Ollscoil na hÉireann, Corcaigh

National University of Ireland, Cork

The contribution of population-based cancer registries to the current knowledge on cancer epidemiology: the example of skin melanoma.

Thesis presented by

Emanuele Crocetti, MD

for the degree of

Doctor of Philosophy by Prior Published Work

University College Cork

School of Medicine,

Department of Public Health & Epidemiology

Head of Department

Prof. Ivan Perry

Supervisor: Prof. Kerri Clough Gorr

External Supervisor: Dr. Conan Donnelly

2018

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABBREVIATION	TERM
APC	Annual percent change
ASR	Age-standardised incidence rate
BMJ	British Medical Journal
CART	Classification and regression tree analysis
CI	Confidence intervals
CI5	Cancer Incidence in 5 Continents
CI5-X	Cancer Incidence in 5 Continents volume X
CI5-XI	Cancer Incidence in 5 Continents volume XI
CM	Cutaneous melanoma
CONCORD	Cancer survival in five continents: a worldwide population-based study
CR	Population-based cancer registry
CRP	Population-based cancer registry populations
DCI	Death Certificated Identified
DCN	Death Certificate Notified
DCO	Death Certificate Only
ENCR	European network of cancer registries
EUROCARE	European cancer registry based study on survival and care of cancer patients
EUROCOURSE	Europe against Cancer: Optimisation of the Use of Registries for Scientific Excellence in research
FSD	First Significant Digit
GP	General practitioners
GRELL	Group for cancer registration and epidemiology in Latin language countries
HPV	Human Papilloma Virus
HR	Hazard ratio
IACR	International Association of Cancer Registries
IARC	International Agency Against Cancer

ICD	International Classification of Diseases
ICDO	International Classification of Diseases for Oncology
ITT	Tumour Institute of Tuscany
JRC	Joint Research Centre
M:I	Mortality Incidence ratio
M/I	Mortality Incidence ratio
MM	Malignant Melanoma
MV	Microscopic Verification
N-B	Newcomb-Benford law
NAACR	North American Association of Cancer Registries
NM	Nodular melanoma
NORDCAN	Cancer statistics for the Nordic countries
NOS	Not otherwise specified
O&U	Other & Unspecified
QCS	Quality Check Software
r/R	Range between maximum and minimum ASR in sub-area / the ASR of the overall area
RECR	Reggio Emilia Cancer Registry
RTRT	Tuscan Regional Cancer Registry (RTRT)
RTT	Tuscany Cancer Registry
SEER	Surveillance, Epidemiology and End Results programme
SSM	Superficial spreading melanoma
UCC	Union for International Cancer Control
UK	United Kingdom
UKIACR	United Kingdom and Ireland Association of Cancer Registries
US	United Stated of America
UV	Ultraviolet radiation
WHO	World Health Organisation

DECLARATION

I, Emanuele Crocetti, confirms that this thesis is my own work and has not been submitted for another degree, either at University College Cork or elsewhere.

LIST OF PRIOR PUBLISHED PAPERS

The following six papers represent the published work which summarises my research activity in the process of cancer registries data production, diffusion and application. This is the core of the discussion in the present thesis.

- **Paper 1**. Crocetti E, Randi R. Using the Benford's law as a first step to assess the quality of the cancer registry data. Frontiers in Public Health Oct 2016, vol. 4 Article 225.
- **Paper 2**. Crocetti E, Giusti F, Martos C, Randi G, Dyba T, Bettio M. Variability of cancer risk within an area: time to complement the incidence rate. Eur J Cancer Prev. 2017;26;442-446.
- **Paper 3**. Crocetti E, Caldarella A, Chiarugi A, Nardini P, Zappa M. The thickness of melanomas has decreased in central Italy but only for thin melanomas, while thick melanomas are as thick as in the past. Melanoma Research 2010; 20:422-426.
- **Paper 4**. Crocetti E, Carli P, Miccinesi G. Melanoma incidence in central Italy will go on increasing also in the next future: a registry-based, age-period-cohort analysis. Eur J Cancer Prev 2007;16(1):50-4.
- **Paper 5**. Crocetti E, Mangone L, Lo Scocco G, Carli P. Prognosis variables and prognostic groups for malignant melanoma. The information from Cox and Classification and regression tree analysis: an Italian population-based study. Melanoma Research 2006; 15:429-33.
- **Paper 6**. Crocetti E, Caldarella A, Massi D, Sacchettini C, Amunni G, Borgognoni L. Indicators of standard of care for melanoma: Tuscany data. Melanoma Research 2013; 23: 283-9.

RESEARCH AIMS

The aim of this thesis by Prior Published Work is to highlight, in a series of six interlinked papers, the extent to which the quality (paper n.1, 2), the usefulness (n. 2, 4, 5) and the clinical relevance (n. 3, 5, 6) of population-based cancer registries' data can be enhanced.

