Defining Survivorship Trajectories Across Patients With Solid Tumors
An Evidence-Based Approach

Robert L. Dood, MD, MSCE; Yang Zhao, MS; Shannon D. Armbruster, MD, MPH; Robert L. Coleman, MD; Shelley Tworoger, PhD; Anil K. Sood, MD; Keith A. Baggerly, PhD

🔲 **Supplemental content**

IMPORTANCE Survivorship involves a multidisciplinary approach to surveillance and management of comorbidities and secondary cancers, overseen by oncologists, surgeons, and primary care physicians. Optimal timing and coordination of care, however, is unclear and often based on arbitrary 5-year cutoffs.

OBJECTIVE To determine high- and low-risk periods for all tumor types that could define when survivorship care might best be overseen by oncologists and when to transition to primary care physicians.

DESIGN, SETTING, AND PARTICIPANTS In this pan-cancer, longitudinal, observational study, excess mortality hazard, calculated as an annualized mortality risk above a baseline population, was plotted over time. The time this hazard took to stabilize defined a high-risk period. The percent morality elevation above age- and sex-matched controls in the latter low-risk period was reported as a mortality gap. The US population–based Surveillance, Epidemiology, and End Results database defined the cancer population, and the US Census life tables defined controls. Incident cases of patients with cancer were separated into tumor types based on *International Classification of Diseases for Oncology* definitions.

EXPOSURES Population-level data on incident cancer cases was compared with the general US population.

MAIN OUTCOMES AND MEASURES Overall mortality and cause of death were reported on observed cancer cases.

RESULTS A total of 2 317 185 patients (median age, 63 years; 49.8% female) with 66 primary tumor types were evaluated. High-risk surveillance period durations ranged from less than 1 year (breast, prostate, lip, ocular, and parathyroid cancers) up to 19 years (unspecified gastrointestinal cancers). The annualized mortality gap, representing the excess mortality in the stable period, ranged from a median 0.26% to 9.33% excess annual mortality (thyroid and hypopharyngeal cancer populations, respectively). Cluster analysis produced 6 risk cluster groups: group 1, with median survival of 16.2 (5th to 95th percentile range [PR], 10.7-40.2) years and median high-risk period of 2.5 (PR, 0-5.0) years; group 2, 8.3 (PR, 5.1-23.3) and 2.5 (PR, 4.0-8.0) years; group 3, 2.8 (PR, 1.4-3.7) and 7.0 (PR, 6.0-11.1) years; group 4, 1.6 (PR, 1.5-1.8) and 6.0 (PR, 5.1-11.4) years; group 5, 0.8 (PR, 0.5-1.2) and 0.8 (PR, 0.5-1.2) years; and group 6, 0.5 (PR, 0.4-0.8) and 12.0 (PR, 9.3-12.9) years, respectively. Subanalyses of selected tumor types in these groups revealed that stratifying on stage and histologic type can change the risk cluster and guidance for care.

CONCLUSIONS AND RELEVANCE These findings indicate that a standardized 5-year surveillance period is inadequate for some cancers while excessive for others. High-risk cancers require the most resources with the longest high-risk period, highest persistent baseline mortality risk, and longest period of primary cancer mortality, all arguing for longer follow-up with an oncologist in these cancers.

Author Affiliations: Department of Gynecologic Oncology and Reproductive Medicine, The University of Texas MD Anderson Cancer Center, Houston (Dood, Armbruster, Coleman, Sood); Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston (Zhao, Baggerly); Department of Cancer Epidemiology, Moffitt Cancer Center, Tampa, Florida (Tworoger); Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts (Tworoger).

Corresponding Author: Anil K. Sood, MD, Department of Gynecologic Oncology and Reproductive Medicine, The University of Texas MD Anderson Cancer Center, 1155 Herman Pressler, CPB6.3244, Unit 1362, Houston, TX 77030 (asood@mdanderson.org).

