Pizzato M. (2023) **A comprehensive analysis of all-cause and cause-specific excess deaths in 30 countries during 2020** *Eur J Epidemiol* https://doi.org/10.1007/s10654-023-01044-x

CDC, Division of Population Health **Chronic disease indicators**

Centers for Disease Control and Prevention (2023) **Centers for Disease Control and Prevention. 2023. COVID-19 serology surveillance. https://www.cdc.gov/coronavirus/2019-ncov/covid-data/serology.html**

Centers for Disease Control and Prevention (2022) **Weekly United States COVID-19 cases and deaths by state - ARCHIVED**

Chavez-MacGregor M, Lei X, Zhao H, Scheet P, Giordano SH (2022) **Evaluation of COVID-19 Mortality and Adverse Outcomes in US Patients With or Without Cancer** *JAMA Oncol* **8**:69–78

COVID-19 Forecasting Team (2022) **Variation in the COVID-19 infection-fatality ratio by age, time, and geography during the pre-vaccine era: a systematic analysis** *Lancet* **399**:1469–1488

Dagher H *et al.* (2023) **Long COVID in cancer patients: preponderance of symptoms in majority of patients over long time period** *Elife* **12** https://doi.org/10.7554/eLife.81182

Dhana K, Beck T, Desai P, Wilson RS, Evans DA, Rajan KB (2023) **Prevalence of Alzheimer's disease dementia in the 50 US states and 3142 counties: A population estimate using the 2020 bridged-race postcensal from the National Center for Health Statistics** *Alzheimers Dement* **19**:4388–4395

Di Felice G, Visci G, Teglia F, Angelini M, Boffetta P. (2022) **Effect of cancer on outcome of COVID-19 patients: a systematic review and meta-analysis of studies of unvaccinated patients** *Elife* **11** https://doi.org/10.7554/eLife.74634

Doan C, Li S, Goodwin JS (2023) **Breast and Lung Cancer Screening Among Medicare Enrollees During the COVID-19 Pandemic** *JAMA Netw Open* **6**

Fedeli U, Amidei CB, Han X, Jemal A (2024) **Changes in cancer-related mortality during the COVID-19 pandemic in the United States** *JNCI* **116**:167–169 https://doi.org/10.1093/jnci/djad191

Fernandes GA, Junior APN, Azevedo E Silva G, Feriani D, França E Silva ILA, Caruso P, Curado MP (2021) **. Excess mortality by specific causes of deaths in the city of São Paulo, Brazil, during the COVID-19 pandemic** *PLoS One* **16**

Gobiņa I, Avotiņš A, Kojalo U, Strēle I, Pildava S, Villeruša A, Briģis Ģ. (2022) **Excess mortality associated with the COVID-19 pandemic in Latvia: a population-level analysis of all-cause and noncommunicable disease deaths in 2020** *BMC Public Health* **22**

Grande E, Fedeli U, Pappagallo M, Crialesi R, Marchetti S, Minelli G, Iavarone I, Frova L, Onder G, Grippo F (2022) **Variation in Cause-Specific Mortality Rates in Italy during the First Wave of the COVID-19 Pandemic: A Study Based on Nationwide Data** *Int J Environ Res Public Health* **19** https://doi.org/10.3390/ijerph19020805

Han S, Zhuang Q, Chiang J, Tan SH, Chua GWY, Xie C, Chua MLK, Soon YY, Yang VS (2022) **Impact of cancer diagnoses on the outcomes of patients with COVID-19: a systematic review and meta-analysis** *BMJ Open* **12**

Han X, Hu X, Zhao J, Jemal A, Yabroff KR (2022) **Identification of Deaths Caused by Cancer and COVID-19 in the US During March to December 2020** *JAMA Oncol* **8**:1696–1698

Han X *et al.* (2023) **Changes in cancer diagnoses and stage distribution during the first year of the COVID-19 pandemic in the USA: a cross-sectional nationwide assessment** *Lancet Oncol* **24**:855–867

Haribhai S, Bhatia K, Shahmanesh M. (2023) **Global elective breast- and colorectal cancer surgery performance backlogs, attributable mortality and implemented health system responses during the COVID-19 pandemic: A scoping review** *PLOS Glob Public Health* **3**

Islam N *et al.* (2021) **Excess deaths associated with covid-19 pandemic in 2020: age and sex disaggregated time series analysis in 29 high income countries** *BMJ* **373**