THESIS SUMMARY

Introduction

The activity of cancer registries represents a multistep process that starts by gathering information from a variety of sources. Such information is checked, linked, enriched and handled to produce high-quality original data capable of being informative enough to prove useful in answering specific epidemiological and clinical questions.

This thesis is part of a PhD by Prior Publication grounded in six published papers. These papers deal with different steps in the production of cancer registry data, enhancing the contribution of registries to cancer epidemiology. Skin melanoma has been used as an example, but all the presented methods and concepts apply to any cancer type.

Materials and methods

1. The first paper (related to cancer registry data quality) tests the hypothesis whether the distribution of the first digit (from one to nine) of crude incidence rates obeys Benford law. Pearson's coefficient of correlation and different distance measures were applied to compare the theoretical distribution to the observed one in a sample of 43 population-based cancer registry populations randomly drawn from the volume X of Cancer Incidence in 5 Continents.

2. In the second paper, an innovative index for measuring the amount of internal variability among the sub-areas underlying an overall incidence rate is presented. The measure is a ratio, where the numerator is the difference between the highest and the lowest age-adjusted standardised rate in sub-areas. The denominator is the overall area age-adjusted standardised rate. Such measure was applied to age-standardised incidence rates for 'all cancer sites excluding non-melanoma skin cancer', for men, in 2014, for Nordic countries as a whole, for each country (Denmark, Faroe Islands, Finland, Greenland, Iceland, Sweden and Norway) and their regions.

3. In paper three, to make cancer registry data useful in the clinical setting, melanoma incidence during 1985–2004 in the Tuscan cancer registry (Italy) was analysed including both standard (site, morphology, sex, age, calendar period) and clinically relevant variables, as *in situ* melanoma and Breslow's thickness. For the time trend analysis, the annual percent change (APC) of the rates was computed.

4. Paper 4 presents the results of an age-period-cohort model applied to 1977 skin melanomas, incident in the Tuscan cancer registry. Such a method allows us to understand the time trend better and to forecast future change. Moreover, a non-linear regression model was applied to estimate the expected number of new cases in a more recent period.

5. Paper 5 shows a skin melanoma survival analysis based on 1403 patients from two Italian registries (Tuscan and Reggio-Emilia). The focus was on two different approaches: the multivariate Cox proportional hazard model and the Classification And Regression Trees analysis. The latter is an automatic method that splits data through a recursive process creating a 'tree' of groups with different profiles of risk of death. Both ways were applied to the following variables: age, sex, Breslow thickness, Clark level, Registry, sub-site and morphologic type.

6. In Paper 6, the quality of melanoma diagnosis and care in the Tuscan region is measured based on 13 newly realised process indicators, which encompassed early diagnosis, pathology reporting and surgical treatment. We evaluated the clinical adherence to these indicators in two years: 2004 and 2008, using a population-based series of incident skin melanomas, measuring the possible changes in the indexes following the implementation of specific regional recommendations.

Results

1. The distribution of the first significant digits of cancer incidence rates was shown to belong to numbers that abide by Benford law, in the whole dataset (146,590 rates) by sex and cancer registries. The correlation coefficient between observed and expected distributions was extremely high (0.999), and the distance measures very small.

2. The index for internal variability highlighted a quite relevant heterogeneity among Nordic countries (index 57.1% = the difference between the Nordic country with the highest and the one with the lowest rate is 57.1% of the Nordic overall age-adjusted rate). Within countries, the variability was negligible in Iceland (9.6%), and high in Sweden (37.1%).

3. During the four analysed periods standardized melanoma incidence rose significantly, for both invasive (APC = + 5.1%) and *in situ* lesions (APC = + 11.1). Over

time, the median value of thickness decreased from 1.68 to 0.8 mm (P < 0.001), but only for <=1.00 mm melanomas. Although the rates of thin melanoma have increased, rates for thick ones did not decrease.

4. The model that best fitted the data included age and 'drift'. The linear effect ('drift') showed, in each age group, an increase of the risk of malignant melanoma diagnosis of about 36.6% every five years of period or cohort. For the period 2002–2006, 1112 new cases were predicted with a standardised rate (age 15–84 years) of 19.2 × 100.000. In the Tuscany Cancer Registry area, no clues for malignant melanoma incidence rates levelling off were documented. Growing rates and numbers of malignant melanoma are expected soon.

5. The Cox proportional hazard model found sex, age, Breslow thickness, Clark and morphologic type to have a significant independent prognostic value. The Classification And Regression Trees analysis identified six groups of different risks based on Breslow thickness, age and sex. The best prognostic group (5-year observed survival, 98.1%) included those subjects with Breslow less than 0.94 mm and age 19–44 years. The same thickness but an older age (50–69 years) was associated with a statistically significant different prognosis (5-year observed survival, 92.8%).

