Original Research Article

TJ Tumori Journal

Tumori Journal
2021, Vol. 107(4) 311–317

DOI: 10.1177/0300891620961839
journals.sagepub.com/home/tmj


Cancer prevalence by phase of care: an indicator for assessing health service needs

Anna Gigli[1], Silvia Francisci[2], Stefano Guzzinati[3], Aaron Hall[4], Mark Hachey[4], Steve Scoppa[4] and Angela Mariotto[5]

Abstract

Introduction: Cancer prevalence (people alive on a certain date in a population who previously had a cancer diagnosis) is expected to increase in the United States and Europe due to improvements in survival and population aging. Examination of prevalence by phase of care allows us to identify subgroups of patients according to their care trajectories, thus allowing us to improve health care planning, resource allocation, and calculation of costs.

Methods: A new method to estimate prevalence by phase of care using grouped data is illustrated. Prevalence is divided into 3 mutually exclusive phases: initial, continuing, and end-of-life. An application to US and Italian data is applied to prevalent cases diagnosed with colon–rectum, stomach, lung, or breast cancer.

Results: The distribution of phase of care prevalence estimated by cancer type and sex and results from the two datasets are very similar. Most survivors are in the continuing phase; the end-of-life phase is larger for cancers with worse prognosis. All phases prevalence is generally higher in the Italian than in the US dataset, except for lung cancer in women, where prevalence proportion in the Italian dataset is 30% lower than in the United States.

Discussion: Incidence, survival, and population age structure are the main determinants of prevalence and they can affect differences in all phases of prevalence, as well as in discrete phases. Incidence is the most influential determinant. Ours is the first study that compares prevalence by phase of care between two populations in Italy and the United States. Despite great differences in health care management in the two countries, we found extremely similar distribution of survivors by phase of care for most cancer sites under study.

Keywords

Population-based cancer registry, COMPREV software, complete cancer prevalence, phase of care prevalence

Date received: 17 June 2020; revised: 24 July 2020; accepted: 13 August 2020

Introduction

The number of cancer survivors has been increasing and is expected to increase further in the United States and Europe[1,2] mainly due to improvements in survival and population aging. Complete prevalence represents all cancer survivors previously diagnosed with cancer and it is not easily estimated from cancer registry data, since patients diagnosed prior to the start of registration are missed. Limited duration prevalence is directly estimated from cancer registry data and represents survivors by time since diagnosis for the maximum length of the cancer registration. The most common method to estimate complete prevalence is implemented in COMPREV software.[3] The method uses limited

[1]Institute for Research on Population and Social Policies, National Research Council, Rome, Lazio, Italy
[2]National Centre for Disease Prevention and Health Promotion, National Health Institute, Rome, Italy
[3]Veneto Tumour Registry, Azienda Zero, Padova, Italy
[4]Information Management Services, Calverton, MD, USA
[5]Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD, USA

Corresponding author:
Silvia Francisci, National Centre for Disease Prevention and Health Promotion, National Health Institute, viale Regina Elena 299, Rome 00161, Italy.
Email: silvia.francisci@iss.it

duration prevalence coupled with cancer incidence and survival models to estimate the number of missed survivors diagnosed prior to the start of registration.[4,5]

Cancer prevalence includes all survivors irrespective of where in the care trajectory they are. Although an important measure of cancer burden and cancer survivorship, it is not specific enough to be useful in health care planning, resource allocation, or calculation of costs. The breakdown of prevalence into phases of care has been used in several studies.[6–12] Phases of care have usually been classified into three clinically relevant phases: initial (first 12 months after diagnosis), end of life (EOL) (12 months prior to death), and continuing phase, defined as the time in between initial and EOL. Different methods have been used to estimate prevalence by phases of care depending on the type of available data. For individual data, patients are classified into a single phase using the dates of diagnosis and date of death if the patient died.[11,12] If grouped prevalence data are available, methods combining prevalence and survival have been used.[6]

The objective of this study is to introduce a new method to estimate prevalence by phases of care using grouped prevalence data that was recently implemented into a dedicated session of the COMPREV software and to present an application to United States and Italian data. This new method uses the comparison of prevalence in two consecutive calendar years to estimate prevalence by phases of care.