With increasing cancer incidence and improved therapies, the population of cancer survivors has increased substantially. There are more than 15 million individuals living with a cancer diagnosis in the United States, and this number is estimated to reach 20.3 million by 2026.[1,2] In a 2005 report, the Institute of Medicine chronologically defined *cancer survivorship* as the period spanning diagnosis, initial treatment, and eventual palliative treatment.[3] Care should meet the following goals: (1) prevention of primary recurrence, (2) surveillance of primary cancer and identification of secondary cancers and late effects, (3) interventions addressing the late treatment effects and comorbidities, and (4) coordination of care.[3(p3)] Survivorship care coordination involves shifting resources and leadership between the patient's primary care physician, medical oncologist, radiation oncologist, oncologic surgeon, medical subspecialists, and other allied health professionals. This has led to a diverse array of delivery systems of survivorship care, which have been studied.[4,5] No clear or universal guidance exists for the transition from oncologist-coordinated care to primary care physician–coordinated care for most cancers. This is, in part, due to the unique natural history across tumor types, as well as the availability of effective treatment assets and likelihood of response to therapy, which can influence mortality over time from initial diagnosis. A risk-stratified approach to survivorship care could identify patients with select cancers who would be well suited for primary physician–led care[6]; however, these models have only been developed for select cancer types and do not have methods that could be universally applied.[5] Arguments have even been made that certain low-risk cancers may not require any follow-up with an oncologist.[7] Thus, it is crucial to accurately characterize survival trajectories across tumor subtypes to allow for personalized and data-driven survivorship care planning.

The objective of this longitudinal descriptive study was to establish data-informed, tumor-specific boundaries representing different phases of survivorship. Establishing standardized nomenclature for survival trajectories can offer justification for different physicians coordinating survivorship care at different points. Using mortality estimates relative to the general population provides an important benchmark to estimate key periods of cancer survivorship. In this manner, a high-risk period (ie, increased risk of death in the next year) and cancer type (risk cluster) would dictate intensive oncologist-led follow-up, which could be led by a primary care physician once a patient reaches a defined lower-risk period that is dominated by secondary cancers, late-treatment effects, comorbidities, and other environmental factors. We addressed this objective using survival data obtained from the Surveillance, Epidemiology, and End Results (SEER) database in a pan-cancer approach, examining trajectories across 66 different tumor types. We quantified 2 key features of cancer survivorship: (1) the time to reach a stable mortality risk and (2) the amount of increased mortality above the baseline mortality of a general control population. Tumors were then grouped into high- and low-risk clusters based on the duration and extent of elevated mortality risk. To further clarify mortality risks, we defined the risk of primary cancer–related mortality vs other cause mortality.

Key Points

Question How can the survivorship period of various cancers be characterized to understand periods of high risk to guide which practitioners might be involved in patient care?

Findings This cohort study of Surveillance, Epidemiology, and End Results data from 2.3 million patients with 66 cancer types used the point at which mortality following cancer diagnosis stabilized to define a high-risk and low-risk period, and further divided cancers into 6 clusters of increasing mortality concern.

Meaning Oncological review may be necessary throughout the high-risk period, and its duration varies by tumor type.

Methods

Cohort Definition

This study was exempt from institutional review board approval and informed consent waived owing to its use of retrospectively collected, publicly available data. We used SEER version 9 data, which include a population-based sample of US cancer cases diagnosed from 1973 to 2013. We included patients with incident cancer younger than 100 years, and excluded those who had a survival or follow-up time less than 1 month, were missing age or survival data, had a prior personal history of cancer, had "lymph node" or "hematologic" primary malignant neoplasms, or if the tumor-type group had fewer than 200 cases (for a robust survival analysis). The cases were grouped using tissue type categories based on the *International Classification of Diseases for Oncology (ICD-O-3)* codes. Baseline demographic characteristics were recorded for each tumor type, including malignant spread at time of diagnosis (localized vs distant). Primary treatment of surgery or radiation therapy was noted; however, chemotherapy data are not available in SEER.

Outcome: Conditional Probability of Death

Overall survival was defined as the duration of the interval from diagnosis to death or end of follow-up, depending on the vital status noted in SEER. Our approach assessed the conditional probability of death within 1 year in the at-risk population. This annualized mortality risk is similar to the instantaneous hazard function used in Kaplan-Meier analyses but does not exceed 1.0. The conditional probability of death was calculated as $P(t \leq T < t + \Delta t \mid T \geq t) = [S(t) - S(t + \Delta t)]/S(t)$, where $\Delta t = 1$ year, and $S(t)$ is the Kaplan-Meier survival function at time t. This is a discrete version of the instantaneous hazard function used in Cox proportional hazard models. We calculated the conditional probability of death within 1 year in the normal population based on the US Census life tables released in 2011,[8] matched to the cancer population by age and sex. After computing the annualized conditional probabilities, we fit the probabilities with a cubic smoothed spline curve, and used bootstrap resampling to compute 95% confidence intervals.

Defining the High-Risk Period for Increased Mortality

For most cancer types, mortality peaks soon after diagnosis, gradually decreases, and plateaus near that of the general population.[2] We separated this follow-up time into 2 periods: a high-risk period with elevated but decreasing annualized mortality and a stable period with lower and steady annualized mortality, which is close to but never meets that of the general population. This transition from the high-risk to stable period was calculated as the point at which the year-to-year difference of the annualized mortality gap between the cancer population and the age- and sex-matched controls changed less than an arbitrarily chosen α = 0.003 cutoff to reliably define the stable period.

