MATERIALS AND METHODS
MATERIALI E METODI

INTRODUCTION

Until a few years ago, there was no internationally agreed definition of rarity for cancers. According to the European Parliament and Council of the European communities, rare diseases are defined as those with a prevalence of <50 per 100,000,[1] while, in line with the Orphan Drug Act, in the US rare diseases are those affecting <200,000 persons.[2] However, a published analysis of rare cancers in the US employed the definition of <15 incident cases per 100,000 per year.[3] In 2010, the EU funded RARECARE – Surveillance of rare cancers in Europe project[4] proposed a new definition of rarity for cancers and a list of rare cancers in Europe.

The objective of this Italian Network of Cancer Registries (AIRTUM) monograph is to describe the burden of rare cancers in Italy, in order to give information on rare cancers to the community at large (oncologists, general practitioners, researchers, health authorities, patients and their families).

In detail, this monograph provides estimations of incidence, survival, and prevalence of rare cancers in Italy, based on the definition and list of rare cancers proposed by RARECARE and recently (2012) updated by the RARECAREnet – Information Network on Rare Cancers project.[5]

This is the first time that such a comprehensive and detailed description of rare cancers has been provided in Italy and it was possible only thanks to the availability of AIRTUM's large database.[6]

THE RARECARE DEFINITION OF RARE CANCERS

RARECARE defined rare cancers as those with an incidence rate (IR) of <6 per 100,000 per year in the European population. It is important to stress that rare cancers among the RARECARE list of cancers (please refer to Annex 1, see supplementary material on-line) are identified based on the above criterion, in other words, on the basis of the European population and not of a country-specific population. Thus, rare cancers are always the same in all European countries. It is also important to note that this is an incidence-based rather than prevalence-based definition, since incidence was recognised as the best indicator to define rarity for tumours.[4]

INCIDENCE vs. PREVALENCE

Prevalence has shortcomings as a measure for rarity for tumours, although it is appropriate for non-neoplastic diseases. Many non-neoplastic diseases are chronic conditions, so prevalence, which reflects the total number of cases at any given time in a population, truly reflects the burden that a disease poses at a population level. On the contrary, tumours are sub-acute diseases in which everything tends to happen once: in the natural history of a tumour, there will be one potentially eradicating surgery, one local radiation therapy, one first chemotherapy, and each of these will take place in a definite time interval. Thus, the total amount of resources that tumours mobilise is proportional to the yearly rate of new diagnoses (incidence) and not to the total number of persons with previous cancer diagnosis (prevalence), some of whom have been cured. Incidence, which reflects the yearly number of new cases occurring in a population, might thus be a better indicator to describe the burden posed by a tumour.

Moreover, the prevalence of a disease depends on two time-dependent characteristics which are independent of one another: incidence and survival. With the prevalence threshold adopted as a definition of cancer rarity, some commonly occurring diseases for which survival is very poor, such as most cancers of the stomach (adenocarcinoma of stomach), pancreas, adenocarcinoma of lung, and squamous cell carcinoma of lung, will be defined as rare, since the proportion of the general population who are survivors is very low. By contrast, some neoplasms that occur very infrequently ("rare" in the sense of incidence) but which have very good survival, such as cancer of the testis and squamous cell carcinoma of the uterine cervix, will be defined as common on the basis of prevalence, because, although they occur infrequently, most people who develop the disease survive for long periods.

For these reasons, incidence seems to be a more useful indicator to select a threshold for rarity in the case of tumours, as opposed to non-neoplastic diseases. In addition, it is worth stressing that:

- incidence is a direct measure of the burden imposed by the need for first-line cancer treatment;
- the number of patients amenable to enter a clinical study is reflected by cancer incidence.

Any threshold for cancer rarity should be considered as merely indicative. The RARECARE rarity threshold at <6 per 100,000 per year might be considered too high. However, if the lower threshold of <3 per 100,000 per year were adopted, glial tumours, epithelial cancers of the oral cavity, epithelial cancers of gallbladder and extrahepatic biliary tract, soft tissue sarcomas, tumours of testis and paratestis, myeloproliferative neoplasms, and acute myeloid leukaemia would all be excluded. Yet these cancers are often inadequately diagnosed and treated in relation to both lack of knowledge and lack of clinical expertise, and clinical trials are rarely performed. They are all diseases that are best treated in specialised centres.

Crocetti E, Trama A, Stiller C, et al. Epidemiology of glial and non-glial brain tumours in Europe. *Eur J Cancer* 2012;48(10):1532-1542.

Stiller CA, Trama A, Serraino D, et al. Descriptive epidemiology of sarcomas in Europe: report from the RARECARE project. *Eur J Cancer* 2013;49(3):684-695.

THE RARECARE LIST OF CANCERS

Usually, cancer statistics are provided for broad cancer categories, based on the anatomic site of the malignancies as defined by the International Classification of Diseases (ICD) codes. Rare tumour entities, because of their specific problems related to health care organisation and clinical management, might be more appropriately defined as a combination of topographical and morphological characteristics, as both defined by the International Classification of Diseases for Oncology, 3rd edition (ICD-O-3).[7]

Each tumour entity in the ICD-O list has a pathologic basis; however, in order to define clinically distinct diseases, the pathological entities have to be grouped. This grouping exercise, necessary to identify a list of clinically distinct entities, was undertaken in the context of the RARECARE project by an international group of experts, including oncologists, epidemiologists, pathologists, and organisations of patients.[4]

As a first step, the two large groups of epithelial and non-epithelial tumours were disentangled and, within them, broad anatomic categories were identified. Thus, the list of rare cancers starts with epithelial tumours of different sites and continues with non-epithelial tumours such as sarcomas, neuroendocrine, central nervous system, and haematological tumours (please refer to Annex 1).