Karlinsky A, Kobak D. (2021) **Tracking excess mortality across countries during the COVID-19 pandemic with the World Mortality Dataset** *Elife* **10** https://doi.org/10.7554/eLife.69336

Kontopantelis E *et al.* (2022) **Excess years of life lost to COVID-19 and other causes of death by sex, neighbourhood deprivation, and region in England and Wales during 2020: A registry-based study** *PLoS Med* **19**

Lee R, Xu W, Dozier M, McQuillan R, Theodoratou E, Figueroa J, UNCOVER and the International Partnership for Resilience in Cancer Systems (I-PaRCS), Breast Cancer Working Group 2 (2023) **A rapid review on the COVID-19's global impact on breast cancer screening participation rates and volumes from January-December 2020** *Elife* **12** https://doi.org/10.7554/eLife.85680

Lee W-E, Woo Park S, Weinberger DM, Olson D, Simonsen L, Grenfell BT, Viboud C (2023) **Direct and indirect mortality impacts of the COVID-19 pandemic in the United States, March 1, 2020 to January 1, 2022** *Elife* **12** https://doi.org/10.7554/eLife.77562

Lemanska A *et al.* (2023) **Healthcare in England was affected by the COVID-19 pandemic across the pancreatic cancer pathway: A cohort study using OpenSAFELY-TPP** *Elife* **12** https://doi.org/10.7554/eLife.85332

Lundberg CE, Santosa A, Björk J, Brandén M, Cronie O, Lindgren M, Edqvist J, Åberg M, Adiels M, Rosengren A (2023) **Age and sex differences in cause-specific excess mortality and years of life lost associated with COVID-19 infection in the Swedish population** *Eur J Public Health* https://doi.org/10.1093/eurpub/ckad086

Msemburi W, Karlinsky A, Knutson V, Aleshin-Guendel S, Chatterji S, Wakefield J (2023) **The WHO estimates of excess mortality associated with the COVID-19 pandemic** *Nature* **613**:130–137

de Lima P Nascimento, van den Puttelaar R, Hahn AI, Harlass M, Collier N, Ozik J, Zauber AG, Lansdorp-Vogelaar I, Rutter CM (2023) **Projected long-term effects of colorectal cancer screening disruptions following the COVID-19 pandemic** *Elife* **12** https://doi.org/10.7554/eLife.85264

National Cancer Institute **Cancer stat facts**

Nickson C *et al.* (2023) **A modelled evaluation of the impact of COVID-19 on breast, bowel, and cervical cancer screening programmes in Australia** https://doi.org/10.7554/eLife.82818

Nonboe MH, Napolitano G, Schroll JB, Vejborg I, Waldstrøm M, Lynge E (2023) **Impact of COVID-19 pandemic on breast and cervical cancer screening in Denmark: A register-based study** *Elife* **12** https://doi.org/10.7554/eLife.81605

Raad II *et al.* (2023) **International multicenter study comparing COVID-19 in patients with cancer to patients without cancer: Impact of risk factors and treatment modalities on survivorship** *Elife* **12** https://doi.org/10.7554/eLife.81127

Rajan KB, Weuve J, Barnes LL, McAninch EA, Wilson RS, Evans DA (2021) **Population estimate of people with clinical Alzheimer's disease and mild cognitive impairment in the United States (2020-2060)** *Alzheimers Dement* **17**:1966–1975

Rüthrich MM *et al.* (2021) **COVID-19 in cancer patients: clinical characteristics and outcome-an analysis of the LEOSS registry** *Ann Hematol* **100**:383–393

Seneviratne SL, Yasawardene P, Wijerathne W, Somawardana B (2022) **COVID-19 vaccination in cancer patients: a narrative review** *J Int Med Res* **50**

Short KR, Kedzierska K, van de Sandt CE (2018) **Back to the Future: Lessons Learned From the 1918 Influenza Pandemic** *Front Cell Infect Microbiol* **8**

Silversmit G, Verdoodt F, Van Damme N, De Schutter H, Van Eycken L. (2021) **Excess Mortality in a Nationwide Cohort of Cancer Patients during the Initial Phase of the COVID-19 Pandemic in Belgium** *Cancer Epidemiol Biomarkers Prev* **30**:1615–1619