6. As regards the quality of care, there were statistically significant increases in the percentage of thin (<= 1 mm) melanomas from 2004 to 2008 (from 50.7 to 61.3%) and in the number of pathology reports that mentioned ulceration (from 61.4 to 84.6%) and margin statuses (from 76.8 to 84.3%). The percentage of patients staged by sentinel lymph node biopsy was stable (63%) and was higher for patients younger than 75 years of age (74%). The number of lymph nodes almost invariably exceeded the proposed site-specific cut-off reference, and, in 2008, the number of nodes removed was always reported for lymphadenectomy. From 2004 to 2008, surgical and pathological waiting times increased.

Conclusions

The six presented papers cope cohesively with consecutive steps in the procedure of cancer registration.

1. The check for Benford law abidance may become a preliminary test in the process of data quality.

2. The heterogeneity index offers a new, simple (to be produced and to be understood) and noteworthy information.
3. The analysis of some clinically crucial variables raises the interest of clinicians and makes cancer registries closer to the real world.
4. The use of methods with higher statistical involvement, e.g., the age-period-cohort model, provided further information on melanoma time trends in the area. Moreover, estimates were projected to a more recent period bridging the Registry's timeliness gap.
5. Prognosis is a piece of vital information for both patients and clinicians. Hazard ratios and patients grouping showed almost the same risk patterns but conveyed by a different message (relative vs absolute), with a different understandability.
6. Population-based quality indexes allow to check the practical application of guidelines and recommendations, highlighting critical situations to be improved.

Cancer registration is a unique process made by different but connected steps. The improvement of each of them positively affects others. It is a sort of virtuous circle in which new methods, new uses and new users are all involved in a common aim: exploiting cancer registries' activity.
Registries can have a real informative power only if the whole process, from the collection of raw data to the provision of relevant information for various stakeholders, is accomplished.

1. INTRODUCTION

1.1. Population-based cancer registries

The usefulness of data series on patients with cancer has been appreciated for a very long time. For example, thanks to the London Bills of Mortality based on the information gathered from parishes at the time of burials, and in which the cause of death was added in 1629, we know that altogether "Cancer, Gangrene, and Fistula" caused 56 deaths in the monitored population in 1665, while “the Great Plague” of the same year had a death toll of 68,596 (Bread, 1908).
Brilliant researchers, as John Grant and, later, Sir William Petty, showed their ability to perform pioneering statistical investigations on the growing batches of information collected in the Bills (Graunt, 1662).

For a very long time, any information on cancer, as on any other diseases, was only available after patients’ death. This limit did not discourage ingenious scientists from speculating on them cleverly. An example, probably one of the first in the world, comes from Italy, where Rigoni-Stern, a medical officer reviewing the deaths caused by uterine cancer in women living in the city of Verona from 1760 to 1839 related the frequencies to personal characteristics in a causal relationship which would become clear only after the identification of the Human papillomavirus (Rigoni-Stern, 1842).

Information from death certificates represents a reliable estimate for the incidence (diagnosis) of lethal diseases, as cancer was considered until recently, but not of benign lesions. Furthermore, death certificates provide no information on diagnosis and clinical course.

Therefore, the interest in collecting information also on cancer diagnosis (incidence) and people affected by cancer (prevalence) grew, and an extensive list of pioneer attempts have been described by some of those who are among the fathers of cancer registration - Clemmenson in Denmark and Wagner in Germany (Clemmenson, 1965; Wagner, 1985; Wagner, 1991).

Their comprehensive descriptions portray many efforts in several European countries, with Germany playing a leading role. However, such attempts were, on the whole, substantially unsuccessful. The results of these laboured European trials did not prevent Americans from entering the struggle, and in 1927, in Massachusetts, they encountered the identical problems of limited participation of medical practitioners and low completeness of cancer series collection already experienced in Europe (Hoffman, 1930).

However, the process had started, and although slowly and with difficulty, precious lessons came from some of these experiences. In Germany, what we can consider a cancer registry was set up in 1926 in Hamburg. There, a group of trained nurses actively collected nominative information for each diagnosed cancer patient from both general practitioners and hospitals (Keding, 1973).

The USA chose a different strategy and performed a survey aimed at collecting cancer incidence, mortality and prevalence data to produce national estimates based on a sample of ten areas. The first survey took place in 1936-37 and again in 1947-48 and 1969-71 (Haenszel, 1975).

Despite the difficulties and the many frustrating failures, some of those attempts and experiences settled durably: e.g. the Connecticut registry in the USA set up in 1935 and the Danish national registry in 1943.

The epochal event which turned cancer registration from an activity of some odd personalities to a harmonised procedure able to produce comparable data started in 1950 when the World Health Organisation (WHO) appointed a committee to address cancer registration and data analysis.

The purpose of an international association committed to cancer registration was, in the idea of its first proponent, Sidney Cutler of the US National Cancer Institute, to improve and harmonise the quality of data on cancer incidence and comparability between registries by standardisation of methods (Whelan, 2010). The process

initiated in 1950 by the WHO, which involved all the most eminent personalities who had been active in the field, moved forward. The committee was established in 1964 by the Union for International Cancer Control (UICC) with the task of producing a technical report on cancer incidence, finally published in 1966. It was the first volume of Cancer Incidence in Five Continents (CI5) and included data from 32 Registries in 29 countries (Doll, 1966).