The RARECARE list is organised into three tiers, as illustrated, for example, in Table 1 for epithelial tumours of nasal cavity and sinuses. The bottom tier (tier 3) on the list is the WHO name of individual cancer entities[8] and its corresponding ICD-O-3 morphology and topography codes.[7] ICD-O-3 entities are grouped into categories (tier 2) of cancers, considered similar from the point of view of clinical management and research. These categories are then further grouped into more general categories of tumours (tier 1), considered to involve the same clinical expertise and patient referral structure.

Tier 2 entities, by definition, include only specific morphologies; thus, rare cancers are identified in this tier. Not Otherwise Specified (NOS) morphology codes (NOS: 8000, 8001 for solid cancers, and 9590,9591,9760,9800,9801,9820,9860 for haematological diseases) are never assigned to tier 2, but to tier 1, which aims at identifying tumours with the same referral structure.

The choice of basing the definition of rare cancer on topography and morphology according to ICD-O-3 was made for two reasons. The first reason was to follow the existing tumour classifications. Any list of rare tumours will always be a subset of a standard list of tumours. International agencies preside over such classifications, constantly updating them. This list of rare tumours was based on the ICD-O-3 classification because this is the worldwide-recognised classification of tumours.

The second reason was data availability. Population-based cancer registry (CR) data, the only data available to calculate population-based incidence and prevalence indicators, refer to cases classified only according to ICD-O. Other, even attractive, classification criteria, such as biomarkers or gene expression, cannot be used for any quantitative description of cancer burden in a wide population.

The new RARECAREnet project on rare cancers reviewed the RARECARE list of cancers in 2012 and identified 198 rare cancers in tier 2 with a European IR (independently of the country-specific IR) <6 per 100,000 per year. The complete list of the 198 tier 2 rare cancers is provided in Annex 1 (rare cancers are identified by the R).

This monograph considers all these 198 rare cancers, including 5 cancers which have an IR that is higher than 6 per 100,000 per year in Italy: • hepatocellular carcinoma of liver (Italy 9.4 vs. EU 3.1), • squamous cell carcinoma of larynx (Italy 7.2 vs. EU 4.6), • carcinoma of thyroid gland (Italy 14.2 vs. EU 3.7), • multiple myeloma (Italy 8.4 vs. EU 5.9), • diffuse and follicular B-cell lymphoma (Italy 9.8 vs. EU 4.9). Thyroid carcinoma is not presented in the site-specific commentaries because of the peculiar characteristics of thyroid carcinoma in Italy (see paragraph «Methodological issue», pp. 20-21). The 198 rare cancers are classified, for this monograph, in 14 major groups (Table 2) considering the clinical referral pattern and the interest of clinicians. Thus, all sarcomas of soft tissue (regardless of the site of origin), bone and gastro-intestinal stromal tumours are grouped together in the sarcoma group. The rationale behind this choice was that all sarcomas should be referred to sarcoma specialists and these experts are very likely interested in having data on all different types of sarcoma. The main rare cancers of the thoracic cavity (thymoma, trachea, ma-

TIER	TUMOUR	TOPOGRAPHY ICD-O-3 CODE	MORPHOLOGY ICD-O-3 CODE
1	**EPITHELIAL TUMOURS OF NASAL CAVITY AND SINUSES**	C30.0, C31	8000, 8001, 8004, 8010, 8011, 8020-8022, 8032, 8050-8076, 8078, 8082-8084, 8123, 8144, 8560, 8980
2	**Squamous cell carcinoma and variants of nasal cavity and sinuses**	C30.0, C31	8004, 8032, 8050-8076, 8078, 8083-8084, 8123, 8560, 8980
3	Squamous cell carcinoma	C30.0, C31	8070
3	Verrucous carcinoma	C30.0, C31	8051
3	Squamous cell carcinoma, spindle cell	C30.0, C31	8004, 8032, 8074, 8980
3	Papillary squamous cell carcinoma	C30.0, C31	8052
3	Adenosquamous carcinoma	C30.0, C31	8560
3	Squamous cell carcinoma, adenoid	C30.0, C31	8075
3	Basaloid squamous cell carcinoma	C30.0, C31	8083
2	**Lymphoepithelial carcinoma of nasal cavity and sinuses**	C30.0, C31	8082
2	**Undifferentiated carcinoma of nasal cavity and sinuses**	C30.0, C31	8020-8022
2	**Intestinal type adenocarcinoma of nasal cavity and sinuses**	C30.0, C31	8144

Table 1. The hierarchical three-tier structure of the RARECARE list of cancers illustrated for epithelial tumours of nasal cavity and sinuses.