Annualized Mortality Gap in the Stable Period

After defining the high-risk and low-risk periods, we calculated the "mortality gap" in the low-risk period. This was reported as the median percent annualized mortality above that of the age- and sex-matched controls in the stable period.

Hierarchical Clustering

We performed hierarchical clustering analyses using data from the first 10 years of follow-up after diagnosis, which includes both the high-risk and stable periods for the majority of cancer subtypes. We used a Euclidean distance metric and the Ward minimum variance-linkage method to cluster the different tissue types,[9] on both the high-risk period duration and the annualized mortality gap, to create clusters of cancers that could be ranked from highest to lowest high-risk duration and elevated annualized mortality risk.[9]

Subgroup Analyses

This pan-cancer analysis did not consider tumor characteristics that are known to influence recurrence and mortality, such as stage and histologic type. To demonstrate the effect of cancer stage, the analyses were repeated after first stratifying on stage. Stage was determined as localized/regional, distantly metastatic, or unstaged as determined by SEER summary staging. A subanalysis of each histologic subtype of all SEER cancers could not be performed and reported efficiently for each cancer type, so we performed this analysis for ovarian cancer as an example. We calculated the high-risk duration and mortality gap for each most common *ICD-O-3* histology codes for ovarian cancer: serous, clear cell, endometrioid, and mucinous ovarian cancers. A permutation *t* test and Cox proportional hazard models were used to determine whether these stratifications resulted in statistically significant model improvement.

Cancer vs Other Cause Mortality

The most common causes of mortality in the cancer survivor populations were reported: primary cancer, secondary cancer, cardiac disease, and all other causes as reported by SEER. To determine the relative contribution of primary cancer–specific mortality in defining a high-risk period duration, we documented the length of the period during which primary cancer mortality was the leading (>50%) contributor to observed mortality, if ever, and reported this as the "cause of death high-risk period."

Figure 1. Cancer Cohort Generation
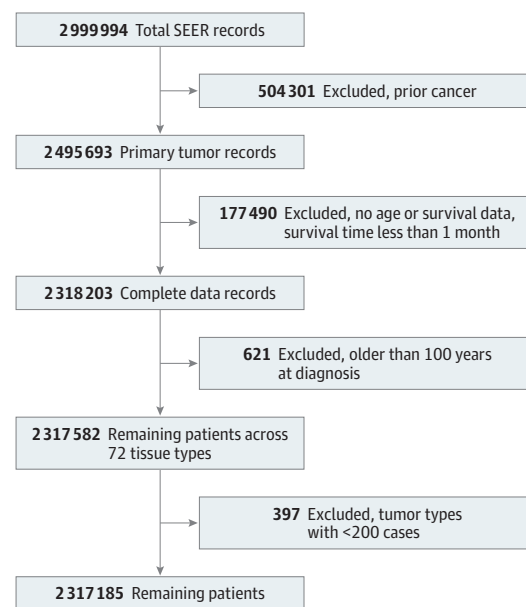
SEER indicates Surveillance, Epidemiology, and End Results.

Results

The SEER database contained data on 2 999 994 people with documented malignant neoplasms. After applying exclusion criteria, 2 317 185 with complete data were available for analysis (**Figure 1**). The final cancer population had a median age of 63 years and was 49.8% female and 83.4% white (**Table 1** and eTable 1 in the Supplement). Total follow-up time ranged from 1 month to 21 years, and 63.6% of patients died during follow-up. A total of 13% had distant disease, 64.0% required surgery, and 27.2% underwent radiation therapy. The cases were sorted into 66 different cancer populations, ranging in size from n = 211 for pituitary gland cancer cases to n = 206 889 for prostate cancer cases.

As an example of the performed high-risk duration analysis (fully outlined in eFigure 1 in the Supplement), **Figure 2** shows a graphical representation of the age- and sex-matched annualized mortality gap in ovarian cancer. Annualized mortality was at maximum immediately following diagnosis, then rapidly decreased over 9 years until it reached stability, defining a 9-year high-risk period. The annualized mortality gap in the stable period remained elevated above the age- and sex-matched control baseline, demonstrating a median mortality gap elevation of 1.2%.