During the same years in France, there was an intense political initiative aimed at fighting cancer which eventually gained the support of President de Gaulle and in May 1965 resulted in the establishment of the International Agency Against Cancer (IARC), which has acted in Lyon as a technical body of the WHO in this specific field (https://www.iarc.fr/en/about/iarc-history.php) ever since.

In 1966 also, the International Association of Cancer Registries (IACR) was founded to foster the aims and activities of cancer registries worldwide (Whelan, 2010).

Therefore, the scenario changed sensibly with the involvement of the World Health Organisation and especially of its specialised cancer agency, IARC. WHO has had an active and growing role in supporting the activities of IARC, and more in general, the registries. IARC has become not only the centre of data collection and analysis for the production of the "Cancer Incidence in 5 Continents" (CI5) but also a renowned centre of expertise in cancer registration and epidemiology providing training courses, software and references (https://www.iarc.fr).

CI5 became the example to follow for all registries, a source of reliable information and a reference for data quality evaluation.

Concurrently with the gradual diffusion of registries, scientific associations were founded to coordinate and support their work. The Group for cancer registration and epidemiology in Latin language countries (Grell) (www.grell-network.com) was founded in 1975, the Association of the Nordic Cancer Registries (https://www.ancr.nu/ancr-p2/ancr/important-dates/) followed in 1984, and the

European network of cancer registries (ENCR) was then established (http://www.encr.eu) in 1989. Finally, in Europe, the United Kingdom and Ireland Association of Cancer Registries (UKIACR) (http://www.ukiacr.org/about/about-ukiacr) was founded in 2002. In 1973 in the USA, the National Cancer Institute established the Surveillance, Epidemiology and End Results (SEER) programme (https://seer.cancer.gov/), and in 1987 the North American Association of Central Cancer Registries (NAACCR) was established (https://www.naaccr.org/).

The number of registries has grown over time and, in 2017, they were overall 343, present in 65 countries over five continents (Bray, 2017).

Quoting the words of Ratnam K. Shanmugaratnam in a book which has been a landmark for generations of registrars (Shanmugaratnam, 1991), a population-based cancer registry, "is the maintenance of a file or register of all cancer cases occurring in a defined population in which the personal particulars of cancer patients and the clinical and pathological characteristics of the cancers, collected continuously and systematically from various data sources, are documented."

To know the source population makes the difference between registries and hospital-based cancer registries which collect the clinical series of a specific department, hospital, or institute (Young, 1991). The information on the population to which the patients belong makes the difference between the two types of registries glaringly obvious: only registries can compute population-based indexes (Shanmugaratnam, 1991).

The newly diagnosed cancers are the base for computing the incidence (of the disease), and once diagnosed a person contributes to prevalence until death because the definition of a cured cancer patient is still under debate (Essig, 2019).
The definition of cancer cure has been proposed differently by oncologists and epidemiologists since it has implemented for individual patients or populations of patients. Furthermore, some definitions apply for specific cancer types, as the complete biochemical remission for thyroid cancers (Pacini, 2012). Moreover, for

some other cancers after initial treatment patients may also remain relapse-free, or without any measurable sign of disease, for several years (e.g., breast cancer patients) with a small long-term excess risk of relapse or death (Janssen-Heijnen, 2014; Mariotto, 2018).

An epidemiological definition is to consider cured those patients who have reached the same death rates of the general population (Dal Maso, 2014). However, although these patients may reach the same risk of death than the general population (at the time-to-cure) and die from other causes than cancer ('cured' from a statistical point of view) (Romain, 2019), some of them may die with cancer (Dal Maso, 2014).

Registries produce incidence, prevalence and survival for any cancer type, and at least in theory for any other collected variable (e.g., sex, age, morphology, as well as stage at diagnosis, sub-groups of patients defined by the combination of biomarkers).

The population represents the denominator of all the measures of frequency and severity produced by registries, and such measures provide, at least theoretically, unbiased estimates of the average values in the whole population. The continuous activity of registries also allows to measure trends of the epidemiological indicators over time.

The information input to registries is clinical and demographic data routinely produced for clinical and administrative needs (e.g., hospital discharge notes, death certificates, pathology reports) (Powell, 1991).

Over time, digitalisation has made it possible to retrieve information from many other sources, both clinical (e.g., laboratory, hospice, radiotherapy, screening, drug prescriptions) and administrative (e.g., insurance records, tax exceptions).

Each source of information is useful for two purposes: to confirm the information coming from other sources and to provide original contribution. Therefore, increasing the number and type of sources improves the completeness of the registry (Parkin & Bray, 2009).

Any source is linked with all the others belonging to the same person, using the most reliable personal identification variable available, then different tumours diagnosed in the same person have to be identified and coded separately if they are independent of each other.

