Annals of Oncology Advance Access published August 22, 2014

Long-term survival, prevalence, and cure of cancer: A population-based estimation for 818,902 Italian patients and 26 cancer types

L. Dal Maso[1], S. Guzzinati[2], C. Buzzoni[3,4], R. Capocaccia[5], D. Serraino[1], A. Caldarella[4], A. P. Dei Tos[2,6], F. Falcini[7], M. Autelitano[8], G. Masanotti[9], S. Ferretti[10], F. Tisano[11], U. Tirelli[12], E. Crocetti[4], R. De Angelis[5], and the AIRTUM Working group[*]

[1]Epidemiology and Biostatistics Unit, CRO Aviano National Cancer Institute IRCCS, Aviano, Italy

[2]Veneto Region Cancer Registry, Padua, Italy

[3]AIRTUM database, Florence, Italy

[4]Tuscany Cancer Registry, Clinical and descriptive epidemiology Unit, Cancer Prevention and Research Institute (ISPO), Florence, Italy

[5]National Centre for Epidemiology, Surveillance and Health Promotion (CNESPS), National Institute of Health (ISS), Rome, Italy

[6]Department of Oncology, Anatomic Pathology Unit, General Hospital of Treviso, Italy

[7]Romagna Cancer Registry, Cancer Institute of Romagna (IRCSS), Meldola, Italy

[8]Milan Cancer Registry, Milan Health Authority, Epidemiology Unit, Milan, Italy

[9]Umbria Cancer Registry, Department of Medical and Surgical Specialties, and Public Health, Section of Public Health, Perugia University, Perugia, Italy

[10]Ferrara Cancer Registry, Ferrara University, Ferrara, Italy

[11]Siracusa Cancer Registry, ASP of Siracusa, Siracusa, Italy

[12]Medical Oncology Unit, CRO Aviano National Cancer Institute IRCCS, Aviano, Italy

Correspondence to: Dr. Luigino Dal Maso, Epidemiology and Biostatistics Unit, CRO Aviano National Cancer Institute IRCCS, Via Franco Gallini 2, 33081 Aviano, Italy. Tel: +39-0434-659354; Fax: +39-0434-659231. E-mail: epidemiology@cro.it

Key Message: "The study provides population-based estimates of several epidemiological indicators of long-term survival and cure among cancer patients, by cancer type, sex, and age. More than a quarter of persons living in Italy after a cancer diagnosis have death rates similar to those of the general population. Moreover, nearly three quarters of them will not die as a result of their cancer. These estimates may be helpful to health care planners, clinicians, and patients."

Abstract

Background Persons living after a cancer diagnosis represent 4% of the whole population in high income Countries. The study aim was to provide estimates of indicators of long-term survival and cure for 26 cancer types, presently lacking.

Patients and Methods Data on 818,902 Italian cancer patients diagnosed at age 15-74 years in 1985-2005 were included. Proportions of patients with the same death rates of the general population (cure fractions) and those of prevalent patients who were not at risk of dying as a result of cancer (cure prevalence) were calculated, using validated mixture cure models, by cancer type, sex, and age group. We also estimated complete prevalence, conditional relative survival (CRS), time to reach five- and ten-year CRS>95%, and proportion of patients living longer than those thresholds.

Results The cure fractions ranged from >90% for patients aged <45 years with thyroid and testis cancers to <10% for liver and pancreatic cancers of all ages. Five- or ten-year CRS>95% were both reached in less than ten years by patients with cancers of the stomach,

colon-rectum, pancreas, corpus and cervix uteri, brain, and Hodgkin lymphoma. For breast cancer patients, five- and ten-years CRSs reached >95% after 19 and 25 years, respectively, and in 15 and 18 years for prostate cancer patients. Five-year CRS remained <95% for >25 years after cancer diagnosis in patients with liver and larynx cancers, non-Hodgkin lymphoma, myeloma, and leukaemia. Overall, the cure prevalence was 67% for men and 77% for women. Therefore, 21% of male and 31% of female patients had already reached five-year CRS>95%, while 18% and 25% had reached ten-year CRS>95%.

Conclusions A quarter of Italian cancer patients can be considered cured. This observation has a high potential impact on health planning, clinical practice, and patients perspective.

Keywords: Survival, prevalence, cancer cure, Italy.

Introduction

In the first decade of the 2000s, persons living after a cancer diagnosis represented 4% of the whole population in high income countries [1-3], and more than 60% of cancer patients survived longer than five years after diagnosis [1,2].

Standard survival indicators, namely five-year or ten-year relative survival (RS) [4,5], do not differentiate patients who, in the long-term, will die because of cancer from those who will be cured and will die of other causes [6]. In cancer patients, the risk of death for specific neoplasm is highest in the initial years after diagnosis, and decreases thereafter until a time period when it becomes negligible, and all the surviving patients reach the same life expectancy of the sex- and age-matched general population [7,8].

The general aim of the present study was to expand the spectrum of indicators of long-term survival and cure among cancer patients, in order to provide helpful information for epidemiologists and healthcare planners, as well as for oncologists [9] and patients [7].

Specific aims of this study were to compute: a) the proportion of cancer cases expected to have the same death rates of the general population (cure fraction); b) the number of years after cancer diagnosis necessary to eliminate excess mortality due to cancer (time to cure); c) the overall proportion of cancer patients not at risk of dying as a result of cancer (cure prevalence); d) the proportion of prevalent cancer patients who survived longer than ‘time to cure’ and who already reached the same death rates of the general population.

Materials and methods

Data collected on 818,902 cancer patients diagnosed at age 15-74 years in 1985-2005 by Italian cancer registries included 454,527 cancer cases were diagnosed in men (85,053 lung, 63,047 prostate, and 56,635 colorectal cancers, Supplementary Data, S-Table1).

Corresponding numbers in women were 364,375, including 128,004 breast, 41,864 colorectal, and 20,398 endometrial cancers (Supplementary Data, S-Table1).

Detailed description of statistical methods is provided in Supplementary Data: Extended description of Statistical Methods. Briefly, the observed RS was calculated for cases diagnosed in 1985-2002 and followed-up until 2007, by cancer type, sex, age at diagnosis, and period of diagnosis. RS were also modelled by means of mixture cure models including continuous age and period of diagnosis effects [1,10,11]. Sex specific model-based survival estimates were also calculated for the overall population (15-74 years), at the average age at diagnosis for each cancer type.

Proportions of patients with the same death rates of the general population (cure fractions Supplementary Data, S-Figure1A) were calculated. In addition, we estimated complete prevalence adjusting the observed prevalence in each registry with the completeness index method [1,11,12], conditional relative survival (CRS), time to reach five- and ten-year CRS>95% [13] (Supplementary Data, S-Figure1B), and proportion of patients living longer than those thresholds (Supplementary Data, S-Figure1C). Finally, we calculated the proportion of prevalent patients who were not at risk of dying as a result of cancer (cure prevalence, Supplementary Data, S-Figure1C) [14].