We illustrate the method using data from the Italian Veneto cancer registry (CR) and from the Surveillance, Epidemiology, and End Results (SEER) Program to calculate complete prevalence by phases of care for (female) breast, colorectal, lung, and stomach cancers. We also investigate the performance of the method using a very short limited-duration prevalence (5 years) versus longer (23 years).

Materials and Methods

Estimating complete prevalence

Complete prevalence (CP) represents the proportion of people of a given age who are alive on a certain date Y with a prior diagnosis of cancer regardless of when the diagnosis occurred. Limited-duration prevalence (LDP) represents the proportion of people of a given age who are alive on a certain date Y and had a cancer diagnosis in a limited period of time. Maximum duration of this period depends on the number of years the registry has been collecting incidence cases. LDP is calculated from cancer registry data using the SEER*Stat software.[13] For patients with cancer lost to follow-up, the method estimates the probability of them being alive at the prevalence date. CP is estimated using LDP and the completeness index method[4] to estimate survivors diagnosed prior to cancer registration.

This method is implemented in the COMPREV software. In a new release of the software, the phase of care session has been added with the following specific input requirements: (1) two LDP data files: the first can refer to any year Y, and the second must be referring to the following year Y + 1; (2) these files must be identical in their settings except for the year covered, and must be stratified by single ages at prevalence.

Prevalence decomposition by phases of care

We define here three mutually exclusive phases of care—initial, continuing, and EOL—and assign each case in the cohort to the phase of care he or she belongs to on prevalence date, in the following way: if he or she had been diagnosed within 12 months before prevalence date and is alive 12 months after prevalence date, he or she belongs to the initial phase; if he or she had been diagnosed more than 12 months before prevalence date and is alive 12 months after prevalence date, he or she belongs to the continuing phase; if he or she died within 12 months after prevalence date, regardless of when he or she had been diagnosed, he or she belongs to the EOL phase. It should be noticed that although during his or her life span each person can contribute to more than one phase of care, on the prevalence date each patient belongs to only one phase of care; that is, the phases of care are mutually exclusive.

Prevalence by phase of care session in COMPREV

We calculate LDP by time since diagnosis and age in single years at 2 consecutive years Y and Y + 1. We then calculate complete prevalence, CP(Y) and CP(Y + 1), using the completeness index method. The comparison between CP(Y) and CP(Y + 1) allows for the estimation of prevalence by phases of care as described below. We illustrate here the algorithm to calculate each phase of care, and details are presented in the appendix.

1. The prevalence of patients in the initial phase includes all patients who were diagnosed within 12 months before Y (1-year LDP) and are still alive at year Y + 1.
2. The continuing phase includes all patients who were diagnosed more than a year before *Y* and are still alive at year Y + 1.
3. The EOL phase includes all patients who will die during year Y + 1, regardless of when they were diagnosed. Patients can die of causes attributed to the cancer or other causes, such as accident or other chronic conditions. Thus, the EOL phase is further classified into patients who will die for a cancer-related cause and those who will die for a non-cancer-related cause. To subdivide the EOL phase in EOL-cancer death and EOL-non-cancer death, we use information on crude probabilities of dying of cancer and dying of other causes by time since diagnosis.[14]