We used this approach for each tumor type. This resulted in a range of high-risk period durations from under 1 year (breast, prostate, lip, ocular, parathyroid cancer populations) to 19 years (unspecified gastrointestinal cancers), reported in eTable 2 in the Supplement. The annualized mortality gap, representing the excess mortality in the stable period, ranged from a median 0.26% to 9.33% excess annual mortality (thyroid and

Table 1. Baseline Demographic Data for Cohort and Most Common Cancer Types[a]

| Tissue | No. | Median (PR), y | | Age, mean (SD), y | No. (%) | | | | | | |
| | | Survival | Follow-up | | Female Sex | Race | | | Treatment | |
						White	Black	Other	Surgery	Radiation Therapy
Prostate	206 889	9.9 (9.9-10.0)	6 (1-19)	69.1 (9.7)	0	156 909 (75.8)	33 755 (16.3)	14 652 (7.1)	99 250 (48.0)	69 813 (33.7)
Localized	110 187	14.7 (14.6-14.8)	7 (1-16)	66.7 (9.5)	0	79 745 (72.4)	20 152 (18.3)	8982 (8.2)	44 796 (40.7)	43 274 (39.3)
Metastatic	5005	2.1 (2.0-2.2)	2 (0-10)	72.1 (11.1)	0	3219 (64.3)	1168 (23.3)	606 (12.1)	676 (13.5)	1196 (23.9)
Lung and bronchus	190 488	0.8 (0.8-0.8)	1 (0-11)	66 (11.3)	73 742 (38.7)	153 457 (80.6)	24 226 (12.7)	12 681 (6.7)	52 773 (27.7)	92 632 (48.6)
Localized	52 080	2.0 (2.0-2.1)	2 (0-14)	67.4 (10.8)	23 950 (46.0)	41 295 (79.3)	6578 (12.6)	4178 (8.0)	26 390 (50.7)	22 232 (42.7)
Metastatic	56 136	0.5 (0.5-0.5)	0 (0-3)	66.5 (11.4)	24 544 (43.7)	42 511 (75.7)	8610 (15.3)	4997 (8.9)	4737 (8.4)	28 788 (51.3)
Breast	188 751	14.7 (14.5-14.8)	9 (1-27)	60.6 (14.2)	187 328 (99.2)	163 802 (86.8)	22 307 (11.8)	2098 (1.1)	176 824 (93.7)	68 451 (36.3)
Localized	171 618	16.1 (16.0-16.2)	10 (1-27)	60.3 (14.0)	170 387 (99.3)	149 508 (87.1)	19 809 (11.5)	1843 (1.1)	167 034 (97.3)	62 962 (36.7)
Metastatic	10 111	1.8 (1.8-1.9)	2 (0-12)	62.9 (14.2)	10 007 (99.0)	8319 (82.3)	1679 (16.6)	95 (0.9)	5909 (58.4)	3904 (38.6)
Colon	172 769	5.3 (5.3-5.4)	4 (0-23)	69.2 (12.5)	88 094 (51.0)	144 501 (83.6)	15 730 (9.1)	12 197 (7.1)	156 886 (90.8)	8001 (4.6)
Localized	132 472	8.3 (8.2-8.4)	6 (0-25)	69.3 (12.3)	67 536 (51.0)	111 117 (83.9)	11 148 (8.4)	9932 (7.5)	129 891 (98.1)	5285 (4.0)
Metastatic	32 307	0.8 (0.8-0.8)	1 (0-6)	67.6 (12.9)	16 170 (50.1)	26 645 (82.5)	3659 (11.3)	1987 (6.2)	23 936 (74.1)	2331 (7.2)
All cancers	2 317 185	7.1 (7.0-7.1)	4 (0-26)	60.9 (17.3)	1 153 342 (49.8)	1 933 332 (83.4)	218 774 (9.4)	148 332 (6.4)	1 482 673 (64.0)	630 689 (27.2)
Localized	1 412 462	13.3 (13.3-13.4)	7 (0-29)	59.8 (16.9)	771 863 (54.6)	1 182 255 (83.7)	127 489 (9.0)	89 535 (6.3)	1 170 266 (82.9)	357 884 (25.3)
Metastatic	301 295	0.7 (0.7-0.7)	1 (0-10)	64.0 (14.6)	158 883 (52.7)	241 681 (80.2)	35 344 (11.7)	23 987 (8.0)	122 231 (40.6)	88 495 (29.4)

Abbreviation: PR, fifth to 95th percentile range.

[a] Unknown and unstaged categories are not displayed.

hypopharyngeal cancer populations, respectively). Details for the most common cancers are reported in **Table 2**.