MAJOR GROUPS	RARE TUMOURS
Epithelial tumours of head and neck	Epithelial tumours of nasal cavity and sinuses, nasopharynx, hypopharynx and larynx, oropharynx, oral cavity and lip, middle ear, major salivary glands and salivary gland type tumours
Tumours of the eye	Epithelial tumours of eye and adnexa, malignant melanoma of uvea
Rare epithelial tumours of the digestive system and mesothelioma of peritoneum	Rare epithelial tumours of stomach, colon, rectum, pancreas, epithelial tumours of oesophagus, small intestine, anal canal, liver and intrahepatic bile tract, gallbladder, and extrahepatic biliary tract, and mesothelioma of peritoneum
Rare epithelial tumours of the thoracic cavity and mesothelioma of pleura and pericardium	Epithelial tumours of trachea, thymus, rare epithelial tumours lung, and mesothelioma of pleura and pericardium
Rare tumours of the female genital system	Rare epithelial tumours of breast, and corpus uteri, epithelial tumours of cervix uteri, ovary and fallopian tube, vulva and vagina, trophoblastic tumours of placenta, non epithelial tumours of ovary and epithelial tumours of males breast
Rare epithelial tumours of the urinary system	Epithelial tumours of renal pelvis and ureter, and urethra; rare epithelial tumours of kidney and bladder
Rare tumours of the male genital system	Epithelial tumours of penis, testicular and paratesticular cancers, extragonadal germ cells tumours, rare epithelial tumours of prostate, and mesothelioma of tunica vaginalis
Rare skin tumours and malignant melanoma of mucosa (extracutaneaous melanoma)	Rare skin tumours (adnexal carcinoma of skin) and malignant melanoma of mucosa (extracutaneous melanoma)
Embryonal tumours	Neuroblastoma and ganglioneuroblastoma, nephroblastoma, retinoblastoma, hepatoblastoma, pleuropulmonary blastoma, pancreatoblastoma, olfactory neuroblastoma, odontogenic malignant tumours
Sarcomas	Soft tissue sarcomas, bone sarcomas, gastrointestinal stromal tumours, Kaposi sarcoma
Neuroendocrine tumours	Neuroendocrine tumours of lung, gastroenteropancreatic tract, skin, thyroid, of other sites, pheochromocytoma, paraganglioma
Tumours of the Central Nervous System (CNS)	Central Nervous System tumours and embryonal tumours of CNS
Tumours of the endocrine organs	Carcinoma of pituitary gland, parathyroid gland, and adrenal cortex
Rare haematological diseases	Rare lymphoid diseases, acute myeloid leukemia and related precursor neoplasms, myeloproliferative neoplasms, myelodysplastic syndrome and myelodysplastic/myeloproliferative diseases, hystiocytic and dendritic cell neoplasms

Table 2. List of the 14 major groups of rare cancers included in this AIRTUM monograph.

lignant pleural mesothelioma, and rare histotypes of lung) are placed together in the group of rare epithelial thoracic cavity tumours (including pleural mesothelioma) because all these tumours should be referred to experts of lung and thoracic cavity cancers. With the same rationale, cancers of the nasal cavity and sinuses, nasopharynx, major salivary glands and salivary gland type tumours, hypopharynx, larynx, oropharynx, oral cavity, and lip are grouped together as head and neck cancers because they should all be referred to head and neck cancer experts. The same rationale was applied to the other rare cancers arising in the major sites of the body: digestive system, female genital system, urological tract, male genital system, central nervous system, eye, skin. In addition to these sites, 4 other groups included are neuroendocrine tumours regardless of their site of origin, all endocrine tumours (carcinoma of thyroid excluded), all embryonal tumours, and all rare haematological diseases.

MATERIALS AND METHODS

Epidemiology of rare cancers presented in this monograph is based on incidence, survival, and prevalence estimates. On the basis of national or local data, mortality rates for each RARECAREnet considered cancer entity cannot be computed. Besides, the available mortality rates by major cancer site pose some limits, as in the case of uterine cancers: the poor specification of the subsites in official death statistics makes it impossible to disentangle mortality between cervix and corpus uteri; this is confirmed by a high proportion of deaths attributed to not otherwise specified uterine cancer.[9]

The AIRTUM database

This monograph is based on the AIRTUM database updated to January 2015. The AIRTUM[6] at present includes 40 general population-based CRs and 5 specialised CRs.

Since 2005, AIRTUM has had a central database, which stores data from all accredited CRs. Accreditation indicates a CR meets the quality standards set by AIRTUM,[10] which verifies data quality and completeness and uses data for collaborative studies on cancer epidemiology in Italy.[11-14]

All cases in the AIRTUM database are coded according to the third edition of the ICD-O.[7] Data are double checked (by the Registry and by the centralised database) with the DEPedits program. In addition, other checks are carried out based on software developed by AIRTUM, CheckAIRTUM, that compares data from a specified registry to the weighted average of the other registries of the database.[6]

Quality checks

The following data quality indicators, usually considered in international population-based survival studies like EUROCARE and RARECARE, were calculated for incident malignant cancers collected between 2000 and 2010, by CR:

1. proportion of cases known from death certificate only (DCO);
2. proportion of microscopically verified (MV) cases;
3. proportion of cases with survival time of zero days (date of diagnosis coincident with date of life status ascertainment);
4. proportion of cases diagnosed incidentally at autopsy;
5. proportion of Not Otherwise Specified (NOS) cases.