For each case (patient and cancer) a minimum set of information to be collected has been identified in technical manuals, with primary and desirable variables (MacLennan, 1991; Menck, 1994; Hutchinson, 1997). However, information availability has varied over time and from country to country due to technical, legal, and economic accessibility.

Although registries generally include all invasive cancers, some of them avoid collecting data on non-melanoma skin cancer (especially basal epitheliomas) (Ferlay, 2014) while others include non-malignant tumours which, for example, are particularly relevant during childhood and are indeed included in the International Classification of Childhood Cancer (Steliarova-Foucher, 2005). Moreover, some registries collect information about *in situ* tumours for those cancers for which organised, or spontaneous screening programmes are active (Bray, 2014). Finally, the inclusion of cancers diagnosed at the autopsy is not homogeneous among countries and registries (Bray, 2014).

Irrespective of these differences in inclusion criteria, which may distinguish one registry from another but are usually minor (Bray, 2014), CRs strategy is to look for the footprints left in the healthcare system by cancer patients during their clinical experience. This information is collected and converted into epidemiological indexes at the end of the registration process. Theoretically, registries aspire to 100% completeness which means identification of all the cancers diagnosed in a defined population. To reach this goal, they have to use as many sources of information as possible. However, clinical paths may be similar, but they are nonidentical for all cancer types, all cancers of the same type (e.g. for differences in the stage of the disease at diagnosis which usually require different treatments) and all patients (e.g.

differences in age at diagnosis, which may cause different intensity of the diagnostic and therapeutic process due to coexistence of comorbidity) (Powell, 1991).
Life has incredibly changed over time, with a succession of technological revolutions. Computers have replaced sheets of papers and calculators, and the web has made traditional ways of communications obsolete and mobile phones have become such smart to be used as health information devices (Min, 2014). In the meantime, while technology has made giant leaps, registries have kept their traditional pace, causing frustrations for the delay in data production, the scarcity of information provided, the scarce use of the data by registries, the weak interest in these data shown by policymakers and even more by clinicians. Moreover, for several years, legislation on data protection hampered registry activity very significantly and only recently clear and standard rules been introduced (EU, 2016).

Registries may be considered systems which monitor cancer at the population level, able to identify risk groups, document changes over time, predict future disease burden, show improvements, but also identify emerging groups at risk and differences in quality of care and unmet needs.

CRs must produce reliable information to carry out all these functions properly. Registrars invest a lot of time and resources for providing valid, comparable and complete data and consequently trustworthy epidemiological measures.

Once the high quality of a dataset is proven, epidemiological parameters may be computed. These may be raw numbers, crude-, specific- and age-adjusted rates, cumulative rates, proportions for incidence and prevalence and different types of survival rates.

Since the beginning of my experience, the expectation of the scientific community from registries was limited, and their fate has seemed doomed: slow but sure death. On the contrary, I had the impression registries were only beginning, and it was up to researchers to exploit their potential.

I accepted this challenge, and I herein present my research, hoping it can help registries to be more actively involved in healthcare systems.

2. Quality of cancer registry data

2.1. Quality of cancer registry data - Introduction

Quality evaluation of registry data includes checking several dimensions: validity (the accuracy of the registered data: the correspondence between what should be measured and what is actually measured); completeness of cancer data collection; timeliness (interval between clinical diagnosis and availability of epidemiological estimates); and comparability (in the same registry over time, and among different registries, which implies the use of standard rules and procedures) (Bray&Parkin, 2009; Parkin&Bray, 2009).

The quality of collected data has been considered a pivotal issue throughout the history of registration. The low completeness achieved weakened the first pioneer attempts at cancer registration (Wood, 1930). Only more recently, data from registries were favourably judged by Sir Richard Doll who wrote: “In many instances, there is now reason to believe that the figures that have been obtained in this way (routine cancer registration) correspond very closely to the true incidence of disease" (Doll, 1972).

The evaluation of the quality of registry data, assessable at the level of each single registry (Ryzhov, 2018), has improved since the implementation of ‘Cancer Incidence in 5 Continents’ that compared data from different registries and developed a section of competence on cancer registration at the IARC.

The evaluation of the quality of registry data, crucial in this comparative publication, produced a series of rules and methods which have spread widely and been increasingly applied, contributing to the establishment of harmonised standards and procedures (Bray, 2014; Bray&Parkin, 2009; Parkin&Bray, 2009).

Recently, a collaborative project coordinated by the ENCR together with the Joint Research Centre (JRC) of the European Commission and with the involvement of the leading European consortia and stakeholders in cancer registration (e.g. IARC, Eurocare, Concord) adopted a list of standard variables to be collected to fulfil various projects (e.g., incidence and survival computation) and corresponding data quality checks (Martos, 2014). Moreover, data checking software includes these agreed rules (JRC-ENCR Quality Check Software - QCS, https://www.encr.eu/download).