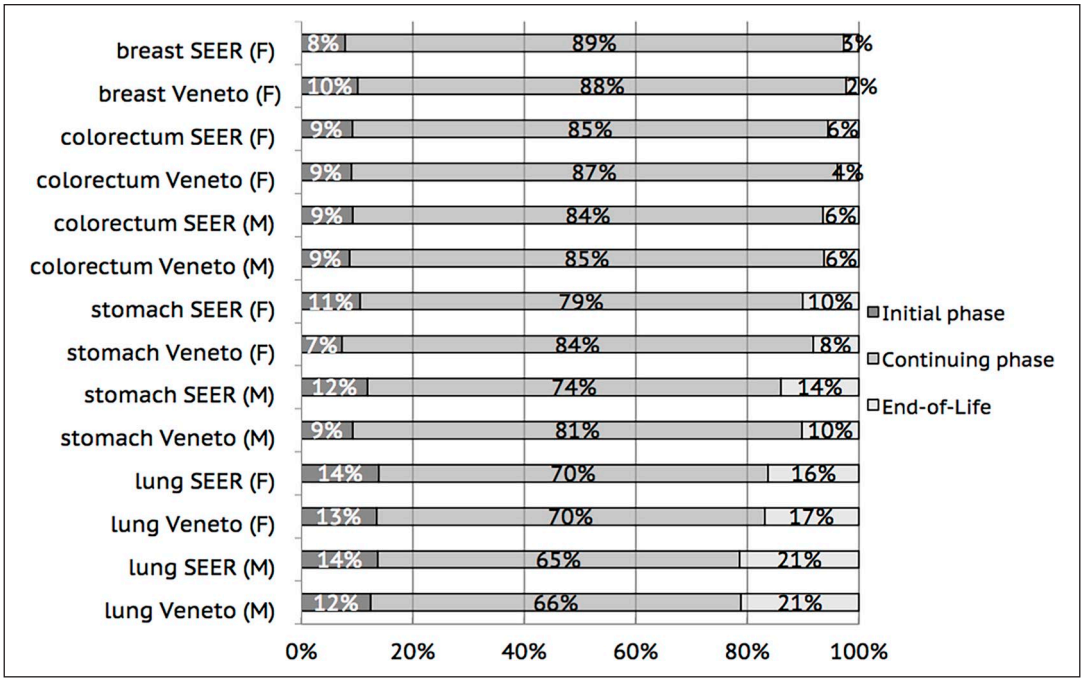

Figure 1. Distribution of complete prevalence by phase of care, cancer type, and sex. All ages (0+). Prevalence date: January 1, 2015. Data from Surveillance, Epidemiology, and End Results (SEER) 13 and Veneto registries.

Data sources

We used data from the Veneto registry,[15] a population-based cancer registry that covers about 1.3 million inhabitants in northeast Italy, and from the US SEER 13 registries,[16] which covers the population of San Francisco–Oakland, Connecticut, Detroit, Hawaii, Iowa, New Mexico, Seattle, Utah, Atlanta (Metropolitan), San Jose–Monterey, Los Angeles, Alaska Natives, and rural Georgia, corresponding to approximately 13.4% of the US population, and selected patients of all ages diagnosed with the following cancer types: colon and rectum, stomach, lung, and breast (females only), classified according to the International Classification of Diseases for Oncology.[17]

From the Veneto registry, 28- and 29-year LDPs were estimated, selecting patients diagnosed in 1987–2014 and followed for vital status up to January 1, 2015, and patients diagnosed in 1987–2015 and followed for vital status up to January 1, 2016, respectively.

From the SEER 13 registries, 23- and 24-year LDPs were estimated, selecting patients diagnosed in 1992–2014 and followed for vital status up to January 1, 2015, and patients diagnosed in 1992–2015 and followed for vital status up to January 1, 2016, respectively. In order to measure the performance of the method, we also estimated 5- and 6-year LDPs from SEER 13 registries, selecting patients diagnosed in 2010–2014 and followed for vital status up to January 1, 2015, and patients diagnosed in 2010–2015 and followed for vital status up to January 1, 2016, respectively.

LDP matrices used as input in COMPREV must be stratified by single ages at prevalence date.

Prevalence decomposition refers to Y = January 1, 2015. For each calendar year Y, we obtain the number of survivors by single year of age up to 84 and all survivors aged 85+.

Results

Figure 1 shows the distribution of age-adjusted prevalence by phases of care by cancer type and sex for all ages combined for survivors in SEER and Veneto CR. Complete prevalence is estimated from 23-year LDP in SEER 13 and 28-year LDP in Veneto CR.

The distributions by phases of care in the two areas are very similar. Most survivors are in continuing phase and the relevance of this phase varies according to the cancer site: cancers with better prognosis (i.e. longer survival) have more survivors (for example, 89% in SEER and 88% in Veneto in female breast cancer), and vice versa for cancers with worse prognosis, such as lung in males (65% in SEER and 66% in Veneto). On the other hand, the EOL phase is larger for sites with worse prognosis: it varies from 16% to 21% in lung cancer to 2% to 3% in female breast cancer; only 1%–3% of cases in the final year of life will die of non-cancer-related causes (data not shown).

The main difference between Veneto and SEER is found in stomach cancer in males, with more cases in the continuing phase in Veneto and fewer in the initial and EOL phases.

Table 1. Complete prevalence (crude and age-adjusted), 5-year age-adjusted relative survival (2010–2016), and age-adjusted incidence rates (2006–2015) by cancer site, phase of care, and sex in Surveillance, Epidemiology, and End Results (SEER) 13 and Veneto cancer registries.

Cancer site, sex, and cancer registry	Prevalence per 100,000 (crude)				Prevalence per 100,000 (age-adjusted)				Age-adjusted rates	
	All phases	Initial phase	Continuing phase	End-of-life phase	All phases	Initial phase	Continuing phase	End-of-life phase	5-Year relative survival (2010–2013)[a]	Incidence per 100,000 (2006–2015)[b]
Breast, F										
Veneto	7119	396	6361	362	2734	210	2439	84	88.6	97.9
SEER	8850	316	7999	535	1905	123	1717	65	89.7	91.4
Colon and rectum, M										
Veneto	4415	301	3631	483	782	61	661	60	66.7	40.8
SEER	3154	154	2714	285	383	31	325	26	62.9	29.6
Colon and rectum, F										
Veneto	2798	186	2464	148	681	52	600	29	67.2	26.4
SEER	2616	129	2272	215	378	28	326	24	65.4	22.4
Lung, M										
Veneto	1064	88	731	245	201	22	134	44	14.8	39.5
SEER	1134	120	749	264	136	18	88	30	18.6	36.5
Lung, F										
Veneto	384	30	258	95	117	13	81	23	22.4	14.2
SEER	984	93	707	184	161	21	113	27	26.7	28.0
Stomach, M										
Veneto	717	28	605	85	121	9	100	13	31.6	9.7
SEER	321	28	245	49	40	4	30	6	30.0	6.4
Stomach, F										
Veneto	425	26	353	46	97	6	82	9	37.1	5.0
SEER	179	12	149	18	29	3	24	3	37.4	3.4

All ages (0+). Prevalence date: January 1st 2015. SEER 13 and Veneto databases.
[a]Corazziari et al.[20] standard population.
[b]World standard population.

Table 1 compares the distribution of prevalence proportions between SEER and Veneto by cancer type, sex, and phase of care. In order to disentangle differences due to the age structure in the two populations from those due to other determinants (i.e. incidence and survival), age-adjusted prevalence proportions and age-adjusted 5-year relative survival and incidence rates[18,19] are also reported.

Crude all-phase prevalence is generally higher in Veneto than in SEER: it varies from 40% more, as in breast cancer, to three times more, as in stomach cancer. The exception is lung cancer in the female population, where prevalence proportion in Veneto is 30% lower than in SEER.

These patterns remain after age adjustment, but differences between Veneto and SEER exist: age-adjusted prevalence (using the Corazziari et al.[20] standard population) in Veneto is 20% higher in breast cancer and is double that of SEER in stomach cancer. In lung cancer, after adjusting by age, the difference among women becomes wider, with prevalence in Veneto 40% lower than prevalence in SEER, while the difference disappears among men.

In general, a similar pattern between the two populations is found across the phases of care. Exceptions are stomach cancer, where differences are lower in the initial and EOL phases, and breast cancer, where differences are higher in the initial phase.

To compare the impact of shorter versus longer LDP (5-year versus 23-year LDP), we estimated CP rates by phase of care and cancer site for all ages combined, using SEER 13 data (Table 2). Although the length in years of the LDP affects the estimation of CP, it does not affect the decomposition by phase of care: the percentage difference between CP distribution estimated from 5-year LDP and CP distribution estimated from 23-year LDP is never higher than 3%.

Discussion

A new method for decomposing prevalence by phase of care is presented. The method has been implemented in the latest release of COMPREV software, freely available at the National Cancer Institute website. The results described are

Table 2. Limited-duration prevalence (LDP) and complete prevalence (CP) (proportions ×100,000) estimated from 5-year LDP and 23-year LDP, by cancer site, phase of care, and sex.

Cancer site and phase of care	23-Year LDP	23-Year CP	5-Year LDP	5-Year CP
Men				
Colon and rectum				
Initial	31.3	31.3	31.7	31.7
Continuing	296.9	325.1	93.5	267.5
End-of-life cancer	13.6	14.9	10.7	18.2
End-of-life other cause	10.0	11.2	3.0	9.6
Total	351.9	382.6	138.9	327.0
Lung				
Initial	17.9	17.9	18.1	18.1
Continuing	75.2	88.3	36.8	97.6
End-of-life cancer	25.7	26.5	22.6	26.7
End-of-life other cause	3.0	3.1	1.5	1.7
Total	121.9	135.7	79.1	144.1
Stomach				
Initial	4.47	4.47	4.44	4.44
Continuing	26.8	30.1	11.0	28.9
End-of-life cancer	4.5	4.7	3.9	4.9
End-of-life other cause	0.9	0.9	0.3	0.4
Total	36.65	40.13	19.70	38.69
Females				
Colon and rectum				
Initial	28.4	28.4	28.3	28.3
Continuing	287.1	325.5	87.1	273.2
End-of-life cancer	12.0	13.5	9.1	15.5
End-of-life other cause	9.0	10.3	2.2	7.8
Total	336.5	377.7	126.7	324.9
Lung				
Initial	20.6	20.6	21.0	21.0
Continuing	101.8	113.3	48.8	113.7
End-of-life cancer	23.8	24.4	20.4	24.6
End-of-life other cause	2.8	2.8	1.3	1.6
Total	149.0	161.0	91.5	160.9
Stomach				
Initial	2.7	2.7	2.7	2.7
Continuing	21.3	23.8	8.7	23.5
End-of-life cancer	2.4	2.5	2.0	2.9
End-of-life other cause	0.4	0.4	0.2	0.3
Total	26.8	29.4	13.6	29.4
Breast				
Initial	122.9	122.9	128.8	128.8
Continuing	1511.6	1716.8	448.6	1551.4
End-of-life cancer	22.5	26.9	11.6	29.7
End-of-life other cause	31.5	38.0	8.7	34.7
Total	1688.5	1904.6	597.7	1744.5

All ages (0+). Prevalence date: January 1, 2015. Surveillance, Epidemiology, and End Results (SEER) 13 database.

based on data from SEER 13 and Veneto CR. Decomposing prevalence by phase of care allows a more accurate estimation of the economic burden of cancer, showing that costs vary according to the phase of care. Many studies adopted this approach to estimate cancer-related costs over time, thus incorporating the natural history of the disease and corresponding patterns of treatment.[7–9,21,22]

In our application, the distribution of phase-of-care prevalence shows a pattern that is common to the different cancer types considered and similar to other studies: a smaller proportion of cases in initial phase and final phase and higher proportion in continuing phase.

Ours is the first study that compares prevalence by phases of care between two international registries: Veneto in Italy

and SEER 13 in the United States. Despite great differences in health care management in the two countries, we found extremely similar distribution of survivors by phase of care for most of the cancer sites under study, as shown in Figure 1.

Incidence, survival, and population age structure are the main determinants of prevalence and they can affect differences in all phases of prevalence, as well as in some specific phases. Incidence is the most influential determinant. In this study, we generally found higher prevalence proportions in Veneto than in SEER (Table 1). This is only partially explained by the older age structure in the Veneto population with respect to the SEER population. After age adjustment, major differences remain, with prevalence in Veneto double that in SEER in stomach cancer and 20%–30% higher in the other cancers. These differences are particularly evident in initial and continuing phases of care and can be attributed to higher incidence found in Veneto with respect to SEER. The only exception to the geographic pattern described is lung cancer female prevalence: crude prevalence of women in SEER is 30% higher than in Veneto and this difference is even wider (40%) after age adjustment. This result is coherent with the differences found in the other two determinants of prevalence presented in Table 1: incidence rate in Veneto is half that of SEER and this is probably related to smoking habits over the past 20 years among Italian women compared to US women, as reported in the Organization for Economic Cooperation and Development data on smoking prevalence[23]; 5-year survival in Veneto is 20% lower than in SEER. The effects of these two determinants reinforce each other and are only partially compensated by the older age structure observed in the Veneto female population.[21,22]

Other studies have shown similar distributions of phase care as ours, using a different estimation method (PIAMOD[24]): Mariotto et al.[6] projected 2010–2020 prevalence in the United States by phase of care, and the distribution of phase-of-care prevalence for stomach, colorectal, and female breast cancer is similar to our findings: the difference never exceeds 3%; the exception is lung cancer, where the difference in prevalence distribution is higher, especially in the initial phase (14% in SEER 13 in this study vs 21% in Mariotto et al.[6]). In the Mariotto et al.[6] study, a different set of registry data has been used (SEER 17) and the results may be affected by the different smoking prevalence in the populations covered by the SEER 13 and SEER 17 registries, particularly relevant in the case of lung cancer results. Yu et al.[11] used Australian data for projecting and decomposing prevalence of colorectal cancer, and the results are comparable with ours: 11% of males and 10% of females are in the initial phase, 84% of males and 86% of females are in the continuing phase, and 5% of males and 4% of females are in the EOL phase.

From Table 2, it is interesting to note that, although slight differences in the total number of prevalent cases occur between CP obtained from 5-year LDP and CP obtained from 23-year LDP, especially for cancer types having better prognosis (colorectal cancer and female breast cancer), there is virtually no difference in the CP decomposition by phases of care, even when looking at age groups 0–49, 50–69, and 70–84 (data not shown).

There are a number of limitations due to the completeness index method. First, complete prevalence strongly depends on the maximum duration of the registry data, as shown in Table 2, and the method is not recommended for registries with only a few years of data available; in the phase of care approach, this limitation affects especially the continuing phase. Moreover, the method reconstructs the unobserved portion of prevalence in the past (before the registry started its activity) but does not allow to project prevalence in the future, unlike the approach implemented in the PIAMOD software. Finally, the estimation of survival parameters must be performed outside the COMPREV software (although default parameters estimated on SEER data are provided).

The method has been validated (results not shown) by comparing prevalence by phases of care estimates with respective estimates assuming that individual data are available and each individual care trajectory is known. The distribution of prevalent cases by phase of care is identical between the two methods for all cancer sites considered in this article.

A strength of our proposed method is that it is based on LDP and its implementation requires only estimation of the completeness index. Furthermore, it has been implemented in a new release of the COMPREV software,[3] freely available and easy to use; the software contains a set of default parameter estimates, obtained from SEER data. Other software, such as PIAMOD, relies more strongly on statistical modeling of incidence and survival.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship and/or publication of this article: This work was supported by National Institutes of Health grant HHSN2612015000362P.

ORCID iDs

Anna Gigli https://orcid.org/0000-0003-4929-505X
Silvia Francisci https://orcid.org/0000-0002-1552-1744

Supplemental material

Supplemental material for this article is available online.

References

1. Parry C, Kent EE, Mariotto AB, et al. Cancer survivors: a booming population. *Cancer Epidemiol Biomarkers Prev* 2011; 20: 1996–2005.