AREA MAJOR GROUPS	MALIGNANT CASES DIAGNOSED BETWEEN 2000 AND 2010	NOS[$]	DCO	AUTOPSY ONLY CASES	CASES WITH ZERO SURVIVAL TIME	MICROSCOPICALLY CONFIRMED CASES	LOST TO FOLLOW-UP CASES
	No.	%	%	%	%	%	%
AIRTUM POOL							
Epithelial tumours of head and neck	43 163		0.3	0.2	0.2	97.1	0.8
Tumours of the eye	1 530		0.6	0.2	0.2	58.9	0.7
Digestive system tumours	358 109		1.4	0.6	0.2	78.7	0.6
Rare epithelial tumours of the digestive system	57 891		0.7	1.2	0.2	73.5	0.6
Thoracic cavity tumours	157 478		1.6	1.0	0.2	74.4	0.5
Rare epithelial tumours of the thoracic cavity	12 027		0.1	2.8	0.2	97.4	0.5
Female genital system tumours	246 903		0.6	0.1	0.2	95.5	1.0
Rare tumours of the female genital system	41 141		0.1	0.2	0.2	99.3	1.0
Urinary system tumours	104 116		0.6	0.4	0.2	89.0	0.9
Rare epithelial tumours of the urinary system	6 394		0.2	0.5	0.2	94.5	0.6
Male genital system tumours	152 102		0.8	0.3	0.2	92.3	0.8
Rare tumours of the male genital system	9 049		0.1	0.1	0.3	95.4	1.8
Tumours of the Central Nervous System	13 071		0.0	0.5	0.2	91.2	1.0
Haematological diseases	123 307		0.8	0.5	0.3	93.1	0.9
Rare haematological diseases	91 094		0.4	0.5	0.3	94.4	0.9
Skin tumours^	33 823		0.1	0.0	0.3	97.5	5.9
Rare skin tumours and malignant melanoma of mucosa	1 699		0.0	0.0	0.5	99.5	1.5
Embryonal tumours	859		0.1	0.0	0.5	93.2	0.9
Sarcomas	20 019		0.0	0.4	0.3	98.6	1.4
Neuroendocrine tumours	9 196		0.0	0.5	0.3	99.6	0.8
Tumours of the endocrine organs*	32 268		0.2	0.2	0.3	95.4	1.2
NORTH-WEST							
All malignant cancers^	445 918	14.7	1.1	0.1	0.3	85.7	1.0
Rare cancers	111 744	–	0.3	0.0	0.4	92.1	1.3
NORTH-EAST							
All malignant cancer^	470 760	13.4	0.8	1.1	0.1	87.3	0.2
Rare cancers	116 808	–	0.1	0.0	0.1	93.5	2.0
CENTRE							
All malignant cancer^	125 671	14.9	0.7	0.0	0.1	85.1	1.6
Rare cancers	31 005	–	0.3	0.2	0.2	90.0	0.8
SOUTH							
All malignant cancer^	282 706	19.2	2.1	0.0	0.3	82.5	0.9
Rare cancers	79 846	–	0.9	0.5	0.0	91.1	0.0
AIRTUM POOL							
All malignant cancer^	1 325 055	15.2	1.2	0.4	0.2	85.5	0.8
Rare cancers	339 403	–	0.9	0.0	0.0	92.1	0.6

[$] **NOS**: ICD-O-3 morphological code: 8000,8001,9800,9590,9820,9760,9860,9800,9801 / * including carcinomas of thyroid gland / ^ excluding non melanoma skin cancer

Table 3. Number of cases diagnosed in 2000-2010 and data quality indicators for the 14 major groups of rare cancers and corresponding common cancers (when the group included also common cancers), and for all malignant tumours vs. all rare cancers, by geographic area and in the overall AIRTUM Pool. Quality indicators include proportion of not otherwise specified (NOS) morphologies (8000,8001 for solid cancers, and 9590, 9591, 9760, 9800, 9801, 9820, 9860 for haematological diseases), death certificate only (DCO) cases, autopsy only cases, cases with zero survival time, microscopically confirmed cases, and lost to follow-up cases (follow-up time <5 years). Pool of 39 general CR of the AIRTUM database.

All standard indicators of data quality for Italian CRs are satisfactory according to international standards.[15]

The new and most relevant indicator to evaluate the accuracy of diagnosis for rare cancers, with respect to the other routinely applied indicators, is the proportion of cases with a NOS category (ICD-O-3 8000-8001 for solid cancers, and ICD-O-3 9590-9591, 9760, 9800-9801, 9820, 9860 for haematological diseases). For rare cancers, the most likely quality problem is lack of specificity of morphology codes, which make it impossible to assign such cases to a specific (rare) cancer entity, resulting in underestimation of the true incidence and prevalence of such entities. Unspecified morphology can be due to genuine difficulty in assigning a specific morphological category or because inadequate documentation was supplied to the CR when the case was registered. The latter problem is registration bias and results in incidence and prevalence underestimation. To assess the extent of registration bias at European level, RARECARE reviewed the original data (mainly pathologic reports) of a selected sample (about 18,000 cases) of eight rare cancers (for details see RARECARE web site).[5] Briefly, the great majority of NOS morphology cases were confirmed as NOS. The few NOS cases that changed to a more specific diagnosis generally increased the incidence of the more common cancer forms. For example, 11% of epithelial oral cavity cancers were reclassified from NOS to more specific diagnoses: 8% were reclassified as squamous

cell carcinoma (more common) and only 3% as adenocarcinoma (rarer). This finding suggests that the problem with poorly specified morphology cases is mainly one of difficulty in reaching a precise diagnosis, not registration bias. However, it raises an important topic for collaboration with pathologists and CRs.

Table 3 shows quality indicators for the 14 major groups of rare cancers included in this monograph, compared, when possible, to common cancers, and the quality indicators for all tumours in the AIRTUM pool and by geographic area. The overall proportion of NOS was 15%, with a higher proportion in the CRs of the South of Italy. Overall, the proportion of DCO cases was 1.2%, with, again, a slightly higher proportion in the CRs of the South of Italy. The proportion of cases discovered at autopsy was 0.4% in total. A high proportion of cases (86% overall) was MV. Follow-up was complete for most CRs, with follow-up censored before 5 years for only 1% of cases overall. These results indicate a high quality dataset.

NOS cases were also analysed within each of the 14 major groups of cancers presented in this monograph. As expected, grouping with major problems were those of the digestive system (mainly liver and pancreas), thoracic cavity, and CNS. NOS cases for these groups were 21%, 29%, and 38%. The difficult access to these sites led to a proportion of MV cases lower than the one observed in other sites. For example, the proportion of MV cases was 58% for tumours of the pancreas and 74% for tumours of the liver. Incidence and prevalence indicators of rare cancers of lung, liver, and pancreas, and specific histotypes of the CNS might therefore be underestimated. Rare cancers always had a higher proportion of MV cases compared to the common counterpart, except in the digestive system. This is due to the fact that rare cancers of the digestive system include all cancers of liver and intrahepatic bile tract and of gallbladder and extrahepatic biliary tract, for which the proportion of MV cases was very low (70% and 58%, respectively), therefore the NOS was high. This influenced the overall proportion of NOS cases for rare digestive system cancers, making it higher than that for common cancers of the digestive system, including cancers of more accessible sites such as stomach and colon.

Cancer registry selection

In order to guarantee homogeneity, all indicators were computed on the same database, therefore only general population-based CRs were considered. Thus, we did not consider in the analyses 3 specialised (Palermo-breast cancer; Modena-colorectal cancers; Liguria-mesothelioma) CRs, and 2 other CRs which collect data on childhood and adolescent cancers only (Piemonte, Marche).

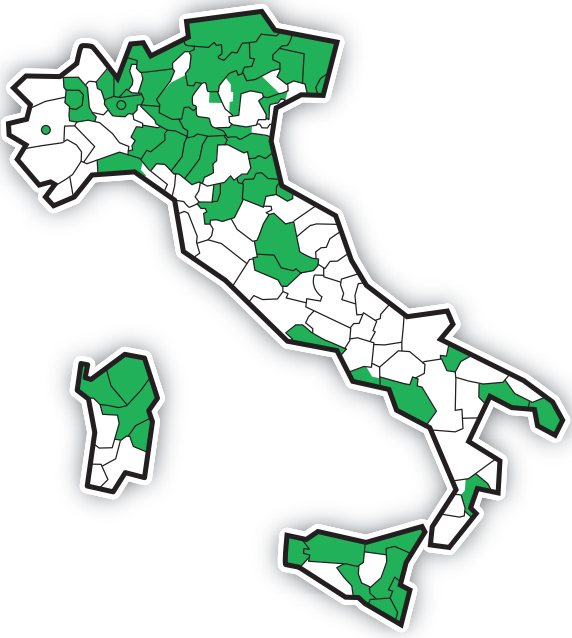
Three CRs (Biella, Napoli, Ragusa) extended their area of registration in recent years: different areas of the same CR are therefore analysed separately in incidence and survival analyses.

To provide estimates of epidemiological indicators, the following inclusion criteria for AIRTUM CRs were applied:

- availability of at least three years of incidence between 2000 and 2010;
- complete follow-up for at least one year after the last year of incidence (i.e. at 31st December 2009) for cases diagnosed between 2000 and 2008 in survival analysis;
- proportion of NOS cases <20%.

All AIRTUM accredited CRs had a proportion of NOS cases

MACROAREA/ REGION	RESIDENT POPULATION (ITALY 2013)	RESIDENTS IN AREAS COVERED BY CANCER REGISTRIES INCLUDED IN THE PRESENT MONOGRAPH (ITALY 2013)		CANCER REGISTRIES
	No.	No.	%	No.
Piemonte	4 374 052	1 229 824	28	2
Valle d'Aosta	127 844	0	0	0
Lombardia	9 794 525	8 307 271	85	10
Liguria	1 565 127	851 283	54	1
NORTH-WEST	15 861 548	10 388 378	65	13
Trentino-Alto Adige	1 039 934	1 039 934	100	2
Veneto	4 881 756	2 346 610	48	1
Friuli Venezia Giulia	1 221 860	1 221 860	100	1
Emilia-Romagna	4 377 487	3 500 936	80	6
NORTH-EAST	11 521 037	8 109 340	70	10
Toscana	3 692 828	1 235 646	33	1
Umbria	886 239	886 239	100	1
Marche	1 545 155	0	0	0
Lazio	5 557 276	552 090	10	1
CENTRE	11 681 498	2 673 975	23	3
Abruzzo	1 312 507	0	0	0
Molise	313 341	0	0	0
Campania	5 769 750	2 262 522	39	2
Puglia	4 050 803	1 776 450	44	3
Basilicata	576 194	0	0	0
Calabria	1 958 238	228 126	12	1
Sicilia	4 999 932	4 381 032	88	5
Sardegna	1 640 379	688 066	42	2
SOUTH AND ISLANDS	20 621 144	9 336 196	45	14
ITALY	59 685 227	30 507 889	52	39



▲ **Figure 1.** Italian geographical areas covered by the general cancer registries included in the present monograph.

◀ **Table 4.** Distribution of the Italian resident population by region, macroarea, and overall. Number (No.) and proportion (%) of the resident population covered by the cancer registries included in this monograph and number of general cancer registries by region, macroarea, and overall. Italy, 2013. (**Source**: ISTAT).[16]

<20%, thus none of the CRs were excluded because of data quality problem. The CR of Macerata was excluded from all analyses because it did not fulfil the first inclusion criterion (it did not have at least 3 years of incidence between 2000 and 2010). General CRs included in this monograph cover more than 30 million people, 52% of the Italian population at 2013 (Table 4 and Figure 1).
In the framework of the RARECARE and RARECAREnet projects, all participating Italian AIRTUM CRs were considered to estimate the burden on rare cancers in Europe because they completely fulfilled the quality criteria and the sensitivity analyses performed.[5]
In order to reduce the uncertainty due to the casual variability of sparse data and the resulting imprecision of the estimates, we considered three different pools of AIRTUM CRs for the different analyses (Table 5); in detail:

- 39 CRs with at least three years of available incidence data on patients registered between 2000 and 2010, for the incidence analysis;
- 37 CRs with cases diagnosed between 2000 and 2008, and followed up to 31st December 2009 or after, for the survival analysis;
- 11 CRs which provided incidence and follow-up data for the period 1992-2006 with a prevalence index date of 1st January 2007 for the prevalence analysis.

Table 5 shows data availability by year of incidence and CRs included in the analyses.

Epidemiological indicators

The epidemiological indicators are estimated considering multiple tumours. The inclusion of multiple tumours in the analyses implies that each single patient may be counted several times. We considered 1,325,055 malignant cancer cases collected by 39 Italian CRs during the 2000-2010 period and included in the AIRTUM database as of January 2015.
For neuronal and mixed neuronal-glial tumours none of the indicators are provided, since they are mainly benign or borderline tumours and thus not available in the AIRTUM database. In addition, incidence of a few entities, including gastrointestinal stromal tumours and several haematological malignancies, is underestimated because the specific ICD-O codes were introduced with the ICD-O-3 in 2000, thus during the study period.
As some rare cancers are extremely rare, estimates stratified by geographic area could not be calculated.
All estimates were computed using SeerStat, version 8.1.2.[17]

Incidence

Crude IRs of rare cancers were calculated as the number of new cancers occurring in 2000-2010 divided by the population at risk (male and female also for gender-specific rare cancers) over the same period, expressed as person-years. In total, 339,403 rare tumours were included in the incidence analysis from 39 CRs. The proportion of rare cancers out of the total cancers (rare and common) by site is also calculated considering rare and common cancers of each specific cancer site. Crude incidence was obtained for rare tumours overall, by sex, and by age class (0-54, 55-64, 65+; 0-4, 5-14, 15+ for embryonal cancers). The normal approximation is used with the standard errors to obtain 95% confidence intervals (95%CI) for incidence rates. Sex- and age-specific incidence rate for 25 rare cancers with less than 15 observed cases between 2000

MACROAREA CANCER REGISTRY/ GEOGRAPHICAL AREA		AVAILABLE INCIDENCE YEARS	INCIDENCE	SURVIVAL	PREVALENCE (OBS)
			PERIOD OF DIAGNOSIS		
			2000-2010	2000-2008 follow-up at 31st Dec 2009	1992-2006 prevalence date at 1st Jan 2007
			CRs INCLUDED No. YEARS INCLUDED		
NORTH-WEST					
Bergamo		2007-2009	✓ (3 yrs)	✓	
Biella	Biella	1995-2009	✓ (9 yrs)	✓	
	Vercelli	2007-2009	✓ (3 yrs)	✓	
Brescia		1999-2006	✓ (7 yrs)	✓	
Como		2003-2009	✓ (7 yrs)	✓	
Cremona		2005-2009	✓ (5 yrs)	✓	
Genova		1986-2007	✓ (8 yrs)	✓	✓
Mantova		1999-2010	✓ (11 yrs)	✓	
Milano (municipality)		1999-2007	✓ (8 yrs)	✓	
Milano 1-2		2007-2009	✓ (3 yrs)	✓	
Monza e Brianza		2007-2009	✓ (3 yrs)	✓	
Sondrio		1998-2010	✓ (11 yrs)	✓	
Torino		1985-2010	✓ (11 yrs)	✓	✓
Varese		1976-2010	✓ (11 yrs)	✓	✓
NORTH-EAST					
Alto Adige		1995-2007	✓ (8 yrs)	✓	
Ferrara		1991-2009	✓ (10 yrs)	✓	✓
Friuli Venezia Giulia		1995-2009	✓ (10 yrs)	✓	
Modena		1988-2010	✓ (11 yrs)	✓	✓
Parma		1978-2011	✓ (11 yrs)	✓	✓
Piacenza		2006-2010	✓ (5 yrs)	✓	
Reggio Emilia		1996-2010	✓ (11 yrs)	✓	
Romagna		1986-2009	✓ (10 yrs)	✓	✓
Trento		1995-2006	✓ (7 yrs)	✓	
Veneto		1987-2007	✓ (8 yrs)	✓	✓
CENTRE					
Firenze-Prato		1985-2005	✓ (6 yrs)		
Latina		1990-2010	✓ (11 yrs)	✓	✓
Umbria		1994-2009	✓ (10 yrs)	✓	
SOUTH AND ISLANDS					
Barletta		2006-2008	✓ (3 yrs)	✓	
Catania-Messina		2003-2008	✓ (6 yrs)	✓	
Catanzaro		2003-2007	✓ (5 yrs)	✓	
Lecce		2003-2007	✓ (5 yrs)		
Napoli	ex ASL 4	1996-2010	✓ (11 yrs)	✓	
	ASL 3 Sud	2008-2010	✓ (3 yrs)	✓	
Nuoro		2003-2008	✓ (6 yrs)	✓	
Palermo		2003-2010	✓ (8 yrs)	✓	
Ragusa	Ragusa	1981-2009	✓ (10 yrs)	✓	✓
	Caltanissetta	2007-2010	✓ (4 yrs)	✓	
Salerno		1996-2009	✓ (10 yrs)	✓	
Sassari		1992-2009	✓ (10 yrs)	✓	✓
Siracusa		1999-2009	✓ (10 yrs)	✓	
Taranto		2006-2008	✓ (3 yrs)	✓	
Trapani		2002-2007	✓ (6 yrs)	✓	
AIRTUM POOL		**1976-2010**	**39**	**37**	**11**

Table 5. Available incidence years by general cancer registries (CRs) considered in this monograph ordered by macroarea. CRs included in incidence analysis with number of incident years considered. Cancer registries included in survival and observed prevalence analyses. AIRTUM database at January 2015.

and 2010 were considered as not estimable (NE). The expected number of new cases in 2015 was estimated assuming incidence in Italy to be the same as that in the AIRTUM sample, and multiplying the age- and sex-specific incidence rate by the corresponding Italian population in 2015 provided by ISTAT.[16]

Survival

Survival analysis was performed for 254,821 rare tumours. One- and 5-year relative survival (RS) estimates[18] in Italy were obtained considering the pool of 37 AIRTUM CRs with cases diagnosed between 2000 and 2008, and followed up to 31st December 2009 or after. Relative survival is defined as the ratio of observed survival to the expected survival in the general population of the same age and sex and it is used to correct for deaths from causes other than the cancer under investigation. RS was calculated for patients aged 0-99. Since all patients are included (not only those followed up for 5 years) we used a *complete* analysis, which is a modification of traditional cohort approach, in which more recently diagnosed patients are also included, even if they could not possibly have completed the entire follow-up interval of interest.[19] Cancers diagnosed only on the basis of DCO, or diagnosed incidentally at autopsy or with survival time of zero days (421 tumours, 0.2%), were excluded from the analysis. Ninety-five% CI are computed through logarithmic transformation, so that the lower bound is always positive and the upper bound can exceed 100%. Whenever it happens, the upper bound is put as equal to 100%. RS for 42 rare cancers with less than 30 observed cases in the period of diagnosis 2000-2008 was considered as not estimable.

A sensitivity analysis of survival was performed restricting the analysis to 29 CRs with cases diagnosed between 2000 and 2010, and follow-up available to 31st December 2011 (data not shown). This analysis was performed to verify whether the use of more updated data would have had an impact on the survival estimates provided. In the sensitivity analysis, 11 CRs were excluded. The analysis performed on this restricted pool (29 CRs) shows percentage differences higher than 10% only for 18 rare cancers among the 198 considered ones in 1-year RS and for 42 rare cancer entities in 5-year RS. However, when considering only 29 CRs the number of analysed cases becomes smaller and the uncertainty in the estimates increases; therefore, considering the rarity of the phenomena and the narrow changes between the two analyses, we decided to present results for the 37 CR pool.

Observed and complete prevalence

To estimate the observed prevalence in Italy – the proportion of cancer patients in a population diagnosed at age x within a given time period (L) and who are still alive at a certain reference date – incidence and follow-up data from 11 CRs for the period 1992-2006 were used, with a prevalence index date of 1st January 2007. Observed prevalence in the general population (male and female also for gender-specific rare cancers) disentangled by time prior to the reference date (<2 years, 2-5 years, <15 years) was calculated using the counting method.[20-22] When including multiple tumours, a patient will not contribute more than one tumour diagnosis to a single prevalence estimate.

The life status of cases lost to follow-up or censored before the prevalence index date was estimated from the survival probability between the censoring and the index date, derived from a subset of cancer patients matched by age and cancer.

The objective of the present monograph is to produce reliable prevalence estimations for all the 198 rare cancers. In order to achieve this objective, a standard methodology, applicable to all these rare cancers, had to be defined. The complete prevalence proportion at 1st January 2007 was estimated overall in Italy, correcting the 15-year observed prevalence by the completeness index,[23-26] to account for those cancer survivors diagnosed before the cancer registry activity started. The completeness indices estimation requires a long stable time series of incidence and survival indicators; the Italian cancer registry database, even though it is continuously increasing, could not guarantee such information for all the 198 analysed rare cancers. Moreover, sparse data did not allow to consider the geographical variability (stratifying the estimates by geographic area) of the prevalence estimates as in the previous AIRTUM monograph on prevalence for common cancers.[14]

To allow more robust estimation, the completeness indices by cancer site and age (0-4, […], 75-99 years), were obtained by means of statistical regression models using incidence and survival data available in the European CRs participating to the RARECARE[4,27] and RARECAREnet projects.

The assumptions are:

- homogeneity of time trends in incidence and survival between Italy and Europe;
- homogeneity of prevalence proportions of rare cancers among geographic areas in Italy.

For cancers with no observed cases within 2, 2-5, or 15 years in the past, prior to 1st January 2007, the observed prevalence was considered as not estimable. If the 15-year observed prevalence is not estimable (NE), then the complete prevalence has to be considered not estimable. This is the case for 13 rare cancers.

Finally, the number of prevalent cases at 1st January 2010 in Italy was calculated assuming the same prevalence proportion as in the AIRTUM sample and applying the obtained complete prevalence proportion by age (0-4, […], 75-99 years) to the corresponding Italian population at 2010 provided by ISTAT.[16]

For the purpose of including as much as possible cancer registries in the prevalence analysis, the reference date is 1st January 2007. Different and more recent reference dates would have determined a restriction of the analysed POOL.

The uncertainty that characterise rare cancers made impossible the projection of prevalence estimates to 2015, therefore only the number of prevalent cases at 1st January 2010 in Italy was calculated.

Methodological issues

The methodological decisions, taken because of the rarity of the majority of the cancers analysed, could have led to prevalence estimates slightly different from those published by the AIRTUM monograph on prevalence for common cancers.[14]

For prevalence, the rationale of using the same methodology for all rare cancers made necessary the assumption of homogeneity between some European and Italian epidemiological indicators.

These assumptions are reasonable for the majority of the analysed rare cancers; nevertheless for diseases with markedly increasing incidence time trends and/or significant differences among geographic areas, some caution in result interpretation should be borne in mind. This

is the case for testicular and paratesticular cancers, rare epithelial tumours of hypopharynx and larynx, carcinoma of the thyroid gland. For cancers of testis and paratestis the estimated number of prevalent cases reported in the present monograph is about 45,000, slightly higher than the national prevalence estimates for testis previously published by AIRTUM[14] (about 38,000). As the incidence rates are higher at younger ages and the prognosis very good, the estimation of prevalent cases is mainly influenced by incidence. In Italy, incidence is lower in the South compared to the Centre-North, therefore the assumption of homogeneity in incidence may have determined an overestimation of prevalence.
We estimated about 53,000 persons to be alive at 1st January 2010 with a previous diagnosis of squamous cell carcinoma with variants of larynx (the majority of larynx cancer), slightly higher than the national prevalence estimates previously published by AIRTUM[14] (about 50,000). These differences may be associated with a difference in time trend among males and females,[28] assumed to be homogeneous for the above-mentioned reasons.
Thyroid cancer is a common cancer in Italy, with an IR higher than 6 per 100,000 per year during the analysed period, affected by a markedly increasing incidence over time in both sexes, with significant differences across areas:[28] the assumptions for rarity are therefore violated. As a consequence, the methodology developed for estimation of epidemiological indicators for rare cancers is not applicable and carcinoma of the thyroid gland was excluded from the specific commentary on endocrine tumours.

REFERENCES

1. European Parliament and Council of the European Communities. Decision no. 1295/1999/EC of the European parliament and of the council of 29 April 1999 adopting a programme of Community action on rare diseases within the framework for action in the field of public health (1999 to 2003). *Official Journal of the European Community*, Volume 42, 22 June 1999.
2. Available from: http://www.fda.gov/orphan/oda.htm
3. Greenlee RT, Goodman MT, Lynch CF, Platz CE, Havener LA, Howe HL. The occurrence of rare cancers in US adults, 1995-2004. *Public Health Rep* 2010;125(1):28-43.
4. Gatta G, van der Zwan JM, Casali PG, et al. Rare cancers are not so rare: the rare cancer burden in Europe. *Eur J Cancer* 2011;47(17):2493-2511.
5. www.rarecarenet.eu
6. Italian Network of Cancer Registries (AIRTUM); available from: http://www.registri-tumori.it
7. Fritz A, Percy C, Jack A, ShanmugaratnamK, Sobin L, Parkin DM, Whelan S (eds). *International classification of disease for Oncology*. 3rd edition. Geneva, World Health Organization, 2000.
8. http://publications.iarc.fr/Book-And-Report-Series/Who-Iarc-Classification-Of-Tumours
9. Capocaccia R, Martina L, Inghelmann R, et al. A method to estimate mortality trends when death certificates are imprecisely coded: an application to cervical cancer in Italy. *Int J Cancer* 2009;124(5):1200-1205.
10. Database AIRTUM Protocol, 2.0. Available from: http://www.registri-tumori.it/cms/files/2010.pdf
11. AIRTUM Working Group. Italian cancer figures, report 2011: Survival of cancer patients in Italy. *Epidemiol Prev* 2011;35(5-6) Suppl 3:1-200.
12. AIRTUM Working Group; CCM; AIEOP Working Group. Italian cancer figures, report 2012: Cancer in children and adolescents. *Epidemiol Prev* 2013;37(1) Suppl 1:1-225.
13. AIRTUM Working Group. Italian cancer figures, report 2013: Multiple tumours. *Epidemiol Prev* 2013;37(4-5) Suppl 1:1-152.
14. AIRTUM Working Group. Italian cancer figures, report 2014: Prevalence and cure of cancer in Italy. *Epidemiol Prev* 2014;38(6) Suppl 1:1-122.
15. Forman D, Bray F, Brewster DH, et al. *Cancer Incidence in Five Continents Vol. X*. IARC Scientific Publications No. 164. Lyon, IARC, 2014. Available from: https://www.iarc.fr/en/publications/pdfs-online/epi/sp164/CI5volX_Full.pdf
16. www.demo.istat.it
17. http://seer.cancer.gov/seerstat/software/
18. Hakulinen T, Seppä K, Lambert PC. Choosing the relative survival method for cancer survival estimation. *Eur J Cancer* 2011;47(14):2202-2210.
19. Brenner H, Gefeller O, Hakulinen T. Period analysis for 'up-to-date' cancer survival data: theory, empirical evaluation, computational realisation and applications. *Eur J Cancer* 2004;40(3):326-335.
20. Feldman AR, Kessler L, Myers MH, Naughton MD. The prevalence of cancer. Estimates based on the Connecticut Tumor Registry. *N Engl J Med* 1986;315(22):1394-1397.
21. Hakama M, Hakulinen T, Teppo L, Saxen E. Incidence, mortality or prevalence as indicators of the cancer problem. *Cancer* 1975;36(6):2227-2231.
22. Adami HO, Gunnarson T, Sparén P, Eklund G. The prevalence of cancer in Sweden 1984. *Acta Oncol* 1989;28(4):463-470.
23. Capocaccia R, De Angelis R. Estimating the completeness of prevalence based on cancer registry data. *Stat Med* 1997;16(4):425-440.
24. Corazziari I, Mariotto A, Capocaccia R. Correcting the completeness bias of observed prevalence. *Tumori* 1999;85(5):370-381.
25. Capocaccia R, Colonna M, Corazziari I, et al. Measuring cancer prevalence in Europe: the EUROPREVAL project. *Ann Oncol* 2002;13(6):831-839.
26. Merrill RM, Capocaccia R, Feuer EJ, Mariotto A. Cancer prevalence estimates based on tumor registry data in the Surveillance, Epidemiology, and End Results SEER Program. *Int J Epidemiol* 2000;29(2):197-207.
27. Mallone S, De Angelis R, van der Zwan JM, et al. Methodological aspects of estimating rare cancer prevalence in Europe: the experience of the RARECARE project. *Cancer Epidemiol* 2013;37(6):850-856.
28. http://itacan.ispo.toscana.it/italian/itacan.htm