Stadlbauer D *et al.* (2021) **Repeated cross-sectional sero-monitoring of SARS-CoV-2 in New York City** *Nature* **590**:146–150

Starkey T *et al.* (2023) **A population-scale temporal case-control evaluation of COVID-19 disease phenotype and related outcome rates in patients with cancer in England (UKCCP)** *Sci Rep* **13**

Tahira AC, Verjovski-Almeida S, Ferreira ST (2021) **Dementia is an age-independent risk factor for severity and death in COVID-19 inpatients** *Alzheimers Dement* **17**:1818–1831

Tope P, Farah E, Ali R, El-Zein M, Miller WH, Franco EL (2023) **The impact of lag time to cancer diagnosis and treatment on clinical outcomes prior to the COVID-19 pandemic: A scoping review of systematic reviews and meta-analyses** *Elife* **12** https://doi.org/10.7554/eLife.81354

U.S. Cancer Statistics Working Group. released in November (2023) **U.S. Cancer Statistics Data Visualizations Tool, based on 2022 submission data (1999-2020): U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute**

Wang Q, Davis PB, Gurney ME, Xu R (2021) **COVID-19 and dementia: Analyses of risk, disparity, and outcomes from electronic health records in the US** *Alzheimers Dement* **17**:1297–1306

Williamson EJ *et al.* (2020) **Factors associated with COVID-19-related death using OpenSAFELY** *Nature* **584**:430–436

Yeh H-C *et al.* (2018) **Diabetes in America**

Zhang Q, Schultz JL, Aldridge GM, Simmering JE, Kim Y, Ogilvie AC, Narayanan NS (2021) **COVID-19 case fatality and Alzheimer's disease** *J Alzheimers Dis* **84**:1447–1452

Editors

Reviewer #1 (Public Review):

Summary:

In the paper, the authors study whether the number of deaths in cancer patients in the USA went up or down during the first year (2020) of the COVID-19 pandemic. They found that the number of deaths with cancer mentioned on the death certificate went up, but only moderately. In fact, the excess with-cancer mortality was smaller than expected if cancer had no influence on the COVID mortality rate and all cancer patients got COVID with the same frequency as in the general population. The authors conclude that the data are consistent with cancer not being a risk factor for COVID and that cancer patients were likely actively shielding themselves from COVID infections.

Strengths:

The paper studies an important topic and uses sound statistical and modeling methodology. It analyzes both, deaths with cancer listed as the primary cause of death, as well as deaths with cancer listed as one of the contributing causes. The authors argue, correctly, that the latter is a more important and reliable indicator to study relationships between cancer and COVID. The authors supplement their US-wide analysis with analysing three states separately.

For comparison, the authors study excess mortality from diabetes and from Alzheimer's disease. They show that Covid-related excess mortality in these two groups of patients was expected to be much higher (than in cancer patients), and indeed that is what the data showed.

https://doi.org/10.7554/eLife.93758.2.sa1

Author response:

The following is the authors' response to the original reviews.

> ***eLife assessment***
>
> *This valuable work explores death coding data to understand the impact of COVID-19 on cancer mortality. The work provides solid evidence that deaths with cancer as a contributing cause were not above what would be expected during pandemic waves, suggesting that cancer did not strongly increase the risk of dying of COVID-19. These results are an interesting exploration into the coding of causes of death that can be used to make sense of how deaths are coded during a pandemic in the presence of other underlying diseases, such as cancer.*

We thank the editor and reviewers for the time they took to review our manuscript and for the thoughtful suggestions they provided. We have completed several revisions based on their feedback and we feel our paper is stronger as a result. However, none of these revisions change the overall conclusions of our study.

> ***Reviewer #1 (Public Review):***
>
> *Summary:*
>
> *In the paper "Disentangling the relationship between cancer mortality and COVID-19", the authors study whether the number of deaths in cancer patients in the USA went up or down during the first year (2020) of the COVID-19 pandemic. They found that the number of deaths with cancer mentioned on the death certificate went up, but only moderately. In fact, the excess with-cancer mortality was smaller than expected if cancer had no influence on the COVID mortality rate and all cancer patients got COVID with the same frequency as in the general population. The authors conclude that the data show no evidence of cancer being a risk factor for COVID and that the cancer patients were likely actively shielding themselves from COVID infections.*
>
> *Strengths:*
>
> *The paper studies an important topic and uses sound statistical and modeling methodology. It analyzes both, deaths with cancer listed as the primary cause of death, as well as deaths with cancer listed as one of the contributing causes. The authors argue, correctly, that the latter is a more important and reliable indicator to study relationships between cancer and COVID. The authors supplement their US-wide analysis by analysing three states separately.*
>
> *Weaknesses:*
>
> *The main findings of the paper can be summarized as six numbers. Nationally, in 2022, multiple-cause cancer deaths went up by 2%, Alzheimer's deaths by 31%, and diabetes deaths by 39%. At the same time, assuming no relationship between these diseases and either Covid infection risk or Covid mortality risk, the deaths should have gone up by 7%, 46%, and 28%. The authors focus on cancer deaths and as 2% < 7%, conclude that cancer is not a risk factor for COVID and that cancer patients must have "shielded" themselves against Covid infections.*
>
> *However, I did not find any discussion of the other two diseases. For diabetes, the observed excess was 39% instead of "predicted by the null model" 28%. I assume this should be interpreted as diabetes being a risk factor for Covid deaths. I think this should*

> *be spelled out, and also compared to existing estimates of increased Covid IFR associated with diabetes.*
>
> *And what about Alzheimer's? Why was the observed excess 31% vs the predicted 46%? Is this also a shielding effect? Does the spring wave in NY provide some evidence here? Why/how would Alzheimer's patients be shielded? In any case, this needs to be discussed and currently, it is not.*

We thank the reviewer for their positive feedback on the paper and for these suggestions. It is true that we have emphasized the impact on cancer deaths, as this was the primary aim of the paper. In the revised version, we have expanded the results and discussion sections to more fully describe the other chronic conditions we used as comparators (lines 267-284;346 – 386).

Note that we are somewhat reluctant to designate any of these conditions as risk factors based solely on comparing the time series model with the demographic model of our expectations. As we mention in the discussion, there is considerable uncertainty around estimates from the demographic model in terms of the size of the population-at-risk, the mean age of the population-at-risk, and the COVID-19 infection rates and infection fatality ratios. Our demographic model is primarily used to demonstrate the effects of competing risks across types of cancers and chronic conditions, since these findings are robust to model assumptions. In contrast, the demographic model should be used with caution if the goal is to titrate the level of these risk factors (as the level of imputed risk is dependent on model assumptions). In the updated version of the manuscript, we have included uncertainty intervals in Table 3, using the upper and lower bounds of the estimated infection rates and IFRs, to better represent this uncertainty. We have also discussed this uncertainty more explicitly in the text and ran sensitivity analyses with different infection rate assumptions in the discussion (lines 354-362; 367 -370).

We would like to note that rather than interpreting the absolute results, we used this demographic model as a tool to understand the relative differences between these conditions. From the demographic model we determined that we would expect to see much higher mortality in diabetes and Alzheimer's deaths compared to cancer deaths due to three factors (1. Size of population-at-risk, 2. Mean age of the population-at-risk, 3. Baseline risk of mortality from the condition), that are separate from the COVID-19 associated IFR. And in general, this is what we observed.

In comparing the results from the demographic model to the observed excess, diabetes does standout as an outlier from cancer and Alzheimer's disease in that the observed excess is consistently above the null hypothesis which does lend support to the conclusion that diabetes is in fact a risk factor for COVID-19. A conclusion which is also supported by many other studies. Our findings for hematological cancers are also similar, in that we find consistent support for this condition being a risk factor. We have commented on this in the discussion and added a few references (lines 346-354; 395-403).

Our hypothesis regarding non-hematological cancer deaths (lower than anticipated mortality due to shielding) could also apply to Alzheimer's deaths. Furthermore, we used the COVID-19 attack rate for individuals >65 years (based on the data that is available), but we estimate that the mean age of Alzheimer's patients is actually 80-81 years, so this attack rate may in fact be a bit too high, which would increase our expected excess. We have commented on this in the discussion (lines 363-377).

> ***Reviewer #2 (Public Review):***
>
> *The article is very well written, and the approach is quite novel. I have two major methodological comments, that if addressed will add to the robustness of the results.*

> *(1) Model for estimating expected mortality. There is a large literature using a different model to predict expected mortality during the pandemic. Different models come with different caveats, see the example of the WHO estimates in Germany and the performance of splines (Msemburi et al Nature 2023 and Ferenci BMC Medical Research Methodology 2023). In addition, it is a common practice to include covariates to help the predictions (e.g., temperature and national holidays, see Kontis et al Nature Medicine 2020). Last, fitting the model-independent for each region, neglects potential correlation patterns in the neighbouring regions, see Blangiardo et al 2020 PlosONE.*

Thank you for these comments and suggestions. We agree there are a range of methods that can be used for this type of analysis, and they all come with their strengths, weaknesses, and caveats. Broadly, the approach we chose was to fit the data before the pandemic (2014-2019), and project forward into 2020. To our knowledge it is not a best practice to use an interpolating spline function to extrapolate to future years. This is demonstrated by the WHO estimates in Germany in the paper you mention. This was our motivation for using polynomial and harmonic terms.

> *Based on the above:*
>
> *a. I believe that the authors need to run a cross-validation to justify model performance. I would suggest training the data leaving out the last year for which they have mortality and assessing how the model predicts forward. Important metrics for the prediction performance include mean square error and coverage probability, see Konstantinoudis et al Nature Communications 2023. The authors need to provide metrics for all regions and health outcomes.*

Thank you for this suggestion. We agree that our paper could be strengthened by including cross validation metrics to justify model performance. Based on this suggestion, and your observations regarding Alzheimer's disease, we have done two things. First, for the full pre-pandemic period (2014-2019) for each chronic condition and location we tested three different models with different degree polynomials (1. linear only, 2. linear + second degree polynomial, 3. linear + second degree polynomial + third degree polynomial) and used AIC to select the best model for each condition and location. Next, also in response to your suggestion, we estimated coverage statistics. Using the best fit model from the previous step, we then fit the model to data from 2014-2018 only and used the model to predict the 2019 data. We calculated the coverage probability as the proportion of weekly observed data points that fell within the 95% prediction interval. For all causes of death and locations the coverage probability was 100% (with the exception of multiple cause kidney disease in California, which is only shown in the appendix). The methods and results have been updated to reflect this change and we have added a figure to the appendix showing the selected model and coverage probability for each cause of death and location (lines 504 – 519; 847-859; Appendix 1- Figure 11).

> *b. In the context of validating the estimates, I think the authors need to carefully address the Alzheimer case, see Figure 2. It seems that the long-term trends pick an inverse U-shape relationship which could be an overfit. In general, polynomials tend to overfit (in this case the authors use a polynomial of second degree).It would be interesting to see how the results change if they also include a cubic term in a sensitivity analysis.*

Thank you for this observation. Based on the changes described above, the model for Alzheimer's disease now includes a cubic term in the national data and in Texas and California. The model with the second-degree polynomial remained the best fit for New York (Appendix 1 – Figure 11).

> *c. The authors can help with the predictions using temperature and national holidays, but if they show in the cross-validation that the model performs adequately, this would be fine.*

At the scale of the US, adding temperature or environmental covariates is difficult and few US-wide models do so (see Goldstein 2012 and Quandelacy 2014 for examples from influenza). Furthermore, because we are looking at chronic disease outcomes, it is unclear that viral covariates or national holidays would drive these outcomes in the same way as they would if we were looking at mortality outcomes more directly related to transmissible diseases (such as respiratory mortality). Our cross validation also indicates that our models fit well without these additional covariates.

> *d. It would be nice to see a model across the US, accounting for geography and spatial correlation. If the authors don't want to fit conditional autoregressive models in the Bayesian framework, they could just use a random intercept per region.*

We think the reviewer is mistaken here about the scale of our national analysis. Our national analysis did not fit independent models for each state or region. Rather, we fit a single model to the weekly-level national mortality data where counts for the whole of the US have been aggregated. We have clarified in the text (lines 156, 464). As such, we do not feel a model accounting for spatial correlation would be appropriate nor would we be able to include a random intercept for each region. We did fit three states independently (NY, TX, CA), but these states are very geographically distant from each other and unlikely to be correlated. These states were chosen in part because of their large population sizes, yet even in these states, confidence intervals were very wide for certain causes of death. Fitting models to each of the 50 US states, most of which are smaller than those chosen here, would exacerbate this issue.

> *(2) I think the demographic model needs further elaboration. It would be nice to show more details, the mathematical formula of this model in the supplement, and explain the assumptions*

Thank you for this comment. We have added additional details on the demographic model to the methods. We have also extended this analysis to each state to further strengthen our conclusions (lines 548-590).

> *Reviewing Editor Recommendations:*
>
> *I think that perhaps something that is missing is that the authors never make their underlying assumption explicit: they are assuming that if cancer increases the risk of dying of COVID-19, this would be reflected in the data on multiple causes of death where cancer would be listed as one of the multiple causes rather than as the underlying cause, and that their conclusions are predicated on this assumption. I would suggest explicitly stating this assumption, as opposed to other reasons why cancer mortality would increase (ex. if cancer care worsened during pandemic waves leading to poorer cancer survival).*

Response: Thank you for this suggestion. We have added a few sentences to the introduction to make this assumption clear (lines 106-112).

> ***Reviewer #1 (Recommendations For The Authors):***
>
> *- It could make sense to add "in the United States" into the title, as the paper only analyses US data.*

> *- It may make sense to reformulate the title from "disentangling the relationship..." into something that conveys the actual findings, e.g. "Lack of excess cancer mortality during Covid-19 pandemic" or something similar. Currently, the title tells nothing about the findings.*

Thank you for these suggestions. We have added “in the US” to the title. However, we feel that our findings are a bit more subtle than the suggested reformulation would imply, and we prefer to leave it in its current form.

> *- Abstract, lines 42--45: This is the main finding of the paper, but I feel it is simplified too strongly in the abstract. Your simulations do *not* "largely explain" excess mortality with cancer; they give higher numbers! Which you interpret as "shielding" etc., but this is completely absent from the abstract. This sentence makes the impression that you got a good fit between simulated excess and real excess, which I would say is not the case.*

Thank you for this comment. We have rephrased the sentence in the abstract to better reflect our intentions for using the demographic model (lines 46-49). As stated above, the purpose of the demographic model was not to give a good fit with the observed excess mortality. Rather, we used the demographic model as a tool to understand the relative differences between these conditions in terms of expected excess mortality given the size, age-distribution, and underlying risk of death from the condition itself, assuming similar IFR and attack rates. And based on this, we conclude that it is not necessarily surprising that we see higher excess mortality for diabetes and Alzheimer’s compared to cancer.

> *- Results line 237: you write that it's "more consistent with the null hypothesis", however clearly it is *not* consistent with the null hypothesis either (because 2% < 7%). You discuss in the Discussion that it may be due to shielding, but it would be good to have at least one sentence about it already here in the Results, and refer to the Discussion.*

We have mentioned this in the results and refer to the discussion (lines 277-278).

> *- Results line 239: why was it closer to the assumption of relative risk 2? If I understand correctly, your model prediction for risk=1 was 7% and for risk=2 it was 13%. In NY you observed 8% (line 187). How is this closer to risk=2?*

Thank you for this observation. We have updated the demographic model with new data, extended the model to state-level data, and included confidence intervals on these estimates. We have also added additional discussion around the differences between our observations and expectations (lines 249-284).

> *- Discussion line 275: "we did not expect to see large increases" -- why exactly? Please spell it out here. Was it due to the age distribution of the cancer patients? Was it due to the high cancer death risk?*

We demonstrate that it is the higher baseline risk of death for cancer that seems to be driving our low expectations for cancer excess mortality (lines 304-320). We have added this to the sentence to clarify our conclusions on this point and have added a figure to better illustrate this concept of competing risks (Figure 6).

> *- Methods, line 405: perhaps it makes sense to cite some other notable papers on Covid excess mortality such as Msemburi et al Nature 2023, Karlinsky & Kobak eLife 2021, Islam et al BMJ 2021, etc.*

Thank you for mentioning this oversight. We certainly should have cited these papers and have included them in the updated version.

> *- Methods line 410: why did you use a 5-week moving average? Why not fit raw weekly death counts? NB regression should be able to deal with it.*

Smoothing time series data with a moving average prior to running regression models is a very common practice. We did a sensitivity analysis using the raw data. This produced excess estimates with slightly larger confidence intervals, but does not change the overall conclusions of the paper.

> *- Methods line 416: please indicate the software/library/package you used for fitting NB regression.*

We fit the NB regression using the MASS package in R version 4.3. We have added this to the methods (line 519).

> *- Line 489: ORCHID -> ORCID*

https://doi.org/10.7554/eLife.93758.2.sa0