Results

Among Italian cancer patients diagnosed at ages 15-74 between 1985 and 2005, the cure fraction was highest in patients below 45 years of age for thyroid (99% in women and 95% in men), testis (94%), and corpus uteri (91%). Conversely, cure fractions <10% emerged, for all ages and sexes, for liver and pancreatic cancer patients and, for patients aged 55 years or older, diagnosed with cancers of the lung, gallbladder, brain, and leukemias (Table 1). Cure fractions were largely higher (by 10% or more) in women than in men for all age groups in

patients with cancer of the oral cavity, skin melanoma, kidney, bladder, and thyroid cancers. A less clear advantage (0 to 10%) for women emerged for patients with stomach and colorectal cancer, while for all other cancer types no difference of cure fraction emerged across age groups (Table 1).

The cure prevalence proportions were particularly high for patients with cancer of testis (98%), thyroid (91% in men and 97% in women), cervix uteri (95%), corpus uteri (93%), and Hodgkin lymphoma (92%, both sexes) (Figure 1). The cure prevalences were 72% for breast cancer patients, 64% for prostate, 83% in men and 87% in women diagnosed with colorectal cancer, whereas it was less than one third for patients with liver cancer, myeloma and leukaemia (Figure 1). The overall cure prevalence proportion for all examined cancer types encompassed 73% (i.e., 67% of men and 77% women) of persons living after cancer diagnosis.

In both sexes, time to cure was reached in less than ten years, consistently so for different definitions used, by patients with cancers of the stomach and colon-rectum (both five to nine years), pancreas (five to six), cervix and corpus uteri (five to nine years), and brain (seven to eight). In particular, time to cure was reached in less than five years by women with thyroid cancer and by men with testicular cancer (Table 2). For patients with liver and larynx cancers, non-Hodgkin lymphoma, myeloma, and leukaemia, time to cure was not reached or it was >15 years for all definitions used. For these cancers CRS remained <90%, in comparison with the general population, for 15 years or more after cancer diagnosis (Table 2).

For other cancer types, the time to reach the different thresholds was heterogeneous according to definition used. In women with breast cancer, a five-year CRS >90% was reached before ten years from diagnosis (14 years for patients ages 15-44) but five-year CRS >95% in nearly 20 years and ten year CRS >95% in 25 years or more. For patients with prostate cancer these times to cure were reached in ten, 15, and 18 years, respectively. Thus, these thresholds were

variable for patients with kidney and bladder cancers. As a consequence (Table 3), proportions of all patients living after a cancer diagnosis who reached time to cure was relatively high in both sexes, and according to different thresholds, for cervix (>70% among 167/100,000 women) and corpus uteri (>50% among 258/100,000 women), testis (>90% among 150/100,000 men), brain (>50% among 56/100,000 men and women), thyroid cancer (>75% among 255/100,000 women), and Hodgkin lymphoma (>50% among 95/100,000 men and 95/100,000 women). Less than 10% of patients with liver, laryngeal, prostate cancers, non-Hodgkin lymphoma, myeloma, and leukaemia had already reached the same or similar death rates of the general population. Among all Italian women, 1709/100,000 were alive after a breast cancer diagnosis, 41% had already reached five-years CRS>90% but only 6% reached ten-year CRS>95%. Similar heterogeneity of proportions of cured patients according to time to cure definition emerged for kidney and, most notably, for bladder cancer patients (five-year CRS>90% reached by 29%, while ten-year CRS>95% by only 1% among 438/100,000 men). The proportion of patients who reached five-year or ten-year CRS>95% was intermediate for colorectal cancer (30% and 27% among 422/100,000 men; 40% and 36% among 352/100,000 men), stomach cancer (38% and 37% among 109/100,000 men; 58% and 49% among 73/100,000 women), and skin melanoma (49% and 41% among 184/100,000 women) (Table 3).

Overall, 27% of all cancer patients (21% in men and 31% in women), or 0.9% of the Italian population, had reached five-year CRS>95% and 22% (18% in men and 25% in women) ten-year RS>95%.

Discussion

This study showed that a quarter (27%) of persons living in Italy in the first decade of the 2000s after a cancer diagnosis had reached a death rate similar to that of the general

population. In addition, nearly three quarters (73%) of them will not die as a result of their cancer. All the indicators of long-term survival showed a huge heterogeneity by cancer type, age group and, less markedly, by sex.

Present findings for breast cancer patients were in substantial agreement with many previous studies, reporting that a small (i.e., <10%) but significant excess mortality remains at least up to 15 years after diagnosis [13,15,16]. However, approximately half of the breast cancer patients will not die as a result of their cancer [6,17], reaching a negligible excess risk of death after approximately 20 years since diagnosis. A similar pattern emerged for men living after a prostate cancer diagnosis [6,7,15], while a more favourable long-term survival emerged for colorectal [7,14,15,18] and invasive cervical cancers [7,13,19] with cure fractions >50% reached in eight years. A cure fraction <10% emerged for lung and pancreas cancer patients, and no excess risk of death remained after nine and six years since diagnosis, including 15-20% of prevalent patients [6,15,20]. As elsewhere, the cure prevalence for patients with non-Hodgkin lymphoma, myeloma, and leukaemia was <50%, and for these patients excess mortality never became negligible [7,13,15].

A substantial advantage in all indicators of long-term cancer survival emerged in women for some cancer types (e.g., colorectal, skin melanoma, bladder, kidney, thyroid, and non-Hodgkin lymphoma) [15,21], variously attributed to lower prevalence of comorbidity, earlier stage at diagnosis, and better resistance to disease than men [22]. Moreover, a poorer long-term survival with increasing age was reported for almost all cancer types, possibly due to late recurrences or adverse treatment effects, secondary tumours or increased comorbidities [7,13].

Strengths and limitations

To the best of our knowledge, this is the first study reporting a wide spectrum of validated indicators of long-term survival and cure of cancer, including cure fraction and cure

prevalence of particular interests to orient public health policies. In addition, for 26 cancer types (representing 95% and 97% of all cancers incidence in men and women, respectively), we estimated the time to reach the same death rates of the general population, which is of overwhelming importance for cancer patients and oncologists.

Potential limitations of cure models are known and should be considered (Supplementary Data: Extended description of Statistical Methods) [23]. In general, cure models may not be appropriate for data with too short a follow-up to identify a plateau in the tail [24]. In this scenario, difficulties in identifying model parameters arose (cancer specific excess mortality did not become negligible) and estimates of time to cure are either longer than the observation time (i.e., leukaemia) or rather sensible to the choice of the CRS threshold (i.e. prostate or breast cancers). However, in the present study, all the used models converged and fittings were graphically assessed (Supplementary Data, S-Figure2 and S-Table2).

The accuracy of the present estimates depended also on the size of the study population and on the follow-up length, which, in turn, were the strengths of our study. Indeed, our population-based survival estimates were based on very large cohorts of patients followed-up for more than 20 years after diagnosis in order to maximize the reliability of the survival parameters estimates.

Histological subtype is an important prognostic factor at diagnosis for many neoplasms, particularly for non-Hodgkin lymphoma and leukaemia [25]. Unfortunately, we were not able to calculate survival and prevalence estimates by histological subtypes [1,7,26]. We also lacked information on other important prognostic factors, in particular, stage and treatment. Previous reports have shown that stage has a prognostic effect, mainly during the first years after diagnosis, which lessens and can disappear for long-term survival [13].

New therapies, in particular biological treatments for solid tumours and lymphomas, have improved the outcome of cancer patients over time. There is a possibility that, for some

neoplasms, adjuvant treatments may prolong survival, but they do not affect lethality [27]. Unfortunately, population-based studies with a long follow-up period could hardly allow these stratifications.

RS could be biased when background cancer risk factors (e.g., smoke, HCV) carry a higher risk of related mortalities. This effect has been shown to be negligible for lung cancer [28] and could result in downward RS estimates and in prolonged time to reach the same mortality of the general population.

The generalization of results herein presented is also questionable even if Italian RS levels were similar to those of most European countries [5] and that differences became smaller with time since diagnosis [13].

Present long-term survival and cancer cure estimates reflected the average survival time of large groups of people (i.e., a population) rather than an individual prognosis. Moreover, they must be interpreted considering that a quantitative estimation of lacking excess mortality is not always the equivalent of well-being. Parallel studies on cancer rehabilitation needs, including indicators of quality of life, are also necessary [29].

The availability of reliable and accurate estimates of long-term survival and cure for the increasing number of persons living many years since cancer diagnosis may be helpful not only to epidemiologists and health care planners, but also to clinicians in developing guidelines to enhance and standardize the long-term follow-up of cancer survivors [9,30]. Most of all, they could be helpful to patients dealing with uncertainty about the future, making important life decisions, and supporting their rehabilitation demands.

Acknowledgements

We thank Marilena Bongiovanni and cancer patients' associations stimulating and supporting the study. The authors also thank Dr Maja Pohar Perme and Dr Enzo Coviello for helpful comments, and Mrs Luigina Mei for editorial assistance.

Funding

This work was supported by the Italian Association of Cancer Research (AIRC) (grant number 11859).

Disclosure The authors have declared no conflicts of interest.

References

1. AIRTUM Working Group. Italian cancer figures, report 2010: Cancer prevalence in Italy. Patients living with cancer, long-term survivors and cured patients. Epidemiol Prev 2010; 34(5-6 Suppl 2): 1-188. http://www.registri-tumori.it/cms/?q=Rapp2010
2. Parry C, Kent EE, Mariotto AB, et al. Cancer Survivors: A Booming Population. Cancer Epidemiol Biomarkers Prev 2011; 20: 1996–2005.
3. Gatta G, Mallone S, van der Zwan JM, et al. Cancer prevalence estimates in Europe at the beginning of 2000. Ann Oncol 2013; 24: 1660-1666.
4. Coleman M, Quaresima M, Berrino F, et al. Cancer Survival in five continents: a worldwide population-based study (CONCORD). Lancet Oncol 2008; 9: 730-756.
5. De Angelis R, Sant M, Coleman MP, et al. Cancer survival in Europe 1999-2007 by country and age: results of EUROCARE-5-a population-based study. Lancet Oncol 2014; 15: 23–34.
6. Francisci S, Capocaccia R, Grande E, et al. The cure of cancer: A European perspective. Eur J Cancer 2009; 45: 1067-1079.
7. Baade PD, Youlden DR, Chambers SK. When do I know I am cured? Using conditional estimates to provide better information about cancer survival prospects. Med J Aust 2011; 194: 73-77. Erratum in Med J Aust 2011; 194: 376.
8. Ellison LF, Bryant H, Lockwood G, Shack L. Conditional survival analyses across cancer sites. Health Rep 2011; 22: 21-25.
9. McCabe MS, Bhatia S, Oeffinger KC, et al. American society of Clinical Oncology Statement: achieving high-quality cancer survivorship care. J Clin Oncol 2013; 31: 631-640.

10. De Angelis R, Capocaccia R, Hakulinen T, et al. Mixture models for cancer survival analysis: application to population-based data with covariates. Stat Med 1999; 18: 441-454.
11. Merrill RM, Capocaccia R, Feuer EJ, Mariotto A. Cancer prevalence estimates based on tumour registry data in the Surveillance, Epidemiology, and End Results (SEER) Program. Int J Epidemiol 2000; 29: 197-207.
12. Capocaccia R, De Angelis R. Estimating the completeness of prevalence based on cancer registry data. Stat Med 1997; 16: 425-440.
13. Janssen-Heijnen MLG, Gondos A, Bray F, et al. Clinical relevance of conditional survival of cancer patients in Europe: Age-specific analyses of 13 cancers. J Clin Oncol 2010; 28: 2520-2528.
14. Gatta G, Capocaccia R, Berrino F, et al. Colon cancer prevalence and estimation of differing care needs of colon cancer patients. Ann Oncol 2004; 15: 1136-1142.
15. Smastuen M, Aagnes B, Johannsen TB, et al. Long-term cancer survival: patterns and trends in Norway 1965-2007. Oslo: Cancer Registry of Norway; 2008. http://www.kreftregisteret.no/Global/Publikasjoner%20og%20rapporter/CIN2007_del2.pdf
16. Janssen-Heijnen MLG, van Steenbergen LN, Voogd AC, et al. Small but significant excess mortality compared with the general population for long-term survivors of breast cancer in the Netherlands. Ann Oncol 2014; 25: 64–68.
17. Woods LM, Rachet B, Lambert PC, et al. 'Cure' from breast cancer among two populations of women followed for 23 years after diagnosis. Ann Oncol 2009; 20: 1331-1336.
18. Lambert PC, Dickman PW, Osterlund P, et al. Temporal trends in the proportion cured for cancer of the colon and rectum: a population-based study using data from the Finnish Cancer Registry. Int J Cancer 2007; 121: 2052-2059.

19. Andrae B, Andersson TML, Lambert PC, et al. Screening and cervical cancer cure: population based cohort study. BMJ. 2012;344:e900.
20. Sinn M, Striefler JK, Sinn BV, et al. Does long-term survival in patients with pancreatic cancer really exist? Results from the CONKO-001 study. J Surg Oncol 2013; 108: 398-402.
21. Cvancarova M, Aagnes B, Fosså SD, et al. Proportion cured models applied to 23 cancer sites in Norway. Int J Cancer 2013; 132: 1700-1710.
22. Micheli A, Ciampichini R, Oberaigner W, et al. The advantage of women in cancer survival: an analysis of EUROCARE-4 data. Eur J Cancer 2009; 45: 1017-1027.
23. Yu XQ, De Angelis R, Andersson TM, et al. Estimating the proportion cured of cancer: Some practical advice for users. Cancer Epidemiol 2013; 37: 836-842.
24. Othus M, Barlogie B, Leblanc ML, Crowley JJ. Cure models as a useful statistical tool for analyzing survival. Clin Cancer Res 2012; 18: 3731-6.
25 Andersson TML, Lambert PC, Derolf AR, et al. Temporal trends in the proportion cured among adults diagnosed with acute myeloid leukemia in Sweden 1973-2001, a population-based study. Br J Haematol 2009; 148: 918-924.
26. AIRTUM Working Group. Italian cancer figures, report 2011. Survival of cancer patients in Italy. Epidemiol Prev 2011; 35(Suppl3): 1-200. http://www.registri-tumori.it/cms/Rapp2011.
27. Early Breast Cancer Trialists' Collaborative Group (EBCTCG), Davies C, Godwin J, et al. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. Lancet 2011; 378: 771-784.
28. Hinchliffe SR, Rutheford MJ, Crowter MJ, et al. Should relative survival be used with lung cancer data. Br J Cancer 2012; 106: 1584-1589.

29. Baili P, Hoekstra-Weebers J, Van Hoof E, et al. Cancer rehabilitation indicators for Europe. Eur J Cancer 2013; 49: 1356-1364.

30. De Moor JS, Mariotto AB, Parry C, et al. Cancer survivors in the United States: prevalence across the survivorship trajectory and implication of care. Cancer Epid Biom Prev 2013; 22: 561-570.

Appendix

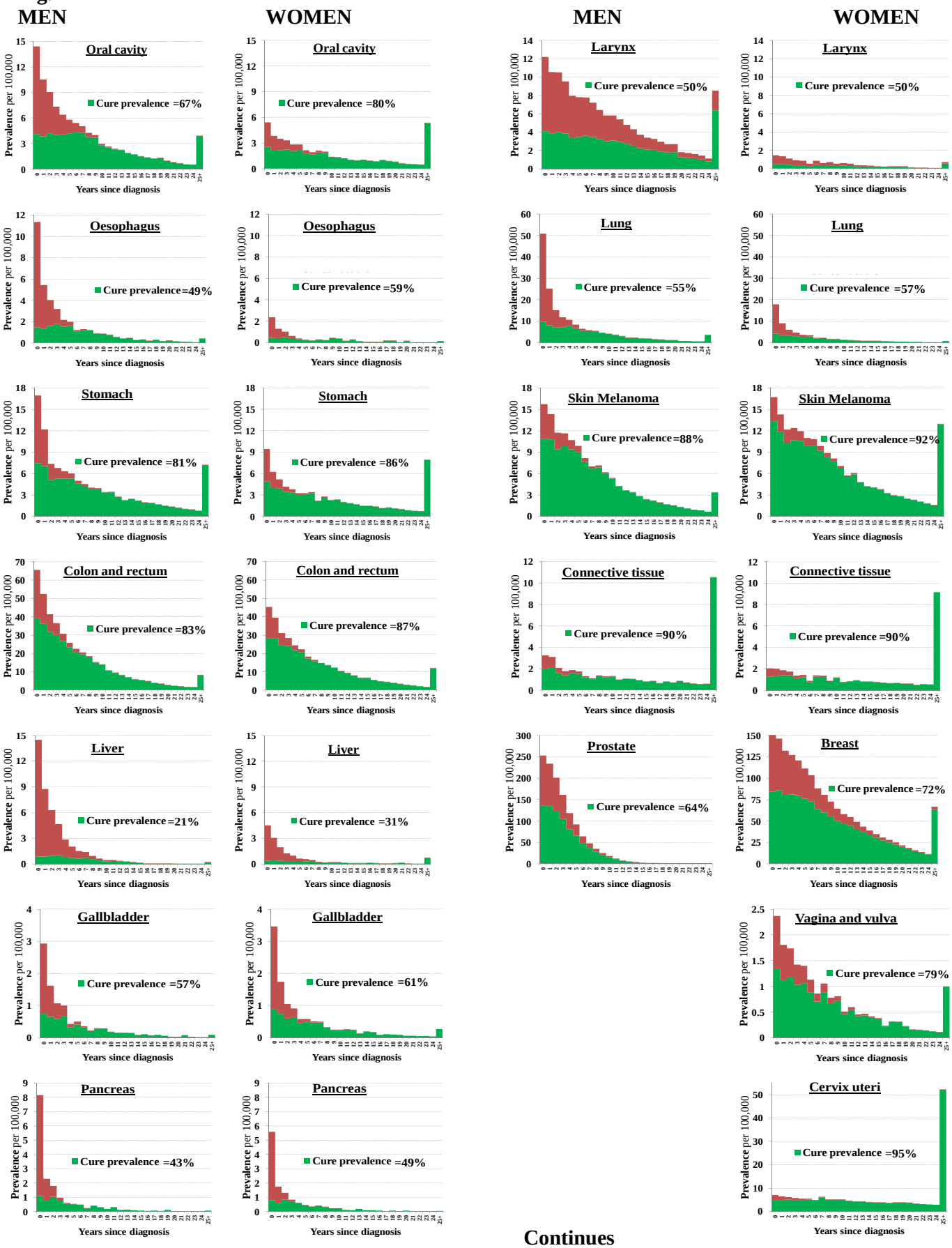
***Members of the AIRTUM Working group**: S. Virdone, A. Zucchetto (CRO Aviano National Cancer Institute of Aviano); A. Gigli (Istitute for Research on Population and Social Sciences, National Research Council, Rome); S. Francisci (National Institute of Health, Rome); P. Baili, G. Gatta (Foundation IRCSS, National Cancer Institute of Milan); M. Castaing (Integrated Cancer Registry of Catania-Messina-Siracusa-Enna, Catania), R. Zanetti (Piedmont Cancer Registry, Torino); P. Contiero (Varese Cancer Registry, Foundation IRCSS, National Cancer Institute of Milan); E. Bidoli (Friuli Venezia Giulia Cancer Registry); M. Vercelli (Liguria Cancer Registry, Genova); M. Michiara (Parma Cancer Registry); M. Federico (Modena Cancer Registry); G. Senatore (Salerno Cancer Registry); F. Pannozzo (Latina Cancer Registry); M. Vicentini (Reggio Emilia Cancer Registry); A. Bulatko (Alto Adige/Sudtirol Cancer Registry); D. R. Pirino (Sassari Cancer Registry); M. Gentilini (Trento Cancer Registry); M. Fusco (Napoli Cancer Registry); A. Giacomin (Biella Cancer Registry); A. C. Fanetti (Sondrio Cancer Registry); R. Cusimano (Palermo Cancer Registry).

Figure legend

Figure 1. Overall prevalence and cure prevalence (proportion of cancer patients who will not die of their disease)[1] overall and by year since diagnosis, sex, and cancer type. Italy 1985-2005

[1]At 1/1/2006 in patients aged 15-74 years living after a cancer diagnosis, calculated as sum of age specific estimates.

Fig.1



Continues

Fig. 1. continued

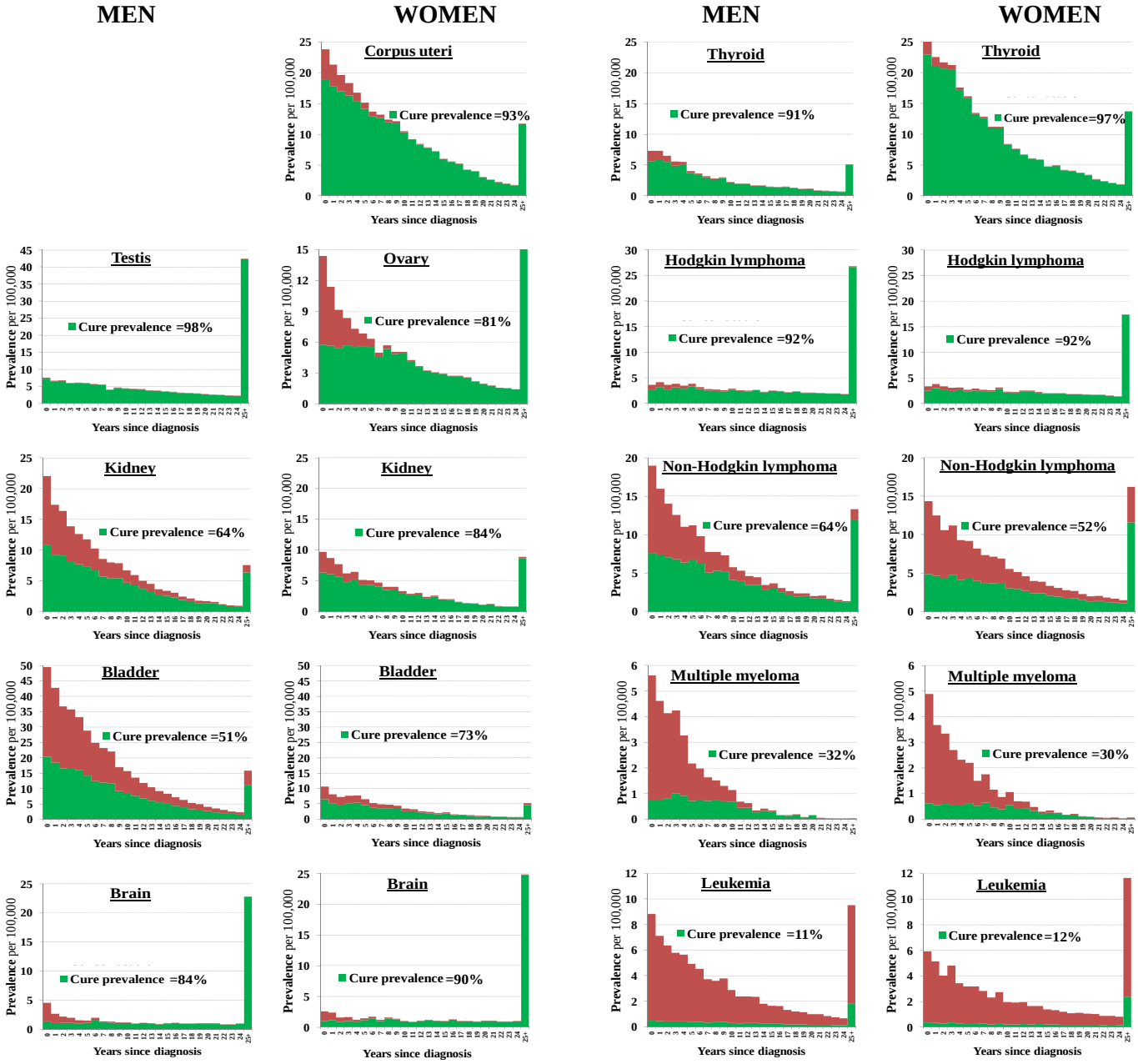


Table 1. Estimated cure fraction by cancer type, sex, and age at diagnosis.[1] Italy 1985-2005

Cancer type, ICD10	**Cure fraction (%)**							
	Men age (years)				**Women age (years)**			
	15-44	**45-54**	**55-64**	**65-74**	**15-44**	**45-54**	**55-64**	**65-74**
Oral cavity, C01-14	42	27	19	13	60	45	37	29
Oesophagus, C15	16	9	6	4	10	9	8	8
Stomach, C16	44	31	25	19	42	35	31	28
Colon and rectum, C18-21	54	50	48	45	57	53	51	49
Liver, C22	7	3	2	1	12	6	4	2
Gallbladder, C23-24	26	15	11	7	18	12	9	7
Pancreas, C25	6	4	3	2	17	7	4	2
Larynx, C32	49	39	34	29	49	39	34	29
Lung, C33-34	17	11	8	6	30	16	11	7
Skin melanoma, C43	77	67	61	54	85	78	73	68
Connective tissue, C47,C49	67	54	47	39	60	52	47	43
Breast, C50					40	60	54	56
Vagina and vulva, C51-52					75	60	51	41
Cervix uteri, C53					77	64	56	46
Corpus uteri, C54					91	83	76	66
Ovary, C56					62	40	28	17
Prostate, C61	47	49	50	50				
Testis, C62	94	90	86	83				
Kidney, C64-66,68	65	51	43	35	75	62	55	46
Bladder, C67, D09.0, D30.3, D41.4	81	63	50	35	89	75	63	48
Brain, C70-72	35	12	5	2	41	15	7	2
Thyroid, C73	95	75	60	43	99	94	84	65
Hodgkin lymphoma, C81	80	63	38	35	79	74	50	33
Non-Hodgkin lymphoma, C82-85,96	45	35	30	26	56	40	31	22
Multiple myeloma, C88-90	20	11	12	11	35	23	9	12
Leukaemia, C91-95	18	12	9	7	18	12	9	7

[1] Age 15-74 years.

Table 2. Estimates of time to cure according to different thresholds for 5-year and 10-year conditional relative survival (CRS)[1] by cancer type, sex and age. Italy 1985-2005

Cancer type, ICD10	Age at diagnosis[2] (years)	Men			Women		
		Years to 5-yr CRS		Years to 10-yr CRS	Years to 5-yr CRS		Years to 10-yr CRS
		>90%	>95%	>95%	>90%	>95%	>95%
Oral cavity, C01-14	15-44	7	9	10	6	9	10
	45-54	9	11	11	8	11	12
	55-64	9	11	12	9	12	13
	65-74	10	12	12	10	12	13
	15-74	*9*	*11*	*12*	*9*	*11*	*12*
Oesophagus, C15	*15-74*	*8*	*9*	*9*	*7*	*8*	*8*
Stomach, C16	15-44	5	7	7	5	7	7
	45-54	6	8	8	5	7	7
	55-64	6	8	9	5	7	8
	65-74	7	9	9	6	8	8
	15-74	*6*	*8*	*9*	*6*	*7*	*8*
Colon and rectum, C18-21	15-44	5	7	8	5	7	7
	45-54	5	8	8	5	7	7
	55-64	6	8	8	5	7	8
	65-74	6	8	9	5	7	8
	15-74	*6*	*8*	*9*	*5*	*7*	*8*
Liver, C22	*15-74*	*15*	*18*	*20*	*14*	*18*	*19*
Gallbladder, C23-24	*15-74*	*8*	*9*	*10*	*7*	*9*	*9*
Pancreas, C25	*15-74*	*5*	*6*	*6*	*5*	*6*	*6*
Larynx, C32	15-44	9	22	>25			
	45-54	14	>25	>25			
	55-64	16	>25	>25			
	65-74	19	>25	>25			
	15-74	*17*	*>25*	*>25*	*17*	*>25*	*>25*
Lung, C33-34	15-44	7	9	9	6	8	8
	45-54	7	9	9	7	8	9
	55-64	8	9	10	7	9	9
	65-74	8	10	10	8	9	10
	15-74	*8*	*9*	*10*	*7*	*9*	*9*
Skin melanoma, C43	15-44	3	6	6	1	5	7
	45-54	5	7	8	3	7	9
	55-64	6	8	9	4	9	10
	65-74	6	9	9	6	10	11
	15-74	*5*	*8*	*8*	*3*	*7*	*9*
Connective tissue, C47,C49	*15-74*	*6*	*9*	*10*	*6*	*10*	*11*
Breast, C50	15-44				14	>25	>25
	45-54				6	17	24
	55-64				9	19	25
	65-74				9	17	21
	15-74				*9*	*19*	*25*
Vagina and vulva, C51-52	*15-74*				*8*	*12*	*15*
Cervix uteri, C53	15-44				3	5	6
	45-54				5	7	7
	55-64				6	8	8
	65-74				6	8	9
	15-74				*5*	*7*	*8*

survival (CRS)[1] by cancer type, sex and age. Italy 1985-2005. Continued

Corpus uteri, C54	15-44				1	1	2
	45-54				2	4	5
	55-64				3	6	7
	65-74				5	7	8
	15-74				*4*	*6*	*7*
Ovary, C56	15-44				5	7	8
	45-54				7	9	9
	55-64				8	10	10
	65-74				9	11	11
	15-74				*8*	*10*	*10*
Prostate, C61	45-54	10	15	18			
	55-64	10	15	18			
	65-74	10	15	18			
	15-74	*10*	*15*	*18*			
Testis, C62	15-44	1	1	1			
	45-54	1	1	2			
	15-74	*1*	*1*	*1*			
Kidney, C64-66,68	15-44	3	8	15	2	5	8
	45-54	5	13	22	4	8	11
	55-64	7	16	>25	5	10	13
	65-74	10	20	>25	7	12	15
	15-74	*8*	*17*	*>25*	*5*	*10*	*14*
Bladder, C67, D09.0, D30.3, D41.4	15-44	1	1	8	1	1	2
	45-54	2	13	>25	1	5	12
	55-64	7	22	>25	3	11	19
	65-74	15	>25	>25	8	18	>25
	15-74	*10*	*>25*	*>25*	*5*	*14*	*23*
Brain, C70-72	*15-54*	*7*	*8*	*8*	*7*	*8*	*8*
Thyroid, C73	15-44				1	1	1
	45-54				1	1	1
	55-64				1	4	4
	65-74				3	5	5
	15-74	*1*	*9*	*14*	*1*	*1*	*1*
Hodgkin lymphoma, C81	*15-74*	*1*	*7*	*9*	*1*	*6*	*11*
Non-Hodgkin lymphoma, C82-85,96	15-44	8	15	22	4	14	25
	45-54	10	18	25	10	23	>25
	55-64	12	20	>25	14	>25	>25
	65-74	13	22	>25	18	>25	>25
	15-74	*11*	*20*	*>25*	*>25*	*>25*	*>25*
Multiple myeloma, C88-90	55-64	18	22	24	24	>25	>25
	65-74	17	20	21	17	21	22
	15-74	*17*	*20*	*21*	*17*	*21*	*22*
Leukaemia, C91-95	15-44	25	>25	>25	25	>25	>25
	45-54	>25	>25	>25	>25	>25	>25
	55-64	>25	>25	>25	>25	>25	>25
	65-74	>25	>25	>25	>25	>25	>25
	15-74	*>25*	*>25*	*>25*	*>25*	*>25*	*>25*

[1] Estimates are based on the relative survival function for each type and sex parameterized using mixture cure models.
'1' year is reported when time to reach thresholds is <one year; '>25'when threshold is not reached within 25 years.
[2] Estimates are shown by age groups when overall annual incidence rates>10/100.000.

Table 3. Complete prevalence (CP)[1] and proportion (%) of patients who reached different levels of conditional relative survival (CRS) by cancer type and sex. Italy 1985-2005

Cancer type, ICD10	**Men**				**Women**			
	CP	5-yr CRS	5-yr CRS	10-yr CRS	*CP*	*5-yr CRS*	5-yr CRS	10-yr CRS
	x 100.000	>90%	>95%	>95%	*x 100.000*	>90%	>95%	>95%
Oral cavity, C01-14	*99*	30%	23%	22%	*49*	43%	35%	32%
Oesophagus, C15	*13*	18%	15%	15%	*4*	29%	26%	26%
Stomach, C16	*109*	47%	38%	37%	*73*	58%	49%	47%
Colon and rectum, C18-21	*422*	41%	30%	27%	*352*	52%	40%	36%
Liver, C22	*46*	1%	1%	1%	*16*	9%	8%	7%
Gallbladder, C23-24	*10*	20%	14%	13%	*12*	26%	20%	20%
Pancreas, C25	*17*	18%	16%	16%	*13*	22%	19%	19%
Larynx, C32	*141*	15%	0%	0%	*14*	14%	0%	0%
Lung, C33-34	*174*	23%	18%	17%	*60*	21%	17%	15%
Skin melanoma, C43	*140*	54%	38%	36%	*184*	76%	49%	41%
Connective tissue, C47,49	*42*	67%	58%	55%	*35*	70%	59%	55%
Breast, C50					*1709*	41%	12%	6%
Vagina and vulva, C51-52					*19*	40%	25%	18%
Cervix uteri, C53					*167*	82%	75%	73%
Corpus uteri, C54					*258*	71%	56%	50%
Ovary, C56					*136*	51%	43%	42%
Prostate, C61	*592*	4%	0%	0%				
Testis, C62	*150*	95%	94%	93%				
Kidney, C64-66,68	*181*	39%	12%	2%	*97*	57%	36%	26%
Bladder, C67, D09.0, D30.3, D41.4	*438*	29%	5%	1%	*98*	58%	25%	10%
Brain, C70-72	*57*	58%	54%	53%	*56*	67%	63%	63%
Thyroid, C73	*74*	74%	58%	54%	*255*	87%	78%	78%
Hodgkin lymphoma, C81	*95*	85%	65%	59%	*78*	88%	68%	52%
Non-Hodgkin lymphoma, C82-85,96	*176*	31%	14%	4%	*159*	34%	7%	1%
Multiple myeloma, C88-90	*35*	2%	1%	0%	*29*	2%	1%	0%
Leukemia, C91-95	*87*	7%	0%	0%	*69*	13%	0%	0%

[1] At 1/1/2006 in patients aged 15-74 years living after a cancer diagnosis, calculated as sum of age specific estimates of prevalence.
[2] Estimates are based on the relative survival function for each type and sex parameterized using mixture cure models.

Supplementary Data: Extended description of Statistical Methods

1. *Patients*

   This study took advantage of data collected by a network of 23 population-based Italian cancer registries covering 28% of the Italian population, which included more than 12 million people aged 15-74 years. Routine indicators of data completeness and quality of Italian cancer registries were deemed satisfactory [26, additional Ref. 1]. Between 1985 and 2005, 454,527 cancer cases were diagnosed in men aged 15-74 years, including 85,053 lung, 63,047 prostate, and 56,635 colorectal cancers. Corresponding numbers in women were 364,375 cancer diagnoses, including 128,004 breast, 41,864 colorectal, and 20,398 endometrial cancers (Table S1).

2. *Relative Survival (RS)* is the ratio of the proportion of observed survivors in our cohort of cancer patients to the proportion of expected survivors in a comparable set of cancer free individuals. The formulation is based on the assumption of independent competing causes of death. The relative survival (RS) adjusts for the general survival of the Italian population for that Cancer Registry area, sex, age, and date at which the age was coded. Death certificates only, cases incidentally diagnosed at autopsy only, and cases not actively followed-up, were excluded from the analysis.

   The observed RS was calculated for cases diagnosed in 1985-2002 and followed-up until 2007 (until 18 years of incidence and 23 years of follow-up), using the cohort method and the Ederer II approach by means of SeerStat software. Nine cancer registries providing data for such a long time span (18 years of incidence and 23 years of follow-up) were eligible for this estimation, including 70% of all incident cases.

   Relative survivals were calculated by cancer type, sex, age at diagnosis (15–44, 45–54, 55–64, 65–74 years), and period of diagnosis (in three-years periods from 1985-1987 to 2000-2002) using the life tables provided by National Institute of Statistics for each CR area, broken down by age in years, sex, and calendar year.

   RS estimates may be biased when there is a high proportion of deaths due to a specific cancer in the external group [additional Ref. 2] However, at least in the age range of the present study (15-74 years), the impact specific cancer deaths, in population mortality figures, could have on the estimate of relative survival was so small, in comparison with total mortality, to make little difference in relative survival estimates.

   It has been shown that observed RS estimated according to Ederer II method [26] could be biased [additional Ref. 3], even when differences versus unbiased method are likely to be limited, in particular for ages below 75 years [additional Ref. 4].

3. *Conditional relative survival (CRS)* of a cohort of patients is the probability of surviving an additional y years, given that patients already survived x years. $CRS_{x,y}$ can be calculated as follows:

$$CRS_{x,y} = RS_{x+y} / RS_x \quad (1)$$

   In order to calculate standard error for the CRS, variance was derived by the delta method, as follows:

$$SE\ (CRS_{x,y}) = \frac{\sqrt{Var(RS_{x+y}) - CRS_{x,y}^2\ Var(RS_x)}}{RS_x} \quad (2)$$

4. *Model based relative survival* was calculated using mixture cure models which estimate both the proportion of patients who will reach the same life expectancy of the general population (i.e., C: the cure fraction or statistical cure) and the survival function of the 'not-cured' patients (1-C) [1,10].

   So the cumulative RS probability up to age *x* of a patient diagnosed at age *t* and year *y* is assumed to be:

$$RS(x,t,y) = [C + (1-C) * \exp(-\lambda(x-t)^{\gamma})]^{\exp[\beta_{1*(t-t_0)} + \beta_{2*(y-y_0)}]} \quad (3)$$

   where the follow-up time is x-t, λ and γ are, respectively, the scale and shape parameters of the Weibull distribution used to evaluate the specific differential mortality of fatal cases; β_1 and β_2 express the effects of 'age at diagnosis' and 'year of diagnosis' respectively, on the relative survival function of all patients.

   The model-based RS(x,t) was calculated from the cohort of incident cases in the period 1985-2002 followed up until 2007, by cancer type, sex, age at diagnosis (15–44, 45–54, 55–64, 65–74 years) and period of diagnosis (in three-years periods from 1985-1987 to 2000-2002). We used the simplest mixture cure model including relative risks for marginal age (β_1) and period effects (β_2) for any cancer type in men and women, assuming linearity (i.e., continuous terms using, as reference, the median for age classes (t_0=65 years) and the central year for periods ($y_{0=}$1992)) in both age and period effects. For the overall population (15-74 years), survival estimates were also calculated by sex at the average age at diagnosis for each cancer type.

Survival distribution was modeled according to the Weibull distribution for most cancer types. For breast and thyroid cancers, Hodgkin lymphoma, and myeloma, a better fit was obtained using separate models with exponential survival distribution for each age class. The goodness of fit of model-based and observed RS were evaluated by likelihood ratio tests and by visual comparison, [23] for each cancer type, sex, and age groups [10,11], showing consistent results.

Observed and model-based RS curves for the cohort diagnosed in 1991-1993 showed consistent results up to 15 years after diagnosis, for each cancer type, sex, and age group, in particular for common cancer types (i.e., breast, lung, colon-rectum, bladder, prostate) (S-Figure2). The observed and model-based RS for each additional year survived, conditioned on being alive at the beginning of each following year (conditioned relative survival, CRS) [13,26] were also compared and showed consistent results (S-Table2 shows observed and model-based five-year CRS estimates in cancer patients alive five years after diagnosis). The parameters C, λ, γ, β_1 and β_2 were then estimated using the SAS NLIN procedure.

In general, in the interpretation of the results of cure models it should be kept in mind that the existence of cured (i.e., supposed to die for causes other than their specific neoplasm) and uncured groups of patients is a necessary though strong assumption.

The cure models have several potential limitations [23]. In particular, they may not converge, or only fit poorly, when survival is either relatively good or relatively poor. We chose the simplest mixture cure model, although a more flexible modelling [additional Ref. 5] could be used. In particular, flexible parametric survival models use splines to model the underlying hazard function; therefore, no parametric distribution must be specified [additional Ref. 6]. Advantages include greater modelling flexibility with respect to the shapes of the survival distributions and greater sensitivity to small excess risk. However, it is also potentially sensitive to the choice of the number and location of knots.

Our assumption of linearity for the variables of age and year at diagnosis in the mixture cure fraction model is questionable for all 26 cancer sites and ages. For this reason, all analyses were restricted to ages 15-74 and, when appropriate, for selected neoplasms we used separate (better fitting) models with exponential survival distribution stratified by age class. In addition, the goodness of fit of model-based RS was evaluated by likelihood ratio tests and by visual comparison.

Unfortunately, criteria to select the better model are still debated and methods for model diagnostics are still needed for future methodological research [23].

5. *Complete Prevalence* estimation was based on observed prevalence and, for the period before the start of registration, by an adjustment (i.e., completeness index) that allows the estimation of the fraction of complete prevalence not observed in the data [1,12].

   Observed prevalence was directly computed from incidence and life status data collected by registries on their target population. The observed prevalence is necessarily incomplete, since it refers to the number and proportion of cases diagnosed after the start of the registry activity. Consequently, the observed prevalence of Italian registries represents a percentage of the complete prevalence so much greater as longer is their observation time; only registries with 40-50 years of follow-up can detect a virtually complete observed prevalence. Italian registries started in different years from 1978 onwards and their observation periods range from 7 to 28 years at the index date of December 31, 2005.

   Estimation of complete prevalence was based partly on observed prevalence and partly, for the period before the start of registration, by modelling a quantity, called the completeness index, that allows the estimation of the fraction of complete prevalence not observed in the recorded data. These indices vary based on the length of the registration period and are specific for tumour sites, since they depend on the incidence of neoplasia as well as on survival.

   Complete prevalence at age x is composed of all incident cases diagnosed at any age t (t<x) surviving up to age x, thus for x-t years. A population registry active since L years can observe only prevalent cases with a disease duration lower than x-L.

   The expected complete prevalence Prev(x) can therefore be broken down into two components, one observed (durations from 0 to x-L) and one unobserved (durations from x-L and x), according to the following relationship:

   $Prev(x) = Prev_L^{obs}(x) + Prev_L^{unobs}(x) = \sum_{t=x-L}^{x} I(t)RS(t,x-t) + \sum_{t=0}^{x-L-1} I(t)RS(t,x-t)$ (4)

   where I(t) is the incidence of the disease at age t and RS(t,x-t) is the relative survival at age x for a person who has been diagnosed with cancer at age t.

The completeness index R_L is defined as the ratio between the expected observed prevalence $Prev_L^{obs}$(x) and the expected complete prevalence Prev(x) for a specific cancer and observation time L, that is:

$$R_L(x) = \frac{Prev_L^{obs}(\mathrm{x})}{\mathrm{Prev(x)}} = \frac{\sum_{t=x-L}^{x} I(t)RS(t,x-t)}{\sum_{t=0}^{x} I(t)RS(t,x-t)} \quad (5)$$

Both survival and incidence functions in the previous equation must be estimated from available data of cancer registries. A systematic estimation was performed of the completeness indices by cancer site using the most current AIRTUM dataset available [1] and the COMPREV software. The pool of nine cancer registries with longer common registration periods was used to this purpose. Example levels of Completeness index by sex, age, and length of observation period (L) have been reported elsewhere [1, page 43]

The incidence function describes the relationship between age and risk of developing and being diagnosed of cancer, as measured along the life span of each birth cohort present in the population at the prevalence date. In the present study, a sixth degree polynomial on age was used for each site:

$$\mathrm{I(x,k)} = \left\{ 1 + e^{-\left[(a_k + \sum_{i=1}^{6} b_i \left(\frac{x-x_0}{m} \right)^i \right]} \right\}^{-1} \quad (6)$$

where I(x,k) is incidence at age x for birth cohort k, the birth cohort covariate was also included, together with age, in the incidence function as a categorical variable to adjust for risk trends across the different birth cohorts.

Parameters of the incidence function were estimated through the SAS logistic procedure by fitting raw incidence rates of patients registered between 1985 and 2004 by the nine Italian CRs in the Pool trend. Incidence data were categorized according to cancer site, gender, five-year age, and birth cohort (<1899, 1900-1904,…, 2000-2004). The goodness of fit of the various incidence models was assessed by Akaike Information Criterion (AIC) as well as by visual comparison between estimated and observed rates [1].

The RS function for each site and gender was parameterized as described before (point 4).

Indicators of long-term survival and cure of cancer

Cure is observable for most cancer types. Not all patients die of cancer and the cured ones, who have the same death risk of the general population, remain alive or die as a result of other causes. When this occurs the RS curve reaches a plateau that represents both the long term RS value and the proportion of cured patients, i.e. the *cure fraction* (Figure S1A). At the same time, the CRS approaches 100%, that observed in the general population who is free from cancer (Figure S1B).

Time to cure is estimated by means of model-based CRS curves assuming that a statistical cure is approached or reached when five-year CRS becomes higher than 90% or 95%. To test the sensitivity of time to cure definition to these assumptions, we considered a more restrictive definition, i.e. time when ten-year CRS becomes higher than 95%. In particular, the following indicators were calculated.

6. *Cure Prevalence (CPrev)* represents the proportion of cancer patients who will not die as a result of cancer. This indicator was computed through the model-based RS (Figure S1C). The overall number of these cured prevalent patients divided by the complete prevalence is called *cure prevalence* [14]. The number of estimated cured prevalent cases at age x that were diagnosed at age t is given by:

$$\mathrm{CPrev_t(x)=C_t * Prev_t(x) / RS_t\ (x\text{-}t)} \quad (7)$$

where C_t is the cure fraction, $Prev_t(x)$ is the complete prevalence at age x of patients diagnosed at age t and RS_t (x-t) is the relative survival of patients diagnosed at age t and follow up time x-t (3). For each cancer type and sex, the overall cure prevalence was calculated summing up the age specific estimates (7) over all age at prevalence (x) and follow up time up (x-t):

$$\mathrm{CPrev} = \frac{\sum_{x=15-44}^{65-74} \sum_{\mathrm{x=0}}^{\mathrm{n}} \mathrm{CPrev_t\ \ (x)}}{Prev_{TOT}} \quad (8)$$

where *n* is up to the maximum duration of time since diagnosis and $Prev_{TOT}$ is the overall complete prevalence for all age classes considered (15-74).

7. *Years to reach 5-year or 10 years CRS>k (YtoCRS>k)* represents the number of years necessary for x-year CRS (x=5 or 10) to reach a fixed clinically or epidemiologically relevant threshold (k) (k=90% or 95%).

8. *Proportion of patients who reached different levels of YtoCRS>k* was calculated as a sum of prevalent patient by a period since diagnosis > k years on the complete prevalence

$$Prev_{(\mathrm{YtoCRS}>k)} = \frac{\sum_{t=15-44}^{65-74} \sum_{x=k+1}^{n} \mathrm{Prev_t} \quad (x>k)}{Prev_{TOT}} \tag{9}$$

Estimates of YtoCRS>k were calculated using age at diagnosis of patients while $\mathrm{Prev_t}$ is based on the index age of prevalent cases. This simplified assumption could lead to a slight underestimation of indicator (9) since the time to reach negligible excess mortality increased with age for most of cancer types, and it was applied to complete prevalence of more advanced (reference) ages.

The proportion of cured among prevalent patients increased with time since diagnosis, and could include patients surviving either less or longer than time to cure. Only prevalent cases surviving longer than time to cure are identifiable to be likely cured from cancer. They are a subset of the overall cure prevalence, varying according to the definition of time to cure (Figure S1C, blue and green rectangles).

Additional References

1. Forman D, Bray F, Brewster DH, et al. eds. *Cancer Incidence in Five Continents*, Vol. X (electronic version). Lyon: IARC; 2013. http://ci5.iarc.fr Last accessed on 4/16/2014. (The printed version of this volume: IARC Scientific Publications No. 164, will be available in 2014.)
2. Hinchliffe SR, Dickman PW, Lambert PC. Adjusting for the proportion of cancer deaths in the general population when using relative survival: a sensitivity analysis. Cancer Epidemiol 2012; 36:148-52.
3. Pohar Perme M, Stare J, Esteve J. On Estimation in Relative Survival. *Biometrics*. 2012;68(1):113–120.
4. Dickman PW, Lambert PC, Coviello E, Rutherford MJ. Estimating net survival in population-based cancer studies. *Int J Cancer*. 2013;133:519-521.
5. Andersson TM, Dickman PW, Eloranta S, Lambert PC. Estimating and modelling cure in population-based cancer studies within the framework of flexible parametric survival models. BMC Med Res Methodol. 2011;11:96
6. Lambert PC, Dickman PW, Weston CL, Thompson JR. Estimating the cure fraction in population-based cancer studies by using finite mixture models. Applied Statistics. 2010; 59:35-55