Using both the morphology and size of this mortality gap, cancers were grouped into 6 clusters (**Table 3**). The annualized mortality gaps for each of the 66 tumor types are shown in eFigure 2 in the Supplement, color-coded by cluster. The lowest-risk cluster 1, which included 18 cancers (eg, breast, prostate, skin), demonstrated a short time to stability (median, 2.5 years) and lower excess mortality (median, 1.4%). Increasing cluster numbers had longer times to stability and increasing excess baseline mortality. For example, the largest mortality gap was observed in cluster 4 (median, 6.1%) and included oral and pharyngeal cancer populations, while the longest high-risk period was observed for cluster 6, which included 3 cancer populations (pancreas, intrahepatic bile duct, and pleural cancer), with a median of 12.0 years.

When stratifying tumors by metastatic vs local disease spread, differences were observed, although many strata were too small to analyze (eTable 2 in the Supplement). The assigned risk cluster was higher in the metastatic clusters for all cancers except retroperitoneal/peritoneal, nasal cavity, pleural, and nasopharyngeal cancers, although the high-risk duration was longer in the metastatic subgroup for each. In the example secondary analysis of ovarian cancer histological subtypes, serous ovarian cancers, known to have a poorer survival,[10] emerged as having a longer high-risk period than other histologic subtypes (eFigure 3 in the Supplement).

The causes of death from primary cancer are reported for each cancer (Table 2 and eTable 2 in the Supplement) and cluster (Table 3), reporting the year at which primary cancer mortality was no longer above 50% of mortality. Death was observed in 63.4% of the entire cancer survivor population, with primary cancer as the leading cause of death (47.9%), fol-

lowed by other causes (21.7%), secondary cancer (16.7%), and cardiac disease (13.7%) (eTable 2 in the Supplement). Primary cancer was the leading cause of death at some point in a majority of the tumor-specific survivor populations (36 [55%]). In those 36 cancer types, the primary cancer–specific cause of death high-risk period ranged from 0 to 13 years, highest for survivors of the group of brain, central nervous system, or spine malignant neoplasms. This cause-specific method of determining the high-risk period tended to produce a high-risk period that was shorter than that defined by the overall mortality gap analysis, except in breast, retina, brain, central nervous system, and spine, and pineal gland cancer survivor popula-

Figure 2. Annualized Mortality Gap Example: Ovarian Cancer
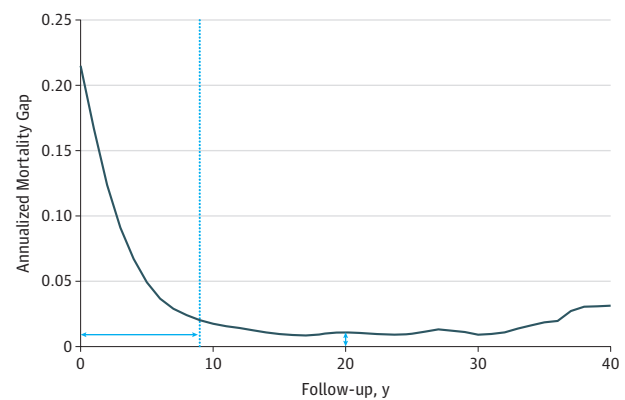
Age- and sex-matched annualized mortality gap for ovarian cancer. The dotted line indicates the high-risk cutoff, and the horizontal and vertical double-headed arrows indicate the high-risk period duration and the persistent mortality gap, respectively.

Table 2. Survivorship Trajectory Details and Cause of Death (COD) for Most Common Cancers

Tissue[a]	No.	High-Risk Duration, y	Mortality Gap, %	Risk Cluster	No. (%)					COD High-Risk Duration, y
					Alive at Censorship	COD				
						Primary Cancer	Secondary Cancer	Cardiac	Other	
Prostate	206 889	0	2.16	1	86 227 (41.7)	36 068 (17.4)	15 858 (7.7)	31 087 (15.0)	37 649 (18.2)	0
Local	110 187	0	0.71	1	77 225 (70.1)	5319 (4.8)	6280 (5.7)	8674 (7.9)	12 689 (11.5)	0
Distant	5005	16	9.44	2	1072 (21.4)	2731 (54.6)	280 (5.6)	388 (7.8)	534 (10.7)	8
Lung and bronchus	190 488	7	6.22	5	15 575 (8.2)	138 116 (72.5)	9066 (4.8)	11 194 (5.9)	16 537 (8.7)	7
Local	52 080	6	6.14	4	11 033 (21.2)	29 195 (56.1)	1981 (3.8)	3711 (7.1)	6160 (11.8)	7
Distant	56 136	11	9.08	6	2895 (5.2)	46 045 (82.0)	2803 (5.0)	1712 (3.0)	2681 (4.8)	7
Breast	188 751	0	1.77	1	89 456 (47.4)	40 871 (21.7)	11 291 (6.0)	19 204 (10.2)	27 929 (14.8)	5
Local	171 618	0	1.58	1	86 912 (50.6)	30 877 (18.0)	10 075 (5.9)	17 766 (10.4)	25 988 (15.1)	0
Distant	10 111	17	2.17	2	1041 (10.3)	7296 (72.2)	532 (5.3)	552 (5.5)	690 (6.8)	13
Colon	172 769	7	1.34	2	39 737 (23.0)	56 845 (32.9)	16 975 (9.8)	25 565 (14.8)	33 647 (19.5)	3
Local	132 472	7	1.29	2	37 335 (28.2)	28 781 (21.7)	12 985 (9.8)	22 896 (17.3)	30 475 (23.0)	3
Distant	32 307	11	2.47	5	1751 (5.4)	24 470 (75.7)	3162 (9.8)	1284 (4.0)	1640 (5.1)	7
All cancers	2 317 185	…	…	…	838 430 (36.3)	704 423 (30.5)	245 056 (10.6)	201 156 (8.7)	319 709 (13.8)	…
Local	1 412 462	…	…	…	217 296 (30.3)	7848 (1.1)	119 520 (16.7)	160 488 (22.4)	211 824 (29.5)	…
Distant	301 295	…	…	…	1224 (20.5)	720 (12.0)	2448 (41.0)	720 (12.0)	864 (14.5)	…

Abbreviation: ellipses, not applicable.

[a] Unstaged cancers are not reported in the local/distant strata.

Table 3. Survival Trajectories and Causes of Death for Risk Clusters

Risk Cluster	Median (PR)			Cause of Death % of At-Risk Population				Metastatic Spread,[a] %		Cancer Types
	Survival, y	High-Risk Period, y	Mortality Gap, %	Primary Cancer	Secondary Cancer	Cardiac	Other	Localized	Distant	
1	16.2 (10.7-40.2)	2.5 (0-5.0)	1.4 (0.3-2.4)	6.7 (1.2-18.0)	10.1 (5.0-27.0)	8.8 (0.5-16.1)	15.1 (5.1-23.7)	83.8 (47.7-95.5)	5.7 (1.1-45.8)	Lip, skin, breast, vagina/labia, vulva, cervix, uterus, placenta, penis/scrotum, prostate, testis, epididymis/spermatic cord/male genital NOS, orbit/lacrimal gland, retina, eyeball, thyroid, parathyroid, pituitary
2	8.3 (5.1-23.3)	6.0 (4.0-8.0)	2.5 (0.9-4.6)	19.1 (4.0-30.0)	18.2 (5.3-35.8)	6.9 (1.2-14.4)	14.1 (8.1-22.2)	66.1 (0-86.0)	13.1 (0.1-44.0)	Tongue, gum/mouth floor, salivary gland, oropharynx, nasopharynx, small intestine, colon, appendix, rectum, anus, nasal cavity, larynx, thymus, mediastinum, bones/joints, spleen, connective soft tissue, other female genitals, kidney, ureter/renal pelvis, urinary bladder, other urinary organs, eye NOS, meninges, ventricle, cerebellum, pineal gland
3	2.8 (1.4-3.7)	7.0 (6.0-11.1)	2.1 (1.3-3.7)	30.8 (6.5-47.8)	30.7 (8.0-45.9)	4.1 (2.8-6.0)	9.6 (6.9-20.4)	31.5 (0-48.7)	34.4 (1.0-62.6)	Middle ear accessory muscles, retroperitoneum/peritoneum, uterus NOS, ovary, other nervous system, adrenal glands, ill-defined
4	1.6 (1.5-1.8)	6.0 (5.1-11.4)	6.1 (4.0-9.0)	18.3 (15.5-23.7)	42.6 (40.4-52.6)	7.3 (7.2-7.8)	14.2 (9.8-15.2)	62.1 (50.9-72.5)	21.3 (15.2-31.7)	Hypopharynx, pharynx, trachea
5	0.8 (0.5-1.2)	9.0 (6.3-16.9)	2.4 (1.1-6.0)	56.8 (19.8-71.2)	12.4 (4.5-56.0)	4.3 (1.8-6.2)	8.9 (8.3-12.7)	49.2 (2.1-65.2)	29.0 (5.9-68.1)	Esophagus, stomach, liver, gallbladder, unspecified digestive organs, lung/bronchus, heart, brain/central nervous system/spine
6	0.5 (0.4-0.8)	12.0 (9.3-12.9)	3.0 (1.1-4.2)	54.5 (19.4-78.8)	29.0 (7.5-65.4)	2.1 (1.7-2.5)	5.7 (4.8-6.5)	38.0 (24.1-46.7)	48.7 (32.5-59.2)	Intrahepatic bile ducts, pancreas, pleura

Abbreviations: NOS, not otherwise specified; PR, 5th-95th percentile range.