As briefly mentioned above, evaluating the quality of registry data remains a complex process requiring the assessment of diverse aspects: validity, comparability, completeness and timeliness.

2.1.1. Quality of cancer registry data - Validity

Validity is the correspondence between what is registered in a registry and the actual value of the feature to be registered (Kearney, 2015). This parameter is evaluated within the dataset measuring the internal correspondence (consistency) among different variables (e.g., sex and cancer sites, or age and cancer sites). The availability of specific software, has simplified this process considerably (e.g., JRC-ENCR QCS, https://www.encr.eu/tools-for-registries; IARCcrgTools, http://www.iarc.com.fr/index.php?option=com_content&view=article&id=72:iarccrgtools&catid=68&Itemid=445).

Moreover, the reliability of information provided by a registry depends on the reliability of the source of such information. Therefore, the higher the proportion of cases confirmed by microscopic verification (MV%, histology or cytology or peripheral blood cells examination), the higher the validity. However, an extremely high MV% (close to 100%) may be the clue for an excessive reliance of registry from such information source with insufficient attention to others, which may result in a selection of a sub-group of cases (Parkin, 1994).

Conversely, the higher the proportion of cases known from the death certificate only (DCO), the lower the validity. However, death certificates are one of the traditional

sources of information and all registries which have access to such data employ it to improve completeness and update follow-up information (Parkin, 1994).

Ordinarily, any death certificate notified case (DCN) should be coded as DCI (cases identified by death certificate). After that, a traceback procedure is performed to detect any missing information in all the other available sources. If traceback is successful, the case is registered with the supplementary information (e.g., date of incidence, source of information). Only those cases for which traceback fails to pinpoint another appropriate date of incidence and source of information become DCO.

Therefore, in case a registry does not register DCI but only DCO, a low proportion of DCO may not be very significant, because we do not know how many of these cases the registry would have initially acknowledged as DCI (Parkin, 1994).

Among the measures of validity, there is also the proportion of missing values for significant variables, usually age at diagnosis and site of the tumours (coded in the group Other & Unspecified cancer site, O&U). A considerable proportion of O&U or DCO may be the indicator of poor cancer control in a jurisdiction, or of limited precision in cancer registration.

2.1.2. Quality of cancer registry data - Comparability

Comparability implies the use by different registries of the same language and grammar, represented by the adoption of the same classifications for defining cancer, and the same definitions for those aspects of cancer registration which may have a substantial effect on the estimate of incidence, like the definition of 'date of diagnosis' or the discrimination between recurrence and second independent cancer diagnosis.

The classifications used for defining cancer have changed and developed over time (WHO, International Classification of Diseases, ICD, http://www.who.int/classifications/icd/en/). Moreover, a specific classification for cancer, the International Classification of Disease for Oncology (ICD-O) has been

introduced and continuously updated (ICD-O 1,2,3, http://codes.iarc.fr/) according to the updates of the ICD to which it is related. The ICDO-3 represents the state-of-the-art for registries and enables them to encode data on cancer site, morphology, behaviour and grade of tumour differentiation (Ferlay, 2014).

The trustworthiness of registry data depends on that of the sources of information. Therefore, such information is necessary and has to be classified and coded, according to IARC (http://www.iacr.com.fr/images/doc/basis.pdf), or ENCR (http://www.encr.eu/images/docs/recommendations/basisd.pdf), or SEER (https://seer.cancer.gov/manuals/2016/SPCSM_2016_maindoc.pdf) in ways which make it possible to find a minimum common denominator.

The rules for the dates of incidence, the occurrence of multiple primaries and cancers detected in asymptomatic people may all affect incidence.

2.1.2.1. Quality of cancer registry data – Comparability – Date of incidence

The availability and application of innovative technology, e.g., Prostate-Specific Testing and subsequent biopsy for prostate cancer, may sensibly anticipate the date of cancer diagnosis (Schroder, 1995).

The Registries collect and code the basis of diagnosis. Some classifications have proposed updating the rules for date of diagnosis definition according to the development of technology. On the other hand, publications like CI5, which compare countries with radically different levels of technology development, apply classifications stable overtime to assure comparability. This reason may have contributed to the adoption of slightly different rules for IARC (Jensen, 1991), and the USA (Adamo, 2018).

In Table 1 the summary recommendations for coding incidence date in SEER (for solid tumours) (Adamo, 2018) and IARC (Jensen, 1991) are compared. Both systems request to choose the date with the highest priority. A broader use of the laboratory testing and/or imaging without biopsy or surgical removal of the lesions or hospitalisation may

increase the completeness and the timeliness with SEER rules for some types of cancer patients (e.g., extremely elderly patients or cases diagnosed in advanced stage).

Table 1: Comparison of the summary recommendations for coding incidence date in SEER and ENCR. The priority decreases from top to bottom.

Priority	SEER	IARC
1 (higher)	Positive histology	Consultation at, or admission to, a hospital, clinic or institution for the cancer in question
2	Positive cytology	First diagnosis by a physician or first pathology report
3	Positive microscopic confirmation (MV), method not specified	Death (when the cancer is first ascertained from the death certificate and follow-back attempts have been unsuccessful)
4	Positive laboratory test / marker study	Death preceding an autopsy, when cancer is first found and was unsuspected clinically
5	Direct visualisation without MV	
6	Radiology and other techniques without MV	
7	Clinical diagnosis only	
8 (lower)	Unknown whether or not MV; death certificate only	

2.1.2.2. Quality of cancer registry data – Comparability – Multiple primaries

The rules for defining when multiple primary cancers diagnosed in the same person may be counted in incidence as independent cases may vary from registry to registry. Also, in this field, the effect of adopting various rules may produce not a negligible variability in incidence rates. The main difference is between the rules of the IACR-

IARC-ENCR (http://www.ia.com.fr/images/doc/MPrules_july2004.pdf) and those adopted by the SEER (Johnson, 2007). The former considers only one tumour for a specific topography (or groups of topography) including bilateral organs unless the morphologies of two or more cancers are different enough to belong to distinct groupings (Berg, 1996). The US rules are more extensive and allow the inclusion of more tumours in bilateral organs (e.g., breast) and in sub-sites of some other tissues (e.g., colon). The application of these two types of rules may cause considerable differences in age-adjusted incidence rates for specific cancers sites (from +2 to +8% for the colon, lung, skin melanoma, testis, female breast, kidney, and from +1% to +3% for all sites except non-melanoma skin cancers) (Bray, 2014).

2.1.2.3. Quality of cancer registry data – Comparability – Screening and autopsy

The last group of conditions includes screening (Morrison, 1985) and the autopsy that both may affect incidence for cancers diagnosed incidentally in people without symptoms or symptomatic but undiagnosed before death.

2.1.2.3.1. Quality of cancer registry data – Comparability – Screening

The effect of screening is to increase both incidence for earlier diagnosis and the detection of some cancers which would not have come to clinical evidence (Morrison, 1985). The latter point is usually called over-diagnosis, "the diagnosis of a "cancer" that would otherwise not go on to cause symptoms or death" (Welch, 2010). Therefore, different cancer screening modes - either organised (mammography for female breast, Pap test or HPV testing for cervix uteri, and faecal blood detection or sigmoidoscopy for colorectal cancers) or spontaneously performed (e.g., for the thyroid, melanoma, lung, or prostate cancer) - may have to be considered when comparing registries. Screening tends to anticipate the time of diagnosis, causing an increase in incidence during the first rounds of its implementation (Bray, 2009).
Moreover, when screening is addressed to detect premalignant lesions (e.g., for the cervix and colorectal cancer), this may cause a decrease in incidence (Bray, 2009). A similar result may be due to variation in the intensity of clinical investigation among jurisdictions (Carsin, 2010).

2.1.2.3.2. Quality of cancer registry data – Comparability – Autopsy

The frequency of post-mortem examination may vary from country to country (Start, 1995; Burton, 2003). In registries which include cases incidentally detected at autopsy - without any suspicion of malignancy before death (Bray, 2009) - incidence rates may inflate when this practice is frequent (Bieri, 2015) or marginally if it is more uncommon.

2.1.3. Quality of cancer registry data – Completeness

Completeness refers to the number of cases diagnosed in the target population in a specified year, included in the registry's dataset. The capability of a registry to express the true incidence in the target population increases with completeness.

A comprehensive set of parameters may be used to evaluate completeness:

2.1.3.1. Quality of cancer registry data – Completeness - Steadiness of rates

Within a registry, the case-series is expected not to change abruptly from one year to the next. Incompleteness may be higher (and the number of cases lower) for data collected in more recent years (Parkin, 1994).

2.1.3.2. Quality of cancer registry data – Completeness - Homogeneity in the area

Cancer incidence is not expected to change from one area to a neighbouring one, unless there are local exposures (Bruno 2014) or actions (e.g., screening) (Waldmann 2012) capable of modifying it. Therefore, a reference for completeness may be the comparison of incidence levels in adjacent registries (Parkin, 1994).

2.1.3.3. Quality of cancer registry data – Completeness - Steadiness in age-specific curves

Sudden drops in age-specific rates for specific cancer sites may be a clue for incompleteness in data collection or a problem in the denominator in specific age-groups (Parkin, 1994).

2.1.3.4. Quality of cancer registry data – Completeness - Number of sources of notification per case

More sources make the information more reliable (if the information is concordant among them). The same is true for the number of independent notifications for each source (which reduces the risk of grounding the decision on the same possible coding error) (Parkin, 1994). This method implies that linkage among sources and notifications avoids duplications and errors (Brenner, 1996).

2.1.3.5. Quality of cancer registry data – Completeness – Childhood cancer incidence

Childhood cancer incidence was traditionally considered quite stable over time and in neighbouring registries. However, recent analysis has documented geographical and temporal variations which call for a reconsideration of this index as a measure of completeness (Steliarova-Foucher, 2017). Several reasons may cause differences in rates between areas. For example, a different prevalence of genetic (Walsh, 2013) or environmental risk factors (Gupta, 2012; Ward, 2009; Hernandez, 2016). Moreover, an important role may be performed by differences in registration techniques and the availability and access to data sources (Steliarova, 2017). Besides, incidence registration and consequently incidence rates may be affected by the availability of diagnostic facilities (Amayiri, 2014; Hadley, 2012) and once available, by socioeconomic barriers to their access (Magrath, 2013, Steliarova, 2017).

2.1.3.6. Quality of cancer registry data – Completeness - Proportion of microscopic verification

This parameter, already described for validity (section 2.1.1.), is equally useful to evaluate completeness comparing the MV%, overall and for specific cancer sites, overtime for the same registry and with neighbouring registries.

2.1.3.7. Quality of cancer registry data – Completeness - The mortality to incidence ratio

The mortality to incidence ratio (M/I) represents the comparison of the number of registered cancer deaths with cancers cases, for all cancer's sites together and for specific cancer sites. It depends on the validity of the definition of the cause of death.

The ratio should be reasonably stable when compared with neighbouring registries in areas with similar organisation of the health care system (e.g., availability of screening programmes, diffusion of campaigns for early diagnosis, etc.) and values above the expected ones are suggestive of incompleteness.

2.1.3.8. Quality of cancer registry data – Completeness - The death certificates method

Cases identified from death certificate (DCO) have been already mentioned for their role in the validity evaluation (section 2.1.1.). The same index may be used to ascertain completeness: the proportion of DCO cases evaluated together with the proportion of cases identified by the death certificate (DCI) in a situation in which death certificates are of good quality may suggest incompleteness.

2.1.3.9. Quality of cancer registry data – Completeness - The DCI and M/I method

This method entails knowing the DCI status and mortality to incidence ratio (M/I) to estimate the fatality rates of originally unregistered cases (Parkin, 1994). Ajiki has also proposed an original formula for such a method (Ajiki, 1998).

2.1.3.10. Quality of cancer registry data – Completeness - Independent case ascertainment method

This method compares the result of the registration procedure with an independent cancer case-series. The missing data proportion is an estimate of completeness (Parkin, 1994). Several different independent sources have been utilized: e.g., in Denmark incompleteness of around 2% was detected by linking the registry with different datasets, e.g., 5,674 Danish invasive cervical cancer patients enrolled in an international clinical follow-up study (Storm, 1988), or Northern Ireland General Practitioners (only 15 out of 17.102 cancer patients were not known by the Northern Ireland Cancer Registry, equal to a 99.9% completeness) (Kearney, 2015).

2.1.3.11. Quality of cancer registry data – Completeness - Capture-recapture method

The method uses the same approach applied initially to estimate the size of wild animal populations. It requires at least two phases: one in which those captured are identified (branded) and then released, and another in which a new catch based on

the numbers of those already known and the new ones constitute the basis for estimates (Parkin, 1994). Over time the method has been employed by registries and has been improved to take into account of the sources of information used by registries are not independent (Brenner, 1995; Crocetti, 2001).

2.1.3.12. Quality of cancer registry data – Completeness - Flow-method

This method (Bullard 2000) estimates the proportions of unregistered patients from the time distributions of three probabilities which can be computed using registry data: the probability of survival, probability of registration of cancer during the patient's life, and probability for cancer to be mentioned on the death certificate for cancer patients who die (Parkin 1994). This method implies a continuous and timely flow of the source information with no delay from their clinical production. Moreover, the improvement in the availability of source information may also improve timeliness of registration (Donnelly, 2017).

2.1.4. Quality of cancer registry data –Timeliness

The process of cancer registration, briefly described before, implies many consecutive phases (e.g., data collection, linkage, patient definition, case/s definition, exclusion of prevalent cases, exclusion of duplicates, data coding, quality evaluation, error amendment). Only after all these steps is it possible to perform data analysis, interpretation and publish results. Moreover, the original data sources may become available for some registries only after some time from their production (weeks, months, one or even more years) (Zanetti, 2015).

Therefore, a certain amount of delay depends on the registration procedure and may vary from registry to registry according to the local organisation and the resources available. Moreover, another part is independent of the registries (being related to the availability of sources) and may vary as well among registries.

Timeliness is related to completeness and to the fact that data are publicly released when they are considered complete (Zanetti, 2015). Therefore, a certain amount of reporting delay is inevitable (Donnelly, 2017). The open issue is how to reduce this

delay and how varies among registries. Also, during the production of incidence data, for a certain number of years, some new cases incident in the previous years are still discovered. This phenomenon is as unavoidable as a certain degree of reporting delay. In Northern Ireland, the proportion of cases reported after the first year of incidence varied between 1% and 13% for different cancer sites (Donnelly, 2017). The SEER programme provides estimates which consider the reporting delