

MATERIALI E METODI
MATERIALS AND METHODS

CAPITOLO 6

e&o

Capitolo 6

Materiali e metodi

Materials and methods

Introduzione

L'Associazione italiana registri tumori

L'Associazione italiana registri tumori, costituitasi formalmente nel 1997, denominata AIRT fino al 2006 e successivamente AIRTUM, coordina l'attività e i programmi di ricerca dei registri tumori di popolazione, sia generali sia specializzati. Nel 2005 l'Associazione si è dotata di una banca dati informatica che raccoglie i dati dei registri tumori, li archivia e, dopo averne verificato qualità e completezza, li utilizza per studi collaborativi di tipo descrittivo e attività di ricerca sull'epidemiologia oncologica in Italia.¹⁻¹³

L'attività di registrazione si è sviluppata nel nostro Paese a partire dagli anni Settanta con una costante crescita della parte di popolazione italiana interessata dalla rilevazione dei tumori. I registri di popolazione generali erano 3 nei primi anni Ottanta, 12 nei primi anni Novanta, 21 nei primi anni Duemila e attualmente sono 29. A questi si aggiungono 5 registri specializzati, per età o per tipologia tumorale (figura 1).

La concentrazione dei registri è maggiore nel Nord, soprattutto nel Nord-Est, rispetto al Centro e soprattutto al Sud Italia (tabella 1). D'altra parte è proprio nel Sud Italia che in anni recenti si è avuto un maggior sviluppo di nuovi registri, che ha permesso di documentare con maggior dettaglio e miglior

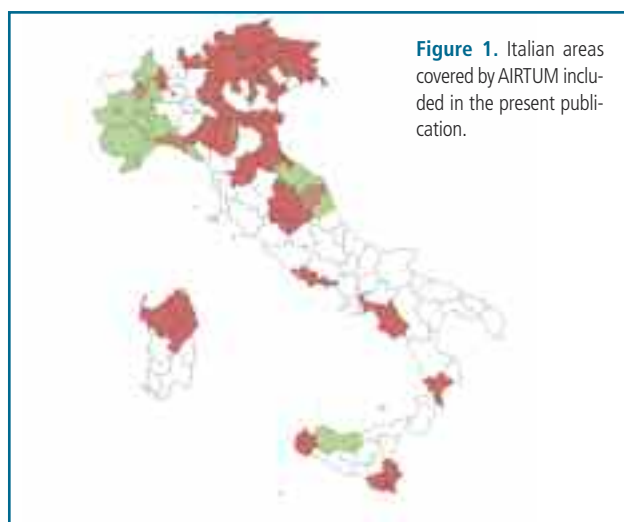


Figure 1. Italian areas covered by AIRTUM included in the present publication.

Introduction

The Italian Network of Cancer Registries

Formally established in 1997, the Italian Network of Cancer Registries, AIRTUM (called AIRT until 2006) promotes and supports activities and research programmes for both general and specialized population-based Italian cancer registries. AIRTUM has developed a central database to collect and store cancer registry data and to make them available for collaborative studies and research activities after an official quality check in terms of data accuracy and completeness.

Cancer registration in Italy began in the 1970s and has steadily developed in terms of experience and numbers, with steadily increasing coverage of the Italian population.

In the early 1980s there were only 3 general cancer registries in Italy; their number grew to 12 in the early 1990s and 21 in the 2000s. Today Italy has 29 general cancer registries. Five age- or cancer-specific cancer registries are also operational (Figure 1). The density of registries is greater in northern Italy, especially in the North-East compared to central and southern Italy (Table 1). On the other hand, cancer registration in the South of Italy has remarkably expanded in recent years, providing greater detail and a better representation of cancer disease in this area: eight general cancer registries and one specialized cancer registry are now active in southern Italy.

Regional coverage varies from 0% for several regions (Puglia, Basilicata, Abruzzo, Molise, and Valle d'Aosta) up to 100% (Umbria, Friuli Venezia Giulia, Trento, and Bolzano).

AIRTUM registries cover more than 19 million subjects, representing 34% of the entire Italian population (38% in the North-West, 68% in the North-East, 26% in the Centre and 18% in the South).

In addition to the general cancer registries, which collect data from the entire population and for all types of cancer, there are also two specific childhood (ages 0 to 14) and adolescence (ages 15-19) registries, one in Piedmont and another in the Marche, which cover about 700,000 children and 250,000 adolescents. There are also three cancer registries for specific cancer types: the colorectal cancer registry of Modena (covering a population of about 270,000), the female breast cancer registry of Palermo (covering a female population of about 640,000)

rappresentatività la patologia oncologica anche in quest'area. Oggi nel Sud Italia sono attivi 8 registri tumori generali e uno specializzato.

Nella tabella 1 è presentata la quota di popolazione monitorata da AIRTUM per singola Regione e per macroaree geografiche (Nord-Ovest, Nord-Est, Centro, Sud). La diversa presenza e la diversa dimensione dei registri determinano una variazione nella copertura regionale che varia dallo 0% per alcune Regioni (Puglia, Basilicata, Abruzzo, Molise e Val d'Aosta), fino al 100% (Umbria, Friuli Venezia Giulia e le province autonome di Trento e Bolzano). Nel Sud Italia si è avuta in anni recenti una notevole estensione delle aree coperte e nuovi registri si stanno sviluppando anche in alcune delle Regioni che ancora ne sono sprovviste.

Oggi oltre un terzo dei cittadini italiani risiede in un'area dove è attivo un registro tumori di popolazione, con percentuali che variano da un'area all'altra del Paese (38% nel Nord-Ovest, 68% nel Nord-Est, 25% nel Centro e 18% al Sud). Complessivamente i registri dell'AIRTUM monitorano oltre 19.000.000 di italiani, pari al 34% della popolazione residente totale.

Oltre ai registri di popolazione (che raccolgono informazioni sui tumori dei cittadini di tutte le fasce di età) sono attivi in Italia due registri specialistici per i tumori infantili e adolescenziali, uno in Piemonte e l'altro nelle Marche, che nel complesso seguono quasi 700.000 residenti tra 0 e 14 anni e circa 250.000 residenti tra i 15 e i 19 anni.

Fanno parte dell'AIRTUM anche tre registri specializzati in singole patologie tumorali: il registro dei tumori coloretali di Modena (circa 270.000 residenti), quello del tumore della mammella femminile di Palermo (circa 640.000 donne residenti) e il registro dei mesoteliomi della regione Liguria (circa 1.600.000 residenti). Quest'ultimo rappresenta l'unico Centro operativo regionale (COR) afferente al registro nazionale dei mesoteliomi (ReNaM) ad avere richiesto e ottenuto l'accreditamento presso la banca dati AIRTUM.

La banca dati AIRTUM

Dal 2005 AIRTUM ha costituito formalmente una banca dati che raccoglie le informazioni prodotte dai singoli registri. Grazie alla realizzazione della banca dati, che rientra tra gli obiettivi statutari dell'Associazione, AIRTUM ha potuto formalizzare la tipologia dei dati raccolti. Tutti i registri tumori trasferiscono alla banca dati le informazioni secondo un tracciato record standard, che ne permette il continuo aggiornamento. I singoli registri possono inviare nuovi dati e integrare i dati già presenti, incluso l'aggiornamento del follow-up e dello stato in vita.

Ogni registro che contribuisce alla banca dati rimane titolare dei propri dati. Per utilizzare tali dati in studi collaborativi è necessaria una specifica autorizzazione da parte dei direttori dei singoli registri.

Nel sito dell'Associazione è disponibile l'attuale protocollo operativo (<http://www.registri-tumori.it/cms/files/Protocollo2010.pdf>) al quale si rimanda per una descrizione dettagliata dei criteri di inclusione nella banca dati, della tipologia e delle modalità dei controlli e dei tracciati record.

and the regional mesothelioma cancer registry of Liguria (covering a population of about 1.6 million), which is the only Regional Operating Centre from the National Register of Mesothelioma (ReNaM) granted accreditation by AIRTUM.

The database

Since 2005, AIRTUM has a central database which stores the data from all accredited registries.

Thanks to the implementation of this database, which is one of the main formal aims of the network, AIRTUM was able to define the type of data to be collected. All cancer registries transfer data to the AIRTUM database following a standard protocol, performing regular updates over time. Each registry can send new data or update old records with new variables, including follow-up for vital status. Each registry is the owner of its own data, so specific approval from registries is required for collaborative studies involving cancer registry data.

The network has a specific protocol, with a description of including criteria, data format, and checks (for detailed information see <http://www.registri-tumori.it/cms/files/Protocollo2010.pdf>).

Region	Italian population 2001	Residents in areas covered by general cancer registries	
	No.	No.	%
Piemonte	4 214 677	1 052 512	25.0
Valle d'Aosta	119 548	0	0.0
Lombardia	9 032 554	3 732 110	41.3
Liguria	1 571 783	878 082	55.9
North-West	14 938 562	5 662 704	37.9
Trentino-Alto Adige	940 016	940 016	100.0
Veneto	4 527 694	2 215 376	48.9
Friuli-Venezia Giulia	1 183 764	1 183 764	100.0
Emilia-Romagna	3 983 346	2 923 606	73.4
North-East	10 634 820	7 262 762	68.3
Toscana	3 497 806	1 161 746	33.2
Umbria	825 826	825 826	100.0
Marche	1 470 581	301 523	20.5
Lazio	5 112 413	491 230	9.6
Centre	10 906 626	2 780 325	25.5
Abruzzo	1 262 392	0	0.0
Molise	320 601	0	0.0
Campania	5 701 931	1 611 503	28.3
Puglia	4 020 707	0	0.0
Basilicata	597 768	0	0.0
Calabria	2 011 466	234 676	11.7
Sicilia	4 968 991	1 116 552	22.5
Sardegna	1 631 880	718 487	44.0
South	20 515 736	3 681 218	17.9
ITALY	56 995 744	19 387 009	34.0

Table 1. Distribution of the Italian resident population overall and in areas covered by general cancer registries (AIRTUM), by region and macroarea. Italy, 2001.

Definizioni

Dati richiesti

Le analisi di sopravvivenza hanno riguardato tutti i casi incidenti dei registri tumori a partire dall'anno di diagnosi 1990 fino al 2007. A tutti i registri tumori partecipanti è stato richiesto l'aggiornamento dello stato in vita al 31.12.2008. Per il solo anno di diagnosi 2004 (ultimo anno dello studio di coorte) e limitatamente ai registri di Catanzaro e Nuoro, attivi solo dall'anno di registrazione 2003, è stato richiesto un follow-up esteso al 31.12.2009 al fine di poter calcolare una stima della sopravvivenza a 5 anni dalla diagnosi. I campi oggetto di analisi sono stati: data di nascita, genere, data di incidenza, topografia ICD-O-3, morfologia ICD-O-3, stato in vita, data di ultimo stato in vita e modalità di diagnosi (o base diagnosi). I valori possibili per quest'ultima variabile sono:

- **Codice 0.** L'informazione disponibile viene da un certificato di morte (DCO, Death Certificate Only);
- **Codice 1.** Diagnosi effettuata prima della morte, ma senza alcuno dei codici 2-7;
- **Codice 2.** Tutte le tecniche diagnostiche senza successivo esame istopatologico, chirurgia esplorativa;
- **Codice 3.** Diagnosi autoptica solo macroscopica;
- **Codice 4.** Diagnosi effettuata tramite marcatori biochimici e/o immunologici specifici per tumore;
- **Codice 5.** Esame citologico di cellule prelevate da una sede tumorale primitiva o secondaria, include l'esame microscopico del sangue periferico e degli aspirati midollari;
- **Codice 6.** Esame istologico effettuato su metastasi, ivi compresi i campioni autoptici;
- **Codice 7.** Esame istologico di tessuto di tumore primitivo;
- **Codice 8.** Autopsia con istologia concomitante o precedente;
- **Codice 9.** Modalità sconosciuta.

Pool dei registri

Al fine di garantire sia la massima rappresentatività geografica sia l'omogeneità interna della casistica analizzata sono stati costituiti tre raggruppamenti diversi (o pool) per tipo di analisi:

- **Analisi di coorte**, che consiste nell'osservazione della sopravvivenza manifestata da un gruppo di pazienti dal momento della diagnosi e per un determinato periodo di tempo (generalmente 5 anni).
- **Analisi del trend**, che mira a raffrontare la sopravvivenza dei pazienti nel corso di diversi periodi di tempo.
- **Analisi condizionata**, che consiste nell'osservazione della sopravvivenza manifestata dai pazienti nell'ipotesi che siano già sopravvissuti alla malattia per un definito periodo di tempo dalla diagnosi.

L'analisi di coorte ha coinvolto 29 registri generali con disponibilità di dati di incidenza nel periodo 2000-2004, a cui si aggiungono i registri specializzati della mammella di Palermo e dei mesoteliomi della Liguria.

L'analisi del trend ha riguardato i dati nel periodo 1990-2007: sono stati considerati i registri con disponibilità di dati almeno nel periodo 1992-2005 per un totale di 11 registri generali.

Definitions

Data

Survival analysis included data from each cancer registry for all tumours in cancer patients diagnosed from 1990 to 2007. Follow-up for death to the end of 2008 was required for all cancer registries. Data from the Catanzaro and Nuoro cancer registries, which only started up in 2003, were supplemented by an additional year of follow-up for death, limited to diagnosis 2004 (i.e., the last calendar year of the cohort study) in order to extend the calculation of five-year survival estimates to 2009 as year of diagnosis. The following items were used in the analysis: date of birth, gender, date of diagnosis, cancer topography and morphology (according to the ICD-O-3 classification), vital status, date of last ascertainment of vital status and basis of diagnosis, defined as follows:

- **Code 0.** *The available information comes from a death certificate only (i.e., DCO);*
- **Code 1.** *Diagnosis before death but with no 2-7 codes;*
- **Code 2.** *All diagnostic techniques without subsequent histological examination, exploratory surgery;*
- **Code 3.** *Autopsy without histological examination;*
- **Code 4.** *Diagnosis made with cancer-specific biochemical and/or immunological markers;*
- **Code 5.** *Cytological examination of primary or metastatic tumour cells, including peripheral blood and bone marrow aspirates;*
- **Code 6.** *Histological examination performed on metastasis, including autopsy specimens;*
- **Code 7.** *Histological examination of primary tumour;*
- **Code 8.** *Autopsy with concurrent or previous histology;*
- **Code 9.** *Unknown method of diagnosis.*

Pools of cancer registries

In order to provide data consistency and maximize the use of descriptive data, three different groups (pools) of cancer registries were defined according to type of analysis:

- **Cohort analysis**, *which consists in the study of survival experiences observed in a group of patients starting from the date of diagnosis and for a fixed period of time;*
- **Time trend analysis**, *focused on comparison of survival over a defined calendar period;*
- **Conditional analysis**, *which consists in the study of additional survival experiences for a group of patients, based on the assumption that all of them have already survived for a certain period of time.*

Cohort analysis included 29 general and two specialized (breast and mesothelioma) cancer registries which provided data from cancer patients diagnosed in 2000-2004. Time trend analysis included 11 general cancer registries which provided data from cancer patients diagnosed in 1990-2007 or, at least, 1992-2005. Finally, conditional analysis included 20 general cancer registries and the mesothelioma registry, which provided data for at least 10 consecutive years of diagnosis within the period 1996-2007.

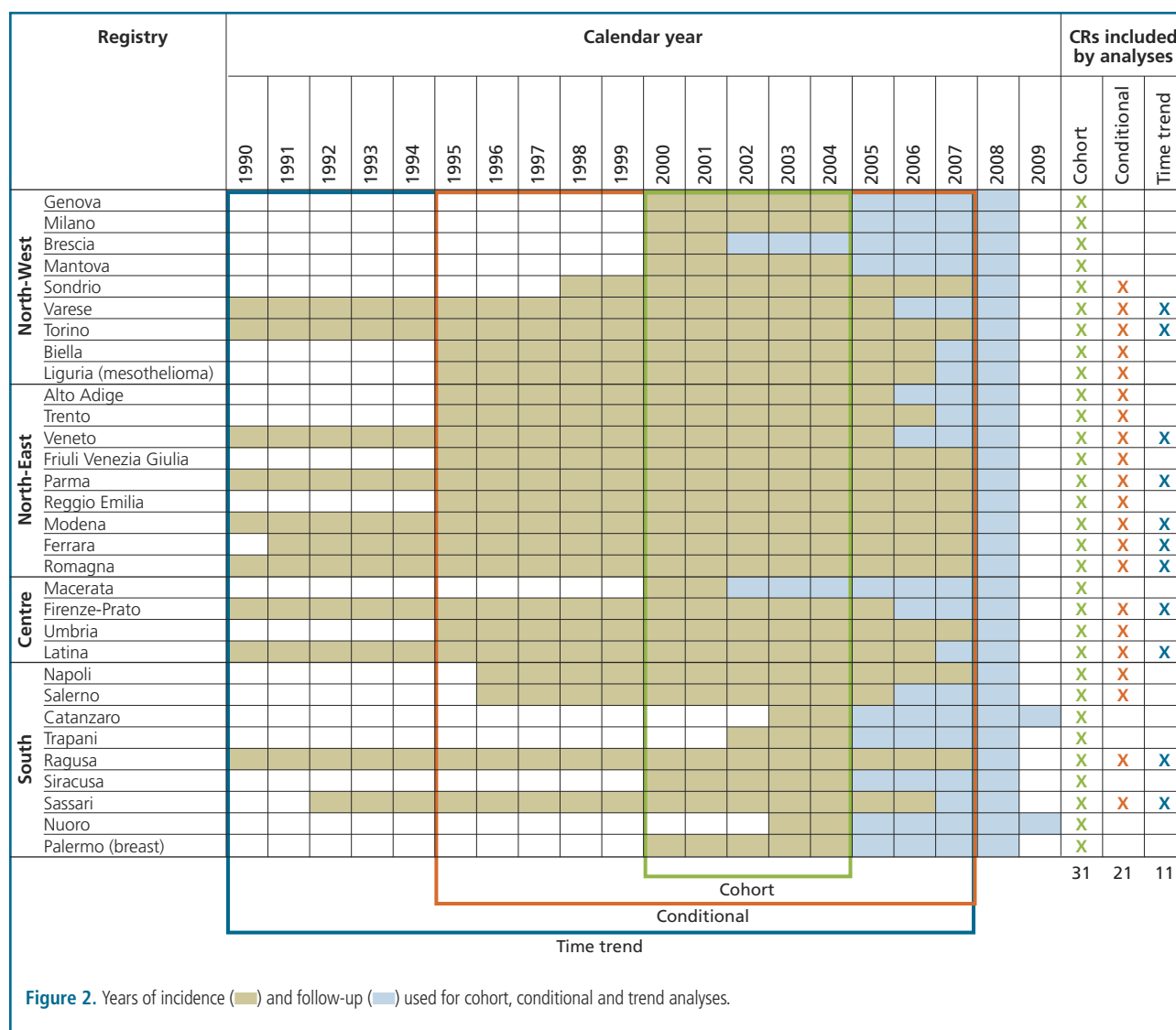


Figure 2. Years of incidence (■) and follow-up (■) used for cohort, conditional and trend analyses.

L'analisi condizionata, infine, ha riguardato i dati nel periodo 1996-2007: sono stati coinvolti i registri con disponibilità di dati per almeno 10 anni in quella finestra temporale, per un totale di 20 registri generali, più il registro dei mesoteliomi della Liguria.

In figura 2 si riportano con dettaglio le composizioni di questi tre pool.

Rispetto al precedente che ha analizzato la sopravvivenza,¹ il presente rapporto AIRTUM è caratterizzato da una maggior numerosità della casistica analizzata e una migliore rappresentatività geografica: i registri generali indagati nell'analisi di coorte 2000-2004 sono passati da 21 a 29 (in particolare quelli attivi nel meridione da 4 a 8), i registri generali indagati nell'analisi del trend da 7 a 11.

Accorpamento per macroarea

Sulla base della definizione di macroarea fornita da ISTAT per il territorio italiano, sono definiti i seguenti accorpamenti di registri: Nord-Ovest, Nord-Est, Centro, Sud e Isole (qui indicato semplicemente come Sud).

A detailed picture of these groups is given in Figure 2.

Compared with the previous AIRTUM publication on survival,¹ both cohort size and geographic coverage have grown: the number of general cancer registries included in 2000-2004 cohort studies increased from 21 to 29 (registries from southern Italy, in particular, increased from 4 to 8), while registries included in the time trend analysis increased from 7 to 11.

Geographical groups

The following geographical areas (macroareas) are defined according to the partition of the country of the Italian Statistical Institute (ISTAT): North-West, North-East, Centre, South-Islands (here designated simply as South).

Cancer sites

Classification

Topography, morphology, and tumour behaviour are defined according to the third edition of the International Classification of Diseases for Oncology (ICD-O-3), edited in 2000.² The criteria adopted for defining the cancer sites are shown in Table 2.

Major site	Cancer site	Topography	Morphology	Behav.	
Head and neck	Tongue	C019-C029	8000-9044, 9060-9133, 9150-9581	3	
	Mouth	C030-C069	8000-9044, 9060-9133, 9150-9581	3	
	Oropharynx	C090-C109	8000-9044, 9060-9133, 9150-9581	3	
	Nasopharynx	C110-C119	8000-9044, 9060-9133, 9150-9581	3	
	Hypopharynx	C129-C139	8000-9044, 9060-9133, 9150-9581	3	
	Pharynx, NOS	C140-C148	8000-9044, 9060-9133, 9150-9581	3	
	Nasal cavity etc.	C300-C319	8000-9044, 9060-9133, 9150-9581	3	
	Larynx	C320-C329	8000-9044, 9060-9133, 9150-9581	3	
	Oesophagus	C150-C159	8000-9044, 9060-9133, 9150-9581	3	
	Stomach	C160-C169	8000-9044, 9060-9133, 9150-9581	3	
Colon rectum	Colon	C180-C189	8000-9044, 9060-9133, 9150-9581	3	
	Rectum	C199-C218	8000-9044, 9060-9133, 9150-9581	3	
	Liver	C220-C221	8000-9044, 9060-9133, 9150-9581	3	
	Biliary tract	C239-C249	8000-9044, 9060-9133, 9150-9581	3	
	Pancreas	C250-C259	8000-9044, 9060-9133, 9150-9581	3	
	Lung	C339-C349	8000-9044, 9060-9133, 9150-9581	3	
	Bone	C400-C419	8000-9044, 9060-9133, 9150-9581	3	
	Skin melanoma	C440-C449	8720-8780	3	
	Mesothelioma	C000-C750	9050-9053	3	
	Kaposi sarcoma	C000-C809	9140	3	
	Soft tissue	C470-C479, C490-C499	8000-9044, 9060-9133, 9150-9581	3	
	Breast	C500-C509	8000-9044, 9060-9133, 9150-9581	3	
	Cervix uteri	C530-C539	8000-9044, 9060-9133, 9150-9581	3	
	Corpus uteri	C540-C549	8000-9044, 9060-9133, 9150-9581	3	
	Ovary	C569-C569	8000-9044, 9060-9133, 9150-9581	3	
	Prostate	C619-C619	8000-9044, 9060-9133, 9150-9581	3	
	Testis	C620-C629	8000-9044, 9060-9133, 9150-9581	3	
	Kidney etc.	C649-C669, C680-C689	8000-9044, 9060-9133, 9150-9581	3	
	Urinary bladder	C670-C679	8000-9044, 9060-9133, 9150-9581	3	
		C670-C679	8000-8010, 8120-8131	0-2	
	Brain and CNS	C700-C729	8000-9044, 9060-9133, 9150-9581	3	
	Choroid melanoma	C690-C699	8720-8780	3	
	Thyroid	C739-C739	8000-9044, 9060-9133, 9150-9581	3	
	Hodgkin lymphoma	C000-C809	9650-9667	3	
	Non Hodgkin lymphoma	C000-C419, C422-C423, C440-C809	9823, 9827	3	
		C000-C809	9590-9596, 9670-9729	3	
	Myeloma	C000-C809	9731-9732, 9734-9734	3	
	Leukaemia	Acute lymphocytic leukaemia	C000-C809	9826, 9835-9837	3
			C420-C421, C424	9827	3
		Chronic lymphocytic leukaemia	C420-C421, C424	9823	3
		Acute myeloid leukaemia	C000-C809	9840, 9861, 9866-9874, 9891-9920, 9931	3
		Chronic myeloid leukaemia	C000-C809	9863, 9875-9876, 9945-9946	3
		Leukaemia, NOS	C000-C809	9733, 9742, 9800-9820, 9832-9834, 9860, 9930	3

Table 2. Tumour sites considered in the present report according to ICD-O-3 topography, morphology, and behaviour.

Sedi tumorali

Classificazione

La definizione della sede anatomica, della morfologia tumorale e del comportamento si basa sulla terza edizione dell' *International Classification of Diseases for Oncology* (ICD-O-3) pubblicata nel 2000.² I criteri adottati per la definizione delle sedi tumorali analizzate sono esplicitati in tabella 2.

Criteri di selezione

Nell'analisi sono stati inclusi tutti i tumori primitivi maligni, a eccezione dei carcinomi cutanei. A differenza dalle passate monografie curate da AIRTUM, così come da quelle prodotte nell'ambito dei protocolli EURO CARE, vengono

Data selection

All primary malignant tumours were included in the analysis, except non-melanomatous skin carcinomas. Thus multiple primaries were considered eligible for the analysis, in contrast with previous AIRTUM and EURO CARE protocols.^{3,4} Cancers notified by death certificate only or by autopsy only (i.e., with date of diagnosis equal to date of death) were excluded. All other cases identified very close to the date of death were included in the data analysis, including patients with autopsy made after clinical diagnosis. Survival was computed for all patients aged 15-99. Separate analyses were performed for children (ages 0-14) and adolescents (ages 15-19).

The inclusion of multiple tumours in the analysis implies that

quindi considerati eleggibili per uno stesso soggetto anche i tumori successivi al primo.^{3,4} Sono stati invece esclusi i casi notificati dal solo certificato di decesso (base di diagnosi 0 -DCO) o autopsia (con diagnosi coincidente con la data di decesso). Gli altri casi identificati in vicinanza della data di morte sono stati inclusi, compresa quindi quella parte di casistica su base autoptica per la quale è stato riconosciuto un precedente momento di diagnosi clinica. La sopravvivenza è stata calcolata per tutti i soggetti con età compresa tra 15 e 99 anni: nella monografia vengono incluse analisi separate riguardanti la sopravvivenza osservata nei bambini (0-14 anni) e negli adolescenti (15-19 anni).

L'inclusione nell'analisi dei tumori multipli comporta che un singolo paziente potrà essere conteggiato più volte sia in differenti analisi, in ragione della presenza di differenti tumori primitivi per lo stesso soggetto, sia all'interno della stessa analisi, quando questa raggruppi più sedi oncologiche. La scelta qui effettuata è coerente con i criteri di inclusione dello studio. Va tuttavia sottolineato che questa impostazione costituisce una variazione rispetto al concetto di sopravvivenza basata sul paziente: comporta infatti l'inclusione di un solo evento diagnostico, generalmente il primo, per il medesimo soggetto.⁴

Casi particolari

In considerazione dell'eterogeneità nella classificazione e nella registrazione del comportamento dei tumori della vescica e delle note problematiche che ne derivano,⁵ le analisi presentate per questa sede includono anche i papillomi e carcinomi a cellule transizionali a comportamento non maligno della vescica (morfologie ICD-O-3; 8000-8010 e 8120-8131). Eventuali neoplasie non maligne papillari della vescica codificate non correttamente nell'ambito dei papillomi a cellule squamose (8050-8052) sono state ricondotte al gruppo più idoneo (8120-8131) e pertanto incluse nell'analisi.

I tumori a comportamento benigno e incerto del sistema nervoso centrale e/o in sede intracranica, registrati in modo eterogeneo nel tempo e nello spazio dai registri tumori italiani, sono stati esclusi dall'analisi.

Metodi di analisi

Analisi di coorte

Sopravvivenza relativa

Per i casi inclusi nell'analisi e relativi alla coorte 2000-2004 è stata calcolata la sopravvivenza relativa fino a 5 anni dalla diagnosi come rapporto tra la sopravvivenza osservata e la sopravvivenza attesa.^{6,7}

La sopravvivenza attesa, basata sulle tavole di mortalità ISTAT suddivise per registro, età e sesso, è stimata utilizzando il metodo Ederer II,⁸ che risulta preferibile al metodo di Hakulinen⁹ come stimatore della sopravvivenza netta,¹⁰⁻¹² cioè della probabilità di sopravvivere se il cancro fosse l'unica causa di morte.



each single patient may be counted several times and in different analyses, because of the presence of different primaries for the same patient, and also within the same analysis, when grouping different cancer sites. The choice made in the latter case is consistent with the inclusion criteria. However, it should be noted that this approach represents a variation on the idea of patient-based survival, which includes only a single primary, usually the first, for the same subject.⁴

Additional criteria

There are several issues related to the current use of different practices for coding behaviour of urinary bladder cancer,⁵ therefore the present publication includes non-malignant bladder transitional cell papillomas and carcinomas (ICD-O-3 morphologies: 8000-8010 plus 8120-8131). Papillary non-malignant bladder neoplasms miscoded into spindle cell papillomas (8050-8052) were revised and coded according to the most appropriate category (8120-8131) and included in the analysis.

Since intra-cranial and/or central nervous system cancers with benign or uncertain behaviour are registered in different ways, both temporally and spatially, they were excluded.

Methods

Cohort analysis

Relative survival

Five-year relative survival was calculated for all cases included in the cohort analysis (years of diagnosis 2000-2004) as the ratio of observed to expected survival.^{6,7}

National life tables by registry, age, and gender were used for calculating expected survival according to the Ederer II method,⁸ which is preferable to the Hakulinen method⁹ when estimating net survival,¹⁰⁻¹² i.e., the probability of survival if cancer is the only possible cause of death.

Cohorts consisting of less than 50 patients were not presented.

Non sono presentate stime di sopravvivenza per coorti con numerosità iniziale inferiore a 50 soggetti.

L'errore standard della sopravvivenza relativa cumulata è calcolato come rapporto tra l'errore standard della sopravvivenza osservata (OS) e la sopravvivenza attesa (ES), che si assume essere costante. L'errore standard della sopravvivenza osservata cumulata dopo un tempo n dalla diagnosi è ottenuto con la formula di Greenwood:¹³

$$SE(OS_n) = OS_n \sqrt{\sum_{i=1}^n \frac{d_i}{\left(l_i - \frac{W_i}{2}\right)\left(l_i - \frac{W_i}{2} - d_i\right)}}$$

Dove l_i esprime il numero dei soggetti viventi all'inizio dell' i -esimo intervallo di tempo, w_i il numero di soggetti non deceduti con tempo di osservazione terminato durante l'intervallo i -esimo, mentre d_i rappresenta il numero di soggetti deceduti durante l'intervallo di tempo i -esimo.

Per il calcolo degli intervalli di confidenza nella presente pubblicazione è stata applicata la trasformazione log-log complementare della sopravvivenza osservata, ottenendo così valori asimmetrici sempre all'interno del *range* dei valori plausibili. Il limite inferiore e superiore della sopravvivenza relativa cumulata dopo un tempo generico n , indicati rispettivamente con LRS_n e USR_n , sono pertanto calcolati nel seguente modo:

$$LRS_n = \frac{e^{-e(LOS'_n)}}{ES_n} \quad USR_n = \frac{e^{-e(UOS'_n)}}{ES_n}$$

Dove LOS'_n e UOS'_n sono rispettivamente i limiti inferiore e superiore della trasformazione log-log complementare della sopravvivenza osservata cumulativa dopo un tempo n , così calcolabili:

$$LOS'_n = \ln(-\ln(OS_n)) - 1.96 \frac{SE(OS_n)}{OS_n \ln(OS_n)}$$

$$UOS'_n = \ln(-\ln(OS_n)) + 1.96 \frac{SE(OS_n)}{OS_n \ln(OS_n)}$$

Sopravvivenza per area geografica

Per il calcolo della sopravvivenza per area geografica è stata applicata una standardizzazione per età, utilizzando l'approccio proposto per il protocollo EURO CARE,¹⁴ con pesi standard per neoplasia e fascia di età. Non sono presentate stime di sopravvivenza nel caso in cui una singola classe di

The standard error of cumulative relative survival is calculated as the ratio of the standard error of observed survival (OS) to expected survival (ES), that is assumed to remain constant. The standard error of cumulative observed survival after an interval n from diagnosis can be derived from the Greenwood formula:¹³

$$SE(OS_n) = OS_n \sqrt{\sum_{i=1}^n \frac{d_i}{\left(l_i - \frac{W_i}{2}\right)\left(l_i - \frac{W_i}{2} - d_i\right)}}$$

where l_i is the number of patients alive at the beginning of the i th interval, w_i is the number of patients alive, with survival experience truncated within the i th interval, while d_i is the number of deaths within the same interval.

In order to obtain non-symmetric values and limit survival within a range of plausible values, a complementary log-log transformation was used to calculate the confidence intervals. The lower and upper limits of cumulative relative survival after the n th interval, denoted LRS_n and USR_n , are calculated as follows:

$$LRS_n = \frac{e^{-e(LOS'_n)}}{ES_n} \quad USR_n = \frac{e^{-e(UOS'_n)}}{ES_n}$$

where LOS'_n and UOS'_n are the lower and upper confidence limits on the log hazard scale:

$$LOS'_n = \ln(-\ln(OS_n)) - 1.96 \frac{SE(OS_n)}{OS_n \ln(OS_n)}$$

$$UOS'_n = \ln(-\ln(OS_n)) + 1.96 \frac{SE(OS_n)}{OS_n \ln(OS_n)}$$

Survival by geographic area

Survival by geographic area is calculated by age-adjusting survival according to standard EURO CARE site- and age-specific weights.¹⁴ Survival was not computed if any age group had no patients. Minor adjustments were introduced, however, to allow calculation for mesothelioma (Table 3).

The standard error for age-standardized survival can be calculated with the formula:

Standard	Age groups	Weights	Cancer sites
1	15-44, 45-54, 55-64, 65-74, 75+	7, 12, 23, 29, 29	all (except those included in standard 1*, 2, 3, 4)
1*	15-54, 55-64, 65-74, 75+	19, 23, 29, 29	mesothelioma
2	15-44, 45-54, 55-64, 65-74, 75+	28, 17, 21, 20, 14	melanoma, cervix uteri, brain and CNS, thyroid
3	15-44, 45-54, 55-64, 65-74, 75+	60, 10, 10, 10, 10	testis, Hodgkin lymphoma, acute lymphocytic leukaemia
4	15-54, 55-64, 65-74, 75-84, 85+	19, 23, 29, 23, 6	prostate

Table 3. Standard population used to calculate standardized relative survival, modified standards are denoted by *.

età sia a numerosità nulla. Sono tuttavia stati introdotti minimi correttivi per consentire il calcolo di alcune sedi specifiche: per il mesotelioma, in particolare, è stato definito un nuovo gruppo di età 15-54 anni, in sostituzione ai gruppi 15-44 e 45-54, numericamente più esigui, mantenendone intatto il peso complessivo (tabella 3). L'errore standard della sopravvivenza standardizzata per età è calcolato con la formula:

$$SE(SRS_n) = \sqrt{\sum_{k=1}^{\alpha} [w_k SE(RS_{n,k})]^2}$$

Sopravvivenza per fascia di età

Per il calcolo della sopravvivenza relativa cumulata per fascia di età, sono state utilizzate le fasce di età definite dalla standardizzazione per età. Non sono presentate stime di sopravvivenza per coorti con numerosità inferiori a 15 soggetti.

Analisi del trend temporale

Modelli di sopravvivenza

La sopravvivenza relativa standardizzata a 5, 10 e 15 anni viene calcolata per i trienni 1990-1992, 1993-1995, 1996-1998, 1999-2001, 2002-2004 e 2005-2007. Per i periodi più recenti, per i quali non è sempre possibile ottenere il calcolo diretto della sopravvivenza utilizzando l'analisi di coorte, si utilizza una stima ibrida come combinazione dell'analisi di coorte e di periodo.¹⁵

Sopravvivenza ibrida

Un esempio di applicazione dell'approccio ibrido è fornito in figura 3 per la sopravvivenza a 10 anni della coorte di soggetti con diagnosi nel periodo 1999-2001 e follow-up al 2008. Indicando per ogni cella gli anni di sopravvivenza completati, si desume che, per calcolare la stima di sopravvivenza a 10 anni dalla diagnosi, sia necessario combinare la sopravvivenza della coorte troncata all'anno di follow-up 2008 con l'analisi di periodo delle esperienze di follow-up recente per gli anni di diagnosi 1996-1998.

$$SE(SRS_n) = \sqrt{\sum_{k=1}^{\alpha} [w_k SE(RS_{n,k})]^2}$$

Age-specific survival

EUROCARE standard age-groups were used for calculating age-specific survival. Survival analysis was not calculated if the number of initial patients was less than 15.

Time trend analysis

Survival models

Five-, ten-, and fifteen-year age-adjusted relative survival was calculated for the following years: 1990-1992, 1993-1995, 1996-1998, 1999-2001, 2002-2004, and 2005-2007. Since traditional cohort-based survival cannot be directly calculated for the most recent periods, a hybrid estimate was used.¹⁵

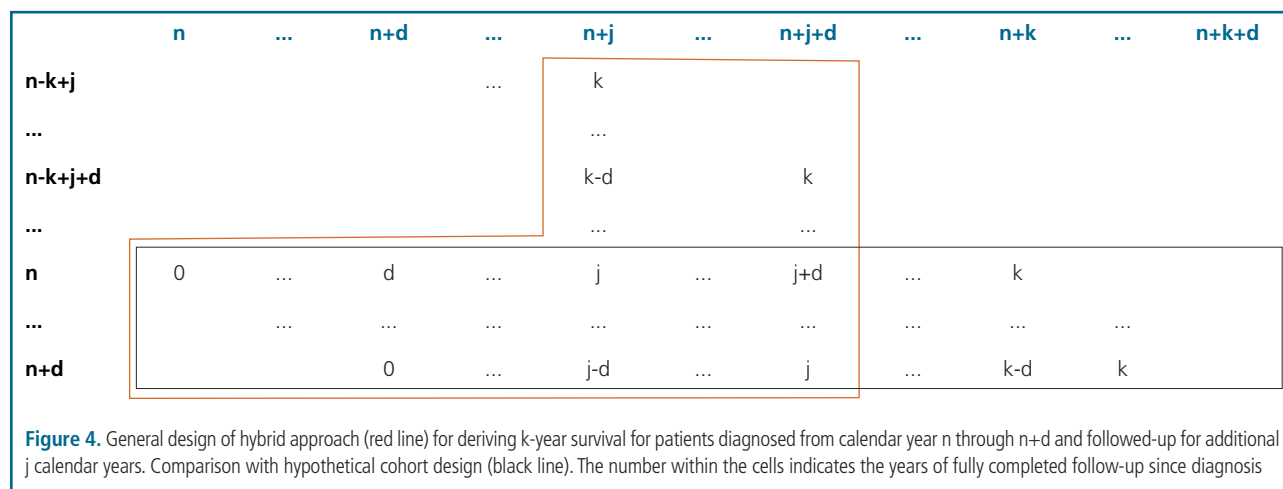
Hybrid survival

Figure 3 gives an example of hybrid approach calculating 10-year survival in a cohort of patients diagnosed in 1999-2001 and followed up until the end of 2008. Denoting each cell with the maximum potential years of observation after the time of diagnosis, it follows that, in order to estimate 10-year survival, cohort survival experiences through the end of 2008 must be combined with the period analysis of most recent survival experiences of cohorts diagnosed in 1996-1998.

In its original form, hybrid analysis is used to estimate the matching period survival, such as 2006-2008 period survival in the previous example. In this publication we used a different approach,¹⁶ in which hybrid analysis is interpreted as a 1990-2001 cohort survival projection. This approach has already been used recently¹⁷ since it produces a more reliable survival estimate. In general, considering a cohort of patients diagnosed between years n and $n+d$ followed-up for at least j additional calendar years, i.e., until calendar year $n+j+d$, their k -year survival, with $k > j$, can be computed combining cohort survival experiences till the end of $n+j+d$ with period survival for years of diagnosis from $n-k+j$ to $n-1$ and years of follow-up from $n+j$ to $n+j+d$ (Figure 4).

		Follow-up years												
		1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Years of diagnosis	1996	3	4	5	6	7	8	9	10					
	1997	2	3	4	5	6	7	8	9	10				
	1998	1	2	3	4	5	6	7	8	9	10			
	1999	0	1	2	3	4	5	6	7	8	9	10		
	2000		0	1	2	3	4	5	6	7	8	9	10	
	2001			0	1	2	3	4	5	6	7	8	9	10

Figure 3. Application of hybrid approach (red line) for deriving 10-year survival for patients diagnosed from 1999 through 2001 and followed-up the end of 2008. Comparison with hypothetical cohort design (black line). The number within the cells indicates the years of fully completed follow-up since diagnosis.



Nella sua formulazione originale, l'analisi ibrida viene interpretata come una stima della corrispondente sopravvivenza di periodo e quindi relativa agli ipotetici anni di diagnosi 2006-2008 (si veda l'esempio precedente). Nella presente pubblicazione viene invece utilizzato un approccio modificato¹⁶ nel quale l'analisi ibrida viene applicata come proiezione della sopravvivenza di coorte più recente, ovvero quella relativa agli anni di diagnosi 1999-2001. Questo genere di approccio è già stato usato recentemente,¹⁷ in quanto produce un'interpretazione della stima di sopravvivenza più attendibile.

In generale, data la coorte di pazienti compresa tra gli anni di diagnosi n e n+d e con follow-up completato per almeno j anni successivi, ovvero fino all'anno n+j+d, la sopravvivenza può essere calcolata a k anni dalla diagnosi, dove ovviamente k>j; è pertanto necessario abbinare alla sopravvivenza di coorte troncata alla fine dell'anno di follow-up n+j+d, la sopravvivenza di periodo relativa agli anni di diagnosi compresi tra n-k+j e n-1 in corrispondenza degli anni di follow-up compresi tra n+j e n+j+d (figura 4).

Come detto, l'uso dell'analisi ibrida comporta un parziale utilizzo dell'approccio di periodo. Questo implica che le esperienze di follow-up considerate possano essere troncate non solo nella loro parte terminale, come accade nella sopravvivenza di coorte, ma anche nella loro parte iniziale. In questo caso la probabilità di sopravvivenza può essere calcolata dalla relazione tra questa e il rischio cumulativo:

$$OS_n = \prod_{i=1}^n \exp\left(-\frac{k_i d_i}{y_i}\right)$$

In questa equazione il rischio cumulativo nell'intervallo i-esimo è ottenuto moltiplicando l'ampiezza dell'intervallo k_i per il tasso medio, stimato dal rapporto d_i/y_i, rispettivamente il numero di morti e il tempo-persona a rischio nell'intervallo. In questo caso l'errore standard della sopravvivenza osservata è dato da:¹⁸

$$SE(OS_n) = OS_n \sqrt{\sum_{i=1}^n \frac{k_i^2 d_i}{y_i^2}}$$

As mentioned before, hybrid analyses involve the use of the period approach. This implies that follow-up experiences can be not only right-truncated at their end, as in cohort survival, but also left-truncated at their beginning. In this case, survival probability can be estimated using the relationship between survivor function and cumulative hazard:

$$OS_n = \prod_{i=1}^n \exp\left(-\frac{k_i d_i}{y_i}\right)$$

In the preceding formula, cumulative hazard for the ith interval is obtained multiplying the length of interval k_i by the average rate of the number of deaths and person-years at risk d_i/y_i. The standard error of observed survival is calculated as follows:¹⁸

$$SE(OS_n) = OS_n \sqrt{\sum_{i=1}^n \frac{k_i^2 d_i}{y_i^2}}$$

where the expression in the square root is the standard error estimate for the cumulative hazard.

Conditional survival

Considering a cohort of patients, their conditional survival CS_{x,y} is the probability of surviving an additional y years, given that patients already survived x years. A simple formula for conditional survival is:¹⁹

$$CS_{x,y} = \frac{RS_{x+y}}{RS_x}$$

Conditional survival has been adjusted for age, with the above mentioned EURO CARE standard weights. In order to calculate standard error for the relative survival rate, variance was derived as follows:

$$Var(RS_{x+y}) = RS_x^2 Var(CS_{x,y}) + CS_{x,y}^2 Var(RS_x)$$

dove la quantità sotto radice è la stima dell'errore standard del rischio cumulativo.

Sopravvivenza condizionata

La sopravvivenza condizionata $CS_{x,y}$ rappresenta la probabilità di una coorte di pazienti, che è già sopravvissuta x anni dopo la diagnosi, di sopravvivere ulteriori y anni. Si può calcolare con la semplice formula:¹⁹

$$CS_{x,y} = \frac{RS_{x+y}}{RS_x}$$

La sopravvivenza condizionata viene standardizzata per età, secondo le stesse modalità descritte precedentemente, per consentire un più agevole confronto per macroarea geografica. Per il calcolo dell'errore standard della sopravvivenza condizionata standardizzata, si ricava la corrispondente varianza:

$$Var(RS_{x+y}) = RS_x^2 Var(CS_{x,y}) + CS_{x,y}^2 Var(RS_x)$$

ottenendo:

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

Aggiustamento per la distribuzione delle sedi

La sopravvivenza complessiva, qui presentata in due forme, come aggregazione di tutte le sedi a esclusione dei carcinomi cutanei e di tutte le sedi a esclusione dei tumori della vescica e dei carcinomi cutanei, è solitamente poco considerata dallo specialista poiché influenzata dalla distribuzione eterogenea delle sedi neoplastiche al suo interno. Nelle tabelle 4 e 5 si illustrano, a titolo di esempio, le distribuzioni della casistica per l'analisi di coorte 2000-2004 relativa ai 29 registri tumori generali e suddivisa per sede e macroarea. Per limitare il problema della distribuzione eterogenea dei casi, sia nello spazio sia nel tempo, è stato definito un aggiustamento per sede neoplastica²⁰ sulla base della casistica 2000-2004, illustrato in tabella 6.

Analisi delle assunzioni

Il valore della stima della sopravvivenza è in parte condizionato anche dalle scelte metodologiche, quali per esempio i criteri di inclusione ed esclusione dallo studio. Molte di queste assunzioni sono adottate convenzionalmente sulla base delle esperienze passate, altre sono invece introdotte a seguito di fondate evidenze.²¹ In questo paragrafo si vogliono analizzare i principali aspetti di potenziale criticità e il loro effetto sulla stima finale della sopravvivenza, in particolare:

- **Tumori multipli.** L'inclusione dei tumori multipli comporta che il calcolo della sopravvivenza non includa solo la prognosi del primo tumore, ma l'impatto di questa modifica della pratica corrente per l'analisi della sopravvivenza è stata valutata tramite il confronto con la sopravvivenza tradizionale limitata al solo primo tumore.
- **Autopsie.** Nel considerare la casistica su base autoptica, è

Then:

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

Adjusting for case-mix

Overall survival, here reported in two different ways, as all sites expect skin carcinomas and all sites except skin carcinomas and bladder cancer, is usually not considered by specialists because it is strongly influenced by case-mix.

Tables 4 and 5 show the distribution of cases from 29 general cancer registries for patients diagnosed in 2000-2004 by site and area. In order to minimize case-mix bias both spatially and temporally, a cancer site adjustment was defined²⁰ using the 2000-2004 cohort dataset as reference (Table 6).

Analysis of assumptions

When estimating survival, decisions concerning methods, including inclusion/exclusion criteria, may in part influence survival estimates. Many of the assumptions are usually based on past knowledge or convincing evidence.²¹ In this section, some of the main critical aspects will be examined, along with their effect on relative survival estimates. They are as follows:

- **Multiple tumours.** The inclusion of multiple tumours implies that survival is not computed only based on the first primary. The influence of this decision was studied comparing results with traditional survival calculated considering only the first primary.
- **Autopsies.** We compared our selection criterion (i.e., to exclude autopsy-only cases) with an alternative criterion, which includes all cancers detected at autopsy.
- **DCO.** Cases known by death certificate only are excluded from analysis. In order to check the maximum error of this assumption we included DCO cases with two variants: assuming zero survival and assuming a survival for at least the survival period in consideration (in this example, 5 years).
- **Expected survival.** The Ederer II method differs from the Hakulinen method, used in most survival studies. Thus, the survival estimate obtained with the Hakulinen method was reported, as well.

We also reported the effect of combining both Hakulinen's expected survival and the exclusion of multiple primaries from calculation.

The effects of these different methods are shown for the cohort dataset from 2000 through 2004 with follow-up for death till the end of 2008 for some characteristic cancers (lung, colon and rectum, prostate, thyroid and female breast).

As expected, the inclusion of subsequent tumours, in general, causes survival estimates to decrease by about one percent, with the exception of lung cancer, the estimate of which slightly increases when considering all lung primaries.

The exclusion of cancer detected at autopsy has a small effect of increase on survival. This effect is more relevant in the elderly and, in particular, for prostate and thyroid cancers, where differences rise to 0.8%.

♂ Males	North-East		North-West		Centre		South		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%
Prostate	26 218	20.6	18 989	20.9	7 631	17.7	5 262	14.4	58 100	19.5
Lung	19 491	15.3	14 955	16.4	6 744	15.7	6 309	17.3	47 499	15.9
Bladder	11 837	9.3	9 274	10.2	4 858	11.3	4 760	13.0	30 729	10.3
Colon	11 509	9.0	7 957	8.7	3 996	9.3	2 756	7.5	26 218	8.8
Stomach	6 947	5.4	4 079	4.5	3 000	7.0	1 476	4.0	15 502	5.2
Rectum	5 171	4.1	3 587	3.9	2 290	5.3	1 596	4.4	12 644	4.2
Liver	5 081	4.0	3 735	4.1	1 236	2.9	2 101	5.7	12 153	4.1
Non Hodgkin lymphoma	4 315	3.4	3 089	3.4	1 304	3.0	1 210	3.3	9 918	3.3
Kidney	4 029	3.2	2 862	3.1	1 529	3.6	851	2.3	9 271	3.1
Pancreas	3 445	2.7	2 286	2.5	1 036	2.4	862	2.4	7 629	2.6
Larynx	2 875	2.3	2 169	2.4	943	2.2	971	2.7	6 958	2.3
Skin melanoma	2 719	2.1	2 031	2.2	939	2.2	553	1.5	6 242	2.1
Head and neck except larynx	2 734	2.1	1 635	1.8	562	1.3	564	1.5	5 495	1.8
Brain and other CNS	1 739	1.4	1 189	1.3	688	1.6	697	1.9	4 313	1.4
Myeloma	1 590	1.2	1 074	1.2	539	1.3	491	1.3	3 694	1.2
Oesophagus	1 701	1.3	928	1.0	277	0.6	264	0.7	3 170	1.1
Biliary tract	1 128	0.9	918	1.0	424	1.0	480	1.3	2 950	1.0
Chronic Lymphocytic Leukaemia	1 111	0.9	857	0.9	299	0.7	405	1.1	2 672	0.9
Testis	1 077	0.8	769	0.8	345	0.8	427	1.2	2 618	0.9
Thyroid	1 116	0.9	664	0.7	398	0.9	408	1.1	2 586	0.9
Acute Myeloid Leukaemia	847	0.7	614	0.7	289	0.7	337	0.9	2 087	0.7
Urinary tract	891	0.7	573	0.6	227	0.5	171	0.5	1 862	0.6
Hodgkin lymphoma	614	0.5	485	0.5	243	0.6	282	0.8	1 624	0.5
Soft tissue	526	0.4	485	0.5	201	0.5	206	0.6	1 418	0.5
Mesothelioma	678	0.5	382	0.4	108	0.3	150	0.4	1 318	0.4
Chronic Myeloid Leukaemia	416	0.3	318	0.3	137	0.3	173	0.5	1 044	0.4
Small intestine	401	0.3	256	0.3	98	0.2	98	0.3	853	0.3
Kaposi sarcoma	247	0.2	367	0.4	63	0.1	122	0.3	799	0.3
Salivary glands	273	0.2	189	0.2	107	0.2	110	0.3	679	0.2
Bone	169	0.1	128	0.1	66	0.2	97	0.3	460	0.2
Acute Lymphocytic Leukaemia	154	0.1	89	0.1	45	0.1	87	0.2	375	0.1
Choroid melanoma	46	0.0	26	0.0	26	0.1	19	0.1	117	0.0
Other	6 472	5.1	4 048	4.4	2 361	5.5	2 264	6.2	15 145	5.1
Total	127 567	100.0	91 007	100.0	43 009	100.0	36 559	100.0	298 142	100.0

Table 4. Number of cases used for cohort analysis (years 2000-2004) by site and macroarea. Males.

possibile utilizzare il criterio applicato nella presente pubblicazione (ovvero escludere i casi diagnosticati incidentalmente con autopsia) oppure includere tutti i casi con base diagnosi autoptica.

■ *DCO.* I casi notificati dal solo certificato di decesso sono esclusi dall'analisi; è possibile tuttavia stimare il peso di questa assunzione includendo i casi DCO e attribuendo loro un tempo di sopravvivenza pari a 0 o in alternativa pari ad almeno l'intera durata del periodo (in questo caso, 5 anni). Saranno valutate entrambe le ipotesi.

■ *Sopravvivenza attesa.* Il metodo applicato, cosiddetto Ederer II, differisce da quello applicato nelle maggiori monografie su questo tema, basate invece sulla stima di Hakulinen. Si riporta pertanto la sopravvivenza ottenuta con l'applicazione di questo metodo.

Viene poi valutato l'effetto dell'utilizzo combinato della metodologia di stima di Hakulinen e dell'esclusione dei tumori multipli.

Le diverse metodologie verranno analizzate sulla coorte 2000-2004 con follow-up al 31.12.2008 per alcune neoplasie rap-

Cases known by death certificate only are commonly excluded from survival analyses. The theoretical maximum overestimate from excluding DCO cases can be computed by attributing zero survival to all DCO cases. Even though the average bias appears quite small, at older ages the differences become relevant, in particular for prostate cancer, where differences reach 4.4% compared to the index scenario.

The theoretical maximum underestimate from excluding DCO cases is estimated by assuming that all DCO cases are alive after 5 years. Again the differences are concentrated in the older age groups where, in prostate cancer, the difference from the index scenario reaches 23.1%.

As mentioned above, expected survival estimates using the Hakulinen method may cause an overcorrection of observed survival and results in a potential overestimate of net survival. From a comparison between the Hakulinen and the method known as Ederer II, we observe that the latter lowers relative survival estimates by about one percent.

Combining inclusion of multiple primaries with application of the Ederer II method seems to add up the already observed

♀ Females	North-East		North-West		Centre		South		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%
Breast	31 075	29.7	24 052	31.1	10 044	28.2	8 413	28.6	73 584	29.8
Colon	9 783	9.4	7 524	9.7	3 517	9.9	2 422	8.2	23 246	9.4
Lung	5 935	5.7	4 717	6.1	1 852	5.2	1 121	3.8	13 625	5.5
Corpus uteri	4 693	4.5	3 570	4.6	1 707	4.8	1 541	5.2	11 511	4.7
Stomach	5 016	4.8	3 045	3.9	2 119	6.0	937	3.2	11 117	4.5
Rectum	3 807	3.6	3 112	4.0	1 642	4.6	1 179	4.0	9 740	3.9
Non Hodgkin lymphoma	3 804	3.6	2 895	3.7	1 197	3.4	1 147	3.9	9 043	3.7
Thyroid	3 482	3.3	2 030	2.6	1 267	3.6	1 508	5.1	8 287	3.4
Ovary	3 264	3.1	2 660	3.4	1 244	3.5	1 106	3.8	8 274	3.4
Pancreas	3 741	3.6	2 564	3.3	999	2.8	837	2.8	8 141	3.3
Bladder	3 118	3.0	2 458	3.2	1 068	3.0	873	3.0	7 517	3.0
Skin melanoma	2 791	2.7	2 058	2.7	962	2.7	549	1.9	6 360	2.6
Liver	2 219	2.1	1 687	2.2	665	1.9	1 002	3.4	5 573	2.3
Kidney	2 383	2.3	1 541	2.0	772	2.2	486	1.7	5 182	2.1
Biliary tract	1 704	1.6	1 339	1.7	542	1.5	739	2.5	4 324	1.8
Cervix uteri	1 660	1.6	1 357	1.8	650	1.8	601	2.0	4 268	1.7
Brain and CNS	1 488	1.4	1 062	1.4	618	1.7	594	2.0	3 762	1.5
Myeloma	1 661	1.6	1 032	1.3	491	1.4	478	1.6	3 662	1.5
Head and neck except larynx	912	0.9	639	0.8	230	0.6	190	0.6	1 971	0.8
Chronic Lymphocytic Leukaemia	848	0.8	580	0.7	242	0.7	264	0.9	1 934	0.8
Acute Myeloid Leukaemia	828	0.8	601	0.8	246	0.7	252	0.9	1 927	0.8
Hodgkin lymphoma	561	0.5	412	0.5	199	0.6	213	0.7	1 385	0.6
Soft tissues	477	0.5	386	0.5	201	0.6	140	0.5	1 204	0.5
Oesophagus	501	0.5	323	0.4	112	0.3	69	0.2	1 005	0.4
Urinary tract	438	0.4	274	0.4	96	0.3	50	0.2	858	0.3
Chronic Myeloid Leukaemia	344	0.3	244	0.3	109	0.3	113	0.4	810	0.3
Larynx	328	0.3	216	0.3	94	0.3	71	0.2	709	0.3
Small intestine	338	0.3	200	0.3	80	0.2	89	0.3	707	0.3
Mesothelioma	214	0.2	222	0.3	42	0.1	65	0.2	543	0.2
Salivary glands	164	0.2	148	0.2	81	0.2	73	0.2	466	0.2
Bone	119	0.1	108	0.1	61	0.2	60	0.2	348	0.1
Acute Lymphocytic Leukaemia	138	0.1	91	0.1	39	0.1	64	0.2	332	0.1
Kaposi sarcoma	90	0.1	126	0.2	19	0.1	84	0.3	319	0.1
Choroid melanoma	40	0.0	25	0.0	16	0.0	19	0.1	100	0.0
Other	6 549	6.3	4 056	5.2	2 348	6.6	2 041	6.9	14 994	6.1
Total	104 513	100.0	77 354	100.0	35 571	100.0	29 390	100.0	246 828	100.0

Table 5. Number of cases used for cohort analysis (years 2000-2004) by site and macroarea. Females.

presentative (polmone, colon retto, prostata, tiroide, mammella femminile).

Come atteso, l'inclusione dei tumori successivi al primo produce, come effetto generale, quello di diminuire la stima di sopravvivenza di circa un punto percentuale. Fanno eccezione i tumori del polmone, la cui stima risulta invece lievemente aumentata quando si considerano tutti i tumori polmonari.

L'esclusione delle diagnosi autoptiche ha globalmente un piccolo effetto di aumento della stima di sopravvivenza, più rilevante nei soggetti anziani e, in particolare, per i tumori della prostata e della tiroide per i quali si osserva una differenza della stima pari a circa lo 0,8%.

La casistica dei casi noti dal solo certificato di decesso viene generalmente esclusa dagli studi di sopravvivenza. La massima sovrastima teorica legata a questa esclusione viene espressa attribuendo ai DCO una sopravvivenza pari a 0. Anche se l'effetto medio appare contenuto, questa differenza diviene rilevante per le classi di pazienti in età avanzata, e raggiunge uno scostamento del 4,4% rispetto all'analisi indice per i tumori della prostata in particolare. La massima sottostima teorica dell'esclusione dei

individual differences, resulting in lower estimates, with a difference compared to the index scenario of 2.7% for prostate cancer and 0.4% for lung cancer.

In conclusion, the inclusion of multiple primaries has, in itself, a variable - though generally small - effect on survival, but this practice is desirable for a better net survival estimate,³ to reduce internal inconsistency due to the heterogeneous pattern of the AIRTUM registries, as well as to improve the reliability of international comparisons.^{3,4}

The exclusion of autopsy only cases has a limited effect on survival, an effect which varies with age, site and, above all, geographic area. Certain areas in north-eastern Italy had a higher proportion of cases diagnosed only by autopsy.²² Excluding these cases is therefore justified in order to reduce possible geographic heterogeneities of survival estimates.

Death certificate only cases are associated with poorer and, sometimes, incorrect diagnosis. We therefore deem it appropriate to exclude them from analysis, especially in the case of elderly patients. As shown in Table 7, both over- and underestimation of real survival are possible, but it is hard to quantify

DCO è stata invece mostrata attribuendo ai DCO una sopravvivenza superiore a 5 anni. Anche in questo caso le differenze si concentrano sulle fasce di soggetti in età anziana raggiungendo, per i tumori prostatici, una differenza rispetto all'analisi indice del 23,1%.

Come già osservato in precedenza, la stima della sopravvivenza attesa con la metodologia di Hakulinen porta, in alcuni casi, a una sovracorrezione della sopravvivenza osservata e una conseguente possibile sovrastima della sopravvivenza netta. Nell'analisi comparativa tra il metodo di Hakulinen e il metodo cosiddetto Ederer II, applicato nel presente studio, si osserva una generale tendenza del secondo a produrre stime di sopravvivenza relative inferiori al primo di circa un punto percentuale. L'uso combinato dell'inclusione dei tumori multipli e dell'adozione della metodologia Ederer II rispecchia in modo apparentemente additivo differenze osservate individualmente per le singole assunzioni, producendo stime della sopravvivenza generalmente più basse, con una differenza rispetto all'analisi di riferimento del 2,7% per i tumori della prostata e del 0,4% per i tumori del polmone.

Concludendo, l'inclusione dei tumori multipli ha di per sé un effetto variabile, ma generalmente contenuto, sulla sopravvivenza; tuttavia questa pratica appare, a nostro giudizio, preferibile sia per una migliore stima della sopravvivenza netta della popolazione,³ sia in quanto riduce l'inconsistenza interna legata all'eterogeneità dei registri presenti nel pool AIRTUM, nonché le problematiche relative ai confronti internazionali.^{3,4}

L'esclusione delle diagnosi autoptiche cosiddette incidentali ha un effetto ridotto sulla stima della sopravvivenza anche se variabile per età, sede neoplastica e, soprattutto, area geografica. E' noto infatti che in alcune aree del Nord-Est si è registrata una quota maggiore di casi di tumore diagnosticati esclusivamente attraverso le autopsie.²² L'esclusione di questa casistica è quindi giustificata al fine di ridurre l'eterogeneità della stima potenzialmente ottenibile con diverse disaggregazioni geografiche.

I casi DCO sono associati a una definizione diagnostica povera

	Males		Females	
	All but skin	All but skin and bladder	All but skin	All but skin and bladder
Head and neck	4.1	4.7	-	-
Stomach	5.2	5.8	4.5	4.6
Colon rectum	13.0	14.5	13.3	13.8
Liver	4.1	4.5		2.3
Pancreas	2.6	2.9	3.3	3.4
Lung	15.9	17.8	5.5	5.7
Skin melanoma	2.1	2.3	2.6	2.7
Breast			29.8	30.7
Corpus uteri			4.7	4.8
Ovary			3.4	3.5
Prostate	19.5	21.7		
Kidney	3.1	3.5		
Bladder	10.3	-	3.0	-
Thyroid			3.4	3.5
NHL/CLL	4.2	4.7	4.5	4.6
Myeloma	-	1.4		
Other	15.9	16.2	22.0	20.4
Total	100.0	100.0	100.0	100.0

Table 6. Weights used in case-mix adjustment for all cancers but non melanoma skin cancer and all cancers but non melanoma skin and bladder cancer. 2000-2004.

these effects. On the other hand, the small number of DCO cases in Italian registries limits the extent of the problem (albeit without completely removing it).

As previously mentioned, there are several reasons to opt for the Ederer II method. However, a comparative study was called for, since the Hakulinen method had been widely used in the past. Our comparison confirmed that the Hakulinen method can sometimes produce survival overestimation even for 5-year survival. Furthermore, age-standardization, used in this report, has been observed to reduce the differences between relative survival estimates based on these methods and net survival.¹⁰

age	Five years survival %														
	Lung			Colon rectum			Prostate			Thyroid			Female Breast		
	15-99	15-74	75-99	15-99	15-74	75-99	15-99	15-74	75-99	15-99	15-74	75-99	15-99	15-74	75-99
Index scenario	13.0	15.7	7.4	57.6	62.6	49.3	88.5	90.3	51.2	93.7	96.4	58.5	86.5	89.0	76.4
Only first primary	12.9	15.4	7.3	58.5	63.4	50.1	90.0	91.8	52.0	94.5	96.9	58.5	87.0	89.3	77.2
Include autopsy only cases	13.0	15.7	7.4	57.4	62.6	49.0	88.2	90.1	50.4	93.4	96.2	57.7	86.6	89.0	76.4
Include DCO with 0 survival	12.8	15.6	7.1	57.1	62.5	48.6	87.8	90.0	46.8	93.5	96.4	57.7	86.0	88.9	74.6
Include DCO with >5 years survival	15.5	16.5	14.0	58.4	62.8	51.5	89.4	90.5	74.3	93.7	96.4	60.0	87.0	89.1	79.4
H.method	13.6	15.8	7.8	58.9	62.8	50.6	89.7	90.7	52.8	94.5	96.5	61.2	87.0	89.1	77.2
Only first primary / H. method	13.4	15.5	7.8	59.8	63.6	51.4	91.2	92.2	53.8	95.2	97.1	61.2	87.5	89.4	78.0

Table 7. Five-year survival (%) in the AIRTUM databases from 2000 through 2004 according to different methods of estimation. H. method = Hakulinen method

e, a volte, mal posta. L'esclusione degli stessi dalle analisi di sopravvivenza è quindi opportuna, a giudizio degli autori, specie nei soggetti anziani. Come mostrato in tabella 7, sono possibili effetti di sovrastima e sottostima della sopravvivenza reale e la portata di questi stessi effetti, spesso combinati, è difficilmente quantificabile con precisione. D'altra parte si rileva che il numero relativamente ridotto di casi DCO riscontrato tra i registri italiani limita (ma non elimina) l'entità del fenomeno. L'adozione del metodo Ederer II, come già detto, è basata su molti e fondati motivi. Tuttavia il radicato utilizzo della metodologia di Hakulinen ne ha consigliato un'analisi comparativa che, di fatto, conferma un possibile effetto di sovrastima di quest'ultima metodologia anche sulla sopravvivenza a 5 anni. È stato osservato che la standardizzazione per età, presente in questo volume, può ridurre la differenza tra sopravvivenza relativa stimata con questi metodi e sopravvivenza netta.¹⁰

Elaborazione dei dati

Tutte le stime di sopravvivenza sono state calcolate usando il comando *stvs* di Stata, versione 11.2 (http://www.pauldickman.com/rsmodel/stata_colon/).

Qualità

Introduzione

La maggior parte dei registri tumori inclusi nella presente monografia ha già pubblicato i propri dati di incidenza in ambito internazionale, mostrando complessivamente un buon livello qualitativo.²³ Dal 2007, l'AIRTUM ha formalizzato l'insieme delle regole relative ai criteri di registrazione che ogni registro tumori italiano deve adottare.²⁴ Esse riprendono e integrano regole e standard nazionali e internazionali oramai consolidati.^{25,26} È comunque stata pianificata una serie di controlli per consentire una verifica ancora più sistematica dei criteri di registrazione e follow-up adottati dai diversi registri tumori. I dati di incidenza destinati alla banca dati sono sottoposti al protocollo operativo che impone due ordini di controllo. Il primo, precedente l'invio dei dati, è a opera del singolo registro che è tenuto a effettuare i controlli previsti dai programmi DEPedit e IARCcrgTools prodotti da IARC,²⁷ verificando l'assenza di errori e certificando che altri avvisi segnalati siano stati verificati e confermati; il secondo avviene a livello della banca dati che provvede a eseguire nuovamente i controlli previsti da DEPedit. Vengono inoltre condotte le verifiche previste da uno specifico software prodotto dall'AIRTUM, CheckAIRTUM, che controlla eventuali difformità dei dati aggregati rispetto a un dato atteso sulla base di un *gold standard* rappresentato dalla casistica della banca dati, eventualmente corretto per area geografica. Vengono effettuati complessivamente 1.431 controlli che vanno dal rapporto mortalità/incidenza alla proporzione delle sedi non definite. Il software è gratuitamente disponibile sul sito AIRTUM (<http://www.registri-tumori.it/cms/?q=software>). Vengono infine effettuati controlli di congruenza e completezza delle variabili

Data analysis

*All survival estimates were computed using the *stvs* command of the Stata package, version 11.2 (http://www.pauldickman.com/rsmodel/stata_colon/).*

Quality

Introduction

Most cancer registries contributing to this monograph have already published their incidence data in international reports, showing an overall good quality of data.²³ Since 2007, AIRTUM has formalized a set of rules concerning registration criteria that each Italian cancer registry should adopt,²⁴ these rules incorporate well-established national and international guidelines and standards.^{25,26} Nevertheless, a series of quality checks were planned to ensure systematic verification of cancer registration and follow-up procedures adopted by each cancer registry.

There are two levels of control in AIRTUM's centralized database. The first, which each registry must follow before sending data, consists in performing data checks according to the DepEdit and IARCcrgTools software produced by the IARC²⁷ and verifying the absence of errors and confirming or correcting warnings. The second level provides further DepEdits controls. In this phase, a specific software provided by AIRTUM, CheckAIRTUM, is used to compare single registry data with a gold standard. CheckAIRTUM performs 1,431 checks, ranging from mortality/incidence ratio to the proportion of undefined cancer sites. The software is available at the AIRTUM website (<http://www.registri-tumori.it/cms/?q=software>). Other specific checks are also performed, such as a number of controls on data consistency and completeness according to the AIRTUM accreditation procedure.

*From the point of view of cancer registration methods, the most relevant problem concerns urinary bladder tumours, especially the coding of *in situ* and uncertain behaviour tumours. The follow-up criteria is essentially homogeneous among Italian cancer registries and is conducted through the health care databases and death files, with active case finding for all patients lost to follow-up in the municipality archives.²⁸ Unfortunately, certain issues such as migration for health care, quality of health services and territorial extension, may affect the accuracy of data and the completeness of follow-up, leading to differences between cancer registries.*

The overall assessment of both registry and area specific quality indicators justified the decision to include all Italian cancer registries in the AIRTUM database for at least the cohort survival analysis for the years 2000-2004.

Quality indicators

DCO and microscopic verifications

A common indicator for cancer registration accuracy and completeness is the proportion of cases known by death certificate only (DCO). The proportion of DCO cases may indicate some

	%DCO					%MV				
	N-E	N-W	Centre	South	Pool	N-E	N-W	Centre	South	Pool
Lip	0.7	0.0	0.0	0.5	0.4	98.5	96.4	93.6	95.9	97.0
Tongue	0.4	0.1	0.3	0.7	0.3	96.9	96.9	89.9	92.6	95.5
Mouth	0.4	0.4	0.0	1.1	0.4	96.8	97.5	92.3	93.8	96.2
Salivary glands	0.9	1.5	0.0	1.1	0.9	94.1	93.6	81.4	84.9	90.4
Oropharynx	0.4	0.5	1.3	1.4	0.6	96.4	96.5	92.3	91.6	95.6
Nasopharynx	0.7	0.5	0.0	0.8	0.5	93.4	94.2	91.2	92.5	93.1
Hypopharynx	0.2	0.7	2.4	1.1	0.6	94.8	96.8	94.0	96.8	95.6
Pharynx NOS	0.8	4.9	0.0	6.3	1.9	95.2	90.2	73.1	81.3	92.1
Oesophagus	1.1	1.8	2.0	5.6	1.8	88.3	88.8	78.4	78.2	86.7
Stomach	1.3	2.1	1.1	3.1	1.6	90.6	90.4	88.1	86.2	89.7
Small intestine	1.2	1.1	0.0	3.1	1.3	90.0	90.1	87.1	83.9	88.9
Colon	0.5	1.2	0.6	1.7	0.9	93.4	92.7	91.3	88.7	92.3
Rectum	0.3	0.8	0.2	0.8	0.5	94.0	94.5	92.9	93.3	93.9
Liver	1.9	3.1	2.3	5.1	2.9	41.2	42.2	55.5	35.0	41.9
Biliary tract	1.0	2.6	1.3	4.3	2.1	57.6	59.9	64.5	53.4	58.5
Pancreas	2.0	3.4	1.1	5.3	2.7	41.5	50.4	41.3	39.2	44.0
Nasal cavity	0.6	1.2	0.8	0.0	0.8	95.4	92.2	84.4	91.9	92.3
Larynx	0.8	0.8	0.3	2.1	0.9	96.6	95.5	89.9	89.4	94.4
Lung	1.2	2.4	1.2	3.9	1.9	73.0	72.5	72.1	68.4	72.2
Other thoracic	2.6	4.6	1.4	6.5	3.8	69.4	49.6	54.2	67.1	59.8
Bone	1.0	1.3	0.8	4.3	1.7	87.3	87.1	74.2	67.1	81.2
Skin melanoma	0.1	0.1	0.3	0.4	0.2	98.9	98.7	94.7	97.7	98.1
Skin non melanoma	0.0	0.1	0.0	0.1	0.1	99.8	91.5	95.5	98.8	97.1
Mesothelioma	0.1	0.2	0.0	0.0	0.1	82.7	88.4	96.7	98.6	87.9
Kaposi sarcoma	0.0	0.2	0.0	0.5	0.2	97.0	94.0	92.7	95.2	95.0
Soft tissue	0.3	0.7	0.2	1.7	0.6	94.9	94.5	88.4	95.2	93.8
Breast	0.4	0.9	0.3	1.3	0.7	97.5	95.6	95.4	93.8	96.1
Cervix uteri	0.2	0.4	0.0	0.2	0.2	98.4	95.8	92.2	94.4	96.0
Corpus uteri	0.1	0.1	0.1	0.8	0.2	97.3	97.3	93.2	95.6	96.5
Uterus NOS	4.5	31.7	10.1	12.7	10.4	74.6	26.8	38.0	59.0	61.0
Ovary	0.8	1.9	0.7	1.9	1.3	86.1	85.3	83.9	82.4	85.0
Other female organs	0.8	1.3	0.7	2.2	1.0	85.6	91.9	87.8	89.7	88.2
Penis	0.0	0.8	0.9	0.9	0.5	97.3	96.1	92.0	88.2	94.2
Prostate	0.6	1.1	0.7	2.0	0.9	92.0	91.5	88.5	85.4	90.8
Testis	0.2	0.1	0.0	0.2	0.2	97.3	94.7	92.2	91.4	94.9
Other male organs	1.8	0.0	0.0	5.3	1.4	91.2	84.0	84.6	94.7	88.5
Kidney, etc.	0.5	0.9	0.4	1.3	0.7	82.1	83.9	83.3	77.7	82.4
Urinary bladder	0.3	0.4	0.3	0.9	0.4	94.8	94.3	92.3	89.6	93.5
Choroid melanoma	5.5	3.8	2.3	2.6	4.0	49.5	28.3	32.6	48.7	41.2
Brain and CNS	1.0	1.3	0.5	3.7	1.4	58.4	54.1	52.8	55.4	55.6
Thyroid	0.1	0.3	0.0	0.2	0.2	97.0	95.6	91.1	93.8	95.2
Other endocrine	0.6	2.6	0.4	2.6	1.4	68.1	70.5	58.9	47.1	61.4
Hodgkin lymphoma	0.1	0.2	0.0	0.0	0.1	98.5	98.8	92.5	98.6	97.7
Non Hodgkin lymphoma	0.3	0.3	0.4	0.7	0.4	96.3	97.2	91.4	94.0	95.6
Myeloma	0.5	0.8	0.8	2.0	0.8	91.6	90.1	77.6	84.1	88.2
ALL	1.0	0.5	2.3	0.0	0.8	94.3	97.3	73.3	100.0	93.7
CLL	0.6	1.0	0.4	0.0	0.6	95.5	93.0	69.1	98.2	92.0
AML	0.6	0.6	0.7	0.0	0.5	93.0	96.3	81.7	99.5	93.4
CML	1.0	1.4	4.3	1.0	1.6	95.3	93.5	68.2	98.6	91.6
Other CMPD and MDS	0.2	0.5	0.0	0.4	0.2	94.8	86.0	54.2	93.9	82.1
Leukaemia NOS	7.2	3.7	4.1	16.6	7.2	85.0	84.9	57.6	63.1	75.1
Miscellaneous	0.9	10.6	2.4	10.1	6.1	78.0	44.0	63.9	55.3	59.7
Ill defined/metastases	5.0	10.0	4.2	7.8	6.7	55.0	44.3	39.6	42.0	48.5
All but skin, non melanoma	0.8	1.5	0.7	2.3	1.2	85.9	85.3	82.9	81.1	84.7

Table 8. Percentage of DCO and microscopically verified cases (MV) for cases diagnosed from 2000 through 2004 for the North-East (N-E), North-West (N-W), Centre, and South macroareas, and the pool of AIRTUM cancer registries.

Registry	Records	DCO	MV	M/I	Lost	Ill	Dead	Alive
	(No.)	(%)	(%)	(%)	<5 years	defined	<1 year	>5 years
Alto Adige	12 440	1.0	90.4	44.2	0.5	2.4	5.8	12.6
Trento	12 880	1.3	87.4	54.5	0.6	5.4	5.4	9.9
Veneto	66 516	1.1	85.6	46.7	0.1	2.5	5.1	10.0
Friuli Venezia Giulia	42 769	0.3	83.6	48.8	0.4	2.9	6.8	12.0
Parma	14 962	0.5	85.5	49.5	1.3	2.4	5.4	11.6
Reggio Emilia	14 659	0.1	85.9	48.5	0.8	2.2	4.9	10.4
Modena	21 015	0.3	88.5	45.7	0.3	2.4	4.7	9.2
Ferrara	14 049	0.8	85.3	49.9	0.9	2.9	6.0	9.8
Romagna	36 902	1.4	86.2	47.0	0.1	2.3	5.1	10.3
North-East	236 192	0.8	85.9	47.8	0.4	2.7	5.5	10.6
Liguria Mesothelioma	745	0.3	78.3	-	0.0	-	-	-
Genova	35 563	1.6	83.9	49.5	0.0	2.6	5.7	8.6
Milano	46 044	1.9	85.6	48.8	3.9	2.3	5.4	13.0
Brescia	12 039	2.3	85.0	46.4	0.0	0.7	4.6	13.2
Mantova	12 056	1.7	81.8	50.9	0.2	1.9	4.9	11.6
Sondrio	5 723	0.5	86.4	48.7	0.1	2.4	6.3	12.6
Varese	24 933	0.4	86.3	51.4	1.0	2.4	5.4	10.4
Torino	30 295	1.4	87.1	48.2	1.9	2.3	5.7	9.8
Biella	6 705	1.4	85.1	50.5	0.1	2.5	5.9	7.0
North-West	174 103	1.5	85.3	49.3	1.5	2.2	5.5	10.9
Macerata	3 838	1.6	82.5	47.4	0.9	2.4	6.6	10.4
Firenze Prato	39 660	0.7	80.1	49.8	1.7	2.2	5.7	10.8
Umbria	26 061	0.3	90.8	49.2	0.2	1.1	4.9	11.3
Latina	10 708	1.6	74.3	54.1	2.1	2.6	4.3	13.1
Centre	80 267	0.7	82.9	50.0	1.2	1.9	5.3	11.2
Napoli	9 962	1.6	77.5	47.1	1.2	3.2	6.2	12.4
Salerno	21 894	2.6	78.9	52.7	0.9	2.5	6.4	10.1
Catanzaro	2 146	1.1	86.1	46.3	0.0	2.8	4.5	8.6
Trapani	5 442	0.7	81.3	52.9	0.1	2.1	5.6	7.6
Palermo Breast	3 188	1.9	95.8	31.5	0.1	-	-	-
Ragusa	7 021	2.3	79.8	49.2	0.4	2.8	6.1	7.9
Siracusa	8 103	3.0	82.3	52.4	0.0	1.7	5.4	8.6
Sassari	11 201	2.9	83.6	47.7	0.0	2.9	5.7	9.7
Nuoro	2 372	2.6	80.3	54.7	0.0	2.5	5.4	6.4
South	71 329	2.3	81.1	49.9	0.5	2.5	5.9	9.8
Total	561 891	1.2	84.7	48.8	0.9	2.4	5.5	10.7

Table 9. Quality indexes by cancer registry for cases diagnosed from 2000 through 2004.

presenti nel tracciato record. I dati di ogni registro sono quindi stati sottoposti ad alcune verifiche di qualità previste per la procedura di accreditamento all'AIRTUM.

Dal punto di vista delle tecniche di registrazione, va ribadito che le difformità più rilevanti riguardano la registrazione dei tumori vescicali, specie nella valutazione dei tumori in situ e incerti. La modalità di effettuazione del follow-up è invece sostanzialmente omogenea tra i registri tumori italiani, perché effettuata tramite le basi dati sanitarie disponibili e le schede di decesso, con ricerca attiva dei casi non rintracciabili presso le anagrafi comunali.²⁸ Fattori come la migrazione sanitaria, l'informatizzazione di fonti sanitarie o anagrafiche, la qualità dell'offerta sanitaria e l'estensione del territorio possono tuttavia influenzare l'accuratezza del dato registrato nonché la completezza del follow-up, producendo alcune differenze tra registri tumori.

La valutazione complessiva degli indicatori, per registro e per area geografica, ha giustificato la decisione di includere

systematic loss of cases (in general those with a bad prognosis, which cause death before they can generate other information in the system) and it may lead to overestimate survival, since DCO cases are not included in the survival analysis.

A good indicator of the quality of the documentation available to the registry is the portion of microscopic verifications (MV), which should be high.

Table 8 shows DCO and MV percentages by site and macroarea.

Both the percentage of DCO and microscopically verified cases are within an overall threshold of acceptability. In particular, with the exception of a number of sites which by definition group together ill-defined tumours, there are no DCO percentage values that deserve special attention. For some sites, such as liver and biliary tract, pancreas, brain, and eye, the portion of microscopic verifications is lower, since clinical diagnostic practice is oriented towards other non-invasive techniques in the absence of surgery. Thus, the table must be

tutti i registri tumori presenti nella banca dati AIRTUM nell'analisi di sopravvivenza di coorte 2000-2004.

Indicatori di qualità

DCO e conferme microscopiche

Un tradizionale indicatore di accuratezza e completezza della registrazione è la percentuale di casi individuati solo mediante il certificato di morte (*Death Certificate Only*, DCO). I DCO possono in parte sottendere un fenomeno di perdita sistematica di casi (in genere a cattiva prognosi, che decedono prima di poter generare altre informazioni all'interno del sistema), che può plausibilmente portare a una sovrastima della sopravvivenza in quanto la casistica DCO viene esclusa dall'analisi. Un buon indicatore della qualità della documentazione a disposizione dei registri è invece dato dalla proporzione di verifiche microscopiche (MV) che è bene mostri valori elevati. In tabella 8 sono elencate le percentuali di casi DCO e MV per sede e macroarea.

Si osserva come sia le percentuali di casi DCO sia le percentuali di casi con verifica microscopica si mantengano per la totalità dei casi all'interno di una soglia complessiva d'accettabilità, essendo per i DCO ben al di sotto del 5% (minimo 0,7% per il Centro Italia, massimo 2,3% per il Sud Italia) e per le verifiche microscopiche superiore all'80% (minimo 81,1% per il Sud Italia, massimo 85,9 per il Nord-Est). In particolare, a esclusione di alcune sedi che raccolgono per definizione neoplasie mal definite, non si osservano valori della proporzione di casi DCO degni di particolare attenzione. Per alcune sedi, quali fegato e vie biliari, pancreas, encefalo e occhio, la proporzione di verifiche microscopiche è più bassa (generalmente inferiore al 60%), poiché in assenza d'intervento chirurgico la pratica clinica diagnostica si è orientata verso tecniche non invasive. In generale, quindi, la tabella dovrà essere letta confrontando le macroaree per sede specifica del tumore.

In tabella 9 vengono mostrati gli stessi valori per registro tumori, riferiti al totale delle sedi a esclusione dei carcinomi della cute nel periodo 2000-2004.

Altri controlli

In tabella 9 sono elencati i principali indicatori per registro sulla base della casistica 2000-2004 dei 31 registri. La percentuale di casi persi al follow-up prima di 5 anni varia da 0% al 3,9% (valore globale 0,9%); la percentuale di casi in sede mal definita è compresa tra 0,7% e 5,4% (globale 2,4%). Escludendo il registro tumori della mammella di Palermo, si osserva un valore del rapporto mortalità/incidenza (M/I) compreso tra 44,2% e 54,7%, piuttosto omogeneo per macroarea.

Sono inoltre stati calcolati indicatori quali la percentuale di decessi entro l'anno per diagnosi a buona prognosi (mammella, tiroide, linfoma di Hodgkin, melanoma, corpo uterino) e la percentuale dei sopravvissuti oltre 5 anni dalla diagnosi per i tumori a cattiva prognosi (esofago, fegato, vie biliari, pancreas, mesotelioma). Entrambi questi indicatori sono risultati adeguatamente contenuti e omogenei.

read comparing the macroareas by specific cancer site. Table 9 reports the same indicators by cancer registry for all sites except skin cancer, for the period 2000-2004.

Other indicators

Table 9 reports other main indicators by registry, from the 2000-2004 AIRTUM database. The percentage of cases lost to follow-up before 5 years ranges from 0% to 3.9% (overall 0.9%). The percentage of ill-defined sites varies from 0.7% to 5.4% (overall 2.4%). Reported mortality/incidence (M/I) ratio, excluding the Palermo breast registry, ranges from 44.2% to 54.7%, with a rather homogeneous distribution by macroarea.

Other indicators were also calculated, such as the percentage of deaths within one year for cancers with good prognosis (breast, thyroid, Hodgkin lymphoma, melanoma, corpus uteri) and the percentage of patients alive after more than 5 years from diagnosis for cancers with poor prognosis (oesophagus, liver, biliary tract, pancreas, mesothelioma). Both of these indicators were appropriately small and homogeneous.

Bibliografia/References

1. AIRTUM Working Group. Italian cancer figures, report 2007: survival. *Epidemiol Prev* 2007; 31(1) Suppl 1: 1-96.
2. Fritz A, Percy C, Jack A, Shanmugaratman K, Sobin L, Parkin DM, Whelan S (eds). *International classification of disease for Oncology. Third Edition*, World Health Organization, Geneva, 2000.
3. Brenner H, Hakulinen T. Patients with previous cancer should not be excluded in international comparative cancer survival studies. *Int J Cancer* 2007; 121: 2274-2278.
4. Rosso S, De Angelis R, Ciccolallo L et al. Multiple tumours in survival estimates. *Eur J Cancer* 2009; 45: 1080-1094.
5. Patriarca S, Gafà L, Ferretti S et al. Coding criteria of bladder cancer: effects on estimating survival. *Epidemiol Prev* 2001; 25(3) Suppl: 42-7.
6. Ederer F, Axtell LM, Cutler SJ. The relative survival rate: a statistical methodology. *Natl Cancer Inst Monogr* 1961; 6: 101-21.
7. Cutler SJ, Ederer F. Maximum utilization of the life table method in analyzing survival. *J Chronic Dis* 1958; 8(6): 699-712.
8. Ederer F, Heise H. *Instructions to IBM 650 programmers in processing survival computations. Methodological note No. 10, End Results Evaluation Section*. National Cancer Institute, Bethesda MD, 1950.
9. Hakulinen T. Cancer survival corrected for heterogeneity in patient withdrawal. *Biometrics* 1982; 38(4): 933-42.
10. Hakulinen T, Seppä K, Lambert PC. Choosing the relative survival method for cancer survival estimation. *Eur J Cancer* 2011 May 4; [Epub ahead of print].
11. Pokhrel A, Hakulinen T. How to interpret the relative survival ratios of cancer patients. *Eur J Cancer* 2008; 44(17): 2661-7.
12. Cho H, Howlader N, Mariotto AB, Cronin KA. Estimating relative survival for cancer patients from the SEER Program using expected rates based on Ederer I versus Ederer II method. *Surveillance Research Program*, NCI, Technical Report No. 2011-01.
13. Greenwood M. *The error of sampling of the survivorship table. Reports on public health and medical subjects*. Vol.33. London, Her Majesty's Stationery Office, 1926.
14. Corazziari I, Quinn M, Capocaccia R. Standard cancer patient population for age standardising survival ratios. *Eur J Cancer* 2004 Oct; 40(15): 2307-16.
15. Brenner H, Rachet B. Hybrid analysis for up-to-date long-term survival rates in cancer registries with delayed recording of incident cases. *Eur J Cancer* 2004 Nov; 40(16): 2494-501.
16. Rashid I, Marcheselli L, Federico M. Estimating survival in newly diagnosed cancer patients: use of computer simulations to evaluate performances of different approaches in a wide range of scenarios. *Stat Med* 2008 May 30; 27(12): 2145-58.
17. Engholm G, Gislum M, Bray F, Hakulinen T. Trends in the survival of patients diagnosed with cancer in the Nordic countries 1964-2003 followed up to the end of 2006. Material and methods. *Acta Oncol* 2010 Jun; 49(5): 545-60.
18. Breslow NE, Day NE (eds). *Statistical Methods in Cancer Research - Volume II - The Design and Analysis of Cohort Studies. IARC Scientific Publications No. 82*, Lyon, 1987.
19. Skuladottir H, Olsen JH. Conditional survival of patients with the four major histologic subgroups of lung cancer in Denmark. *J Clin Oncol* 2003; 21(16): 3035-40.
20. Storm HH, Kejs AM, Engholm G, Tryggvadóttir L, Klint A, Bray F, Hakulinen T. Trends in the overall survival of cancer patients diagnosed 1964-2003 in the Nordic countries followed up to the end of 2006: the importance of case-mix. *Acta Oncol* 2010; 49(5): 713-24.
21. Boer R, Ries L, van Ballegooijen M, Feuer EJ, Legler J, Habbema D. Ambiguities in calculating cancer patient survival: the SEER experience for colorectal and prostate cancer. Statistical Research and Applications Branch, NCI, *Technical Report No. 2002-05*.
22. Parkin DM, Plummer M. Comparability and quality of data. In Parkin DM, Whelan SL, Ferlay J, Teppo L, Thomas DB (eds). *Cancer Incidence in Five Continents Vol. VIII*. IARC Sci Publ. 2002; (155): 57-73.
23. Shin HR, Curado MP, Ferlay J, Heanue M, Edwards B, Storm H. Comparability and quality of data. In Curado MP, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M, Boyle P (eds). *Cancer Incidence in Five Continents Vol. IX*. IARC Sci Publ. 2007; (160): 67-94.
24. Ferretti S, Giacomini A, Gruppo di Lavoro AIRTUM (eds). *Manuale di Tecniche di Registrazione dei Tumori*. Inferenze scari, Milano, 2007.
25. Jensen OM, Parkin DM, MacLennan R, Muir CS, Skeet RG (eds). *Cancer Registration Principles and Methods. IARC Scientific Publication No. 95*, Lyon, 1991.
26. Tyczynski JE, Démaret E, Parkin DM. Standards and Guidelines for Cancer Registration in Europe. *IARC Technical Publication No. 40*, Lyon, 2003.
27. Ferlay J, Burkhard C, Whelan S, Parkin DM. Check and conversion programs for cancer registries (IARC/IACR Tools for Cancer Registries). *IARC Technical Report No. 42*, Lyon, 2005.
28. Ferretti S, Federico M, Contiero P, Stracci F. Comparability issues of survival data. *Epidemiol Prev* 2001; 25(3) Suppl: 37-41.