[a] Unstaged cancers are not reported in the Metastatic Spread columns.

tions. This cause-specific method was also performed on a stage-stratified basis and is reported in Table 2 and eTable 2 in the Supplement. Similar to the overall mortality gap determination of the high-risk periods, the cause of death–determined high-risk period was longer for metastatic vs localized cancer survivor populations.

Discussion

This study developed and implemented a novel analytic method to classify the highest-risk periods for a multitude of tumor types and to quantify the amount of elevated risk in mortality above that of the general population during the stable period. This extends our understanding of cancer survival trajectories beyond the simple 5-year mortality metric or risk of recurrence. The durations of high-risk periods serve as useful cutoffs to define long-term cancer survivors, who represent an important group to study and understand the underlying factors.[11] We then used clustering methods to identify a group of lowest-risk cancer types (ie, short high-risk period of cancer-related death and small mortality gap compared with the general population), as well as groups of high-risk cancer types (ie, long duration of high-risk period and/or large mortality gap). This importantly defines the period when cancer survivor populations might benefit from greater oncology involvement. The cause of death analysis also identifies the large group of cancer survivor populations who never have primary cancer as the leading cause of death, and could have these risks addressed by their primary care physician in shared care, and/or specialists such as cardiologists, oncocardiologists, and others to address their comorbidities and treatment-related effects.[12] Risk clusters 3, 4, 5, and 6 all had high-risk periods longer than 5 years, the highest mortality gap percentage in the stable period, and had primary and secondary cancers contribute to the majority of observed deaths; this argues that they would benefit most from greater oncology involvement. Cluster 1 had the lowest high-risk period duration and median mortality gap, and had cardiac and other causes contributing to a majority of deaths. This cluster had the most surgical management (86.3%), and represents cancer populations that would benefit from early primary care physician–led care after curative cancer treatment from the oncologist and surgical team.

Specific subtypes of each cancer may fall into different risk clusters based on the mortality gap and high-risk period, warranting further study based on type-specific criteria (eg, histologic type, expression of certain tumor markers, grade). The stage-stratified analyses suggest that the localized vs metastatic dichotomy informs the high-risk duration and risk cluster category. Breast cancer populations with metastatic disease, for instance, had a much longer high-risk duration (17 vs 0 years), and existed in risk cluster 2 among metastatic cancers, suggesting that they are a subpopulation whose care plans require a more tailored approach. Pancreatic cancer, however, remained in the highest-risk cluster 6 regardless of stratification. The histology-stratified ovarian cancer example demonstrated that different histologic sub-

types can fall within different risk clusters. These subanalyses wholly demonstrate that there are a number of potential subpopulations for each tumor that may have different high-risk period durations, elevated mortality risk, or risk cluster classification. Although an exhaustive analysis of these tumor subtypes is beyond the scope of this article, the mortality-based clusters provide robust, broad, and accessible categories. We have provided our analysis code so the analysis could be performed for tumors of interest after accessing SEER data using analysis code in R (https://github .com/zhaozhangyangzi/SEER_survivorship.git).

This study additionally identified patients with cancer who have highly increased mortality not due to the primary cancer in the stable period, who could benefit from intense interventions to mitigate posttreatment effects, secondary cancers, and other comorbidities. The cause of death analysis showed that just over half of the cancer populations had their primary cancer as the leading cause of death in the observed follow-up period. This adult study contrasts with an investigation of causes of death of childhood cancer survivors, where primary cancer recurrence remained the leading cause of death even decades after diagnosis.[13] Our study emphasizes the need to focus resources in the survivorship period on other causes of mortality such as comorbidities and heart disease, as well as secondary cancers and treatment effects.

The strengths of this study include its scope and long-term follow-up provided by the SEER database to inform nearly every solid-tumor survivor population in the United States. The use of overall mortality is the most universal approach to incorporate cancer recurrence and mortality, treatment effects, secondary cancer risks, comorbidities, and other related environmental factors, although in a nonspecific manner. This can address the diverse patients encountered by oncologists and primary care physicians. Our analysis uses an unbiased, universal, and reproducible approach relying on mortality data to group cancer survivor populations into clusters of increasing risk, whose survivorship care can then be tailored to the unique care delivery system available to each patient. The additional information provided by the stratified analysis based on stage and the cause of death data can further personalize survivorship care when sufficient resources are available.

Limitations

The exhaustive pan-cancer analysis does serve as a weakness to the study, in that it is not possible to report on specific stage, grade, and histologic type subpopulations within the scope of a single report. The SEER data do not report specific details such as chemotherapy administration, which could further inform late-term treatment effects leading to mortality. Additionally, our primary mortality-based cluster analysis incorporates the balance of cancer mortality and non-primary cancer mortality. However, the reported cause of death secondary analysis findings appear to confirm the cluster's ability to identify populations that could benefit from more oncologist-led or more primary care–led survivorship care. This mortality-based approach also does not address morbidity, as SEER does not contain patient-reported outcomes, or costs.

Conclusions

The analyses presented herein could inform evidence-based follow-up guidelines in the United States and worldwide, and also demonstrate that survivorship care should be individualized on patient-level factors such as stage and histologic subtype. However, additional work is needed to further define survivorship trajectories specifically for subpopulations with characteristics known to influence prognosis.

Better guidelines and care planning can lead to more efficient allocation of public health resources, as well as balancing the time, and insurance reimbursements for care led by oncologists, oncologic surgeons, primary care physicians, medical subspecialists, and allied health professionals. We believe that this pan-cancer analysis can be used to inform explicit and evidence-based survivorship care planning that maximizes both survival and delivery of high-quality cancer survivorship care.

REFERENCES

1. DeSantis CE, Lin CC, Mariotto AB, et al. Cancer treatment and survivorship statistics, 2014. *CA Cancer J Clin*. 2014;64(4):252-271. doi:10.3322/caac .21235

2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin*. 2018;68(1):7-30. doi:10 .3322/caac.21442

3. Hewitt M, Greenfield S, Stovall E, eds. *From Cancer Patient to Cancer Survivor*. Washington, DC: National Academies Press; 2005. doi:10.17226/11468

4. Halpern MT, Viswanathan M, Evans TS, Birken SA, Basch E, Mayer DK. Models of cancer survivorship care: overview and summary of

current evidence. *J Oncol Pract*. 2015;11(1):e19-e27. doi:10.1200/JOP.2014.001403

5. McCabe MS, Partridge AH, Grunfeld E, Hudson MM. Risk-based health care, the cancer survivor, the oncologist, and the primary care physician. *Semin Oncol*. 2013;40(6):804-812. doi:10.1053/j.seminoncol.2013.09.004

6. Nekhlyudov L, O'malley DM, Hudson SV. Integrating primary care providers in the care of cancer survivors: gaps in evidence and future opportunities. *Lancet Oncol*. 2017;18(1):e30-e38. doi:10.1016/S1470-2045(16)30570-8

7. Haymart MR, Miller DC, Hawley ST. Active surveillance for low-risk cancers—a viable solution to overtreatment? *N Engl J Med*. 2017;377(3): 203-206. doi:10.1056/NEJMp1703787

8. Arias E. United States life tables, 2011. *Natl Vital Stat Rep*. 2015;64(11):1-63.

9. Murtagh F, Legendre P. Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? *J Classif*. 2014;31(3): 274-295. doi:10.1007/s00357-014-9161-z

10. Matz M, Coleman MP, Carreira H, Salmerón D, Chirlaque MD, Allemani C; CONCORD Working Group. Worldwide comparison of ovarian cancer survival: histological group and stage at diagnosis (CONCORD-2). *Gynecol Oncol*. 2017;144(2):396-404. doi:10.1016/j.ygyno.2016.11.019

11. Garrido P, Rosell R, Massutí B, et al; Spanish Lung Cancer Group 0008 Investigators. Predictors of long-term survival in patients with lung cancer

included in the randomized Spanish Lung Cancer Group 0008 phase II trial using concomitant chemoradiation with docetaxel and carboplatin plus induction or consolidation chemotherapy. *Clin Lung Cancer*. 2009;10(3):180-186. doi:10.3816/CLC .2009.n.025

12. Petek BJ, Greenman C, Herrmann J, Ewer MS, Jones RL. Cardio-oncology: an ongoing evolution. *Future Oncol*. 2015;11(14):2059-2066. doi:10.2217 /fon.15.89

13. Reulen RC, Winter DL, Frobisher C, et al; British Childhood Cancer Survivor Study Steering Group. Long-term cause-specific mortality among survivors of childhood cancer. *JAMA*. 2010;304(2): 172-179. doi:10.1001/jama.2010.923