2. Guzzinati S, Virdone S, De Angelis R, et al. Characteristics of people living in Italy after a cancer diagnosis in 2010 and projections to 2020. *BMC Cancer* 2018; 18: 169.
3. National Cancer Institute. Complete Prevalence Program, version 3.0.29 (beta), https://surveillance.cancer.gov/comprev/ (November 2019, accessed June 4, 2020).
4. Capocaccia R and De Angelis R. Estimating the completeness of prevalence based on cancer registry data. *Stat Med* 1997; 16: 425–440.
5. Merril RM, Capocaccia R, Feuer EJ, et al. Cancer prevalence estimates based on tumour registry data in the Surveillance, Epidemiology, and End Results (SEER) program. *Int J Epidemiol* 2000; 29: 197–207.
6. Mariotto AB, Yabroff KR, Feuer EJ, et al. Projecting the number of patients with colorectal carcinoma by phases of care in the US: 2000–2020. *Cancer Causes Control* 2006; 17: 1215–1226.
7. Mariotto AB, Yabroff KR, Shao Y, et al. Projections of the cost of cancer care in the United States: 2010–2020. *J Natl Cancer Inst* 2011; 103: 117–128.
8. Francisci S, Guzzinati S, Mezzetti M, et al. Cost profiles of colorectal cancer patients in Italy based on individual patterns of care. *BMC Cancer* 2013; 329.
9. Laudicella M, Walsh B, Burns E, et al. Cost of care for cancer patients in England: evidence from population-based patient-level data. *Br J Cancer* 2016; 1286: 1292.
10. de Oliveira C, Pataky R, Bremner KE, et al. Phase-specific and lifetime costs of cancer care in Ontario, Canada. *BMC Cancer* 2016; 809.
11. Yu XQ, Clements M and O'Connell D. Projections of cancer prevalence by phase of care: a potential tool for planning future health service needs. *J Cancer Surviv* 2013; 7: 641–651.
12. Francisci S, Guzzinati S, Capodaglio G, et al. Patterns of care and cost profiles of women with breast cancer in Italy: EPICOST study based on real world data. *Eur J Health Econ* 2020; 21.
13. National Cancer Institute. SEER*Stat Software, version 8.3.6.1, https://seer.cancer.gov/seerstat/ (accessed June 4, 2020).
14. National Cancer Institute; Howlader N, Noone AM, et al, eds. SEER Cancer Statistics Review, 1975–2016, based on November 2018 SEER data submission, https://seer.cancer.gov/csr/1975_2016/ (April 2019, accessed October 25, 2019).
15. Registro Tumori Veneto, https://registrotumoriveneto.it (accessed June 4, 2020).
16. National Cancer Institute, DCCPS, Surveillance Research Program. Surveillance, Epidemiology, and End Results (SEER) program, SEER*Stat database: incidence–SEER 13 regs research data, sub (1992–2014), Katrina/Rita population adjustment, linked to county attributes, total US, 1969–2015 counties, based on the November 2016 submission, www.seer.cancer.gov (April 2017, accessed October 25, 2019).
17. National Cancer Institute. International Classification of Diseases for Oncology (ICDO-3/WHO 2008), available at https://seer.cancer.gov/siterecode (accessed September 25, 2020)
18. National Cancer Institute; Howlader N, Noone AM, Krapcho, M, et al, eds. SEER Cancer Statistics Review, 1975–2017, based on November 2019 SEER data submission, https://seer.cancer.gov/csr/1975_2017/ (accessed April 30, 2020).
19. Registro Tumori Veneto. Veneto data, https://www.registrotumoriveneto.it/it/dati/dati-del-veneto (accessed June 4, 2020).
20. Corazziari I, Quinn M and Capocaccia R. Standard cancer patient population for age standardising survival ratios. *Eur J Cancer* 2004; 40: 2307–2316.
21. Pisu M, Henrikson NB, Banegas MP, et al. Costs of cancer along the care continuum: what we can expect based on recent literature. *Cancer* 2018; 124: 4181–4191.
22. Goldsbury DE, Yap S, Weber MF, et al. Health services costs for cancer care in Australia: estimates from the 45 and Up Study. *PLoS One* 2018; 13.
23. Organization for Economic Cooperation and Development. Daily smokers (indicator), https://www.oecd-ilibrary.org/social-issues-migration-health/daily-smokers/indicator/english_1ff488c2-en (accessed June 4, 2020).
24. Verdecchia A, De Angelis G and Capocaccia R. Estimation and projections of cancer prevalence from cancer registry data. *Stat Med* 2002; 21: 3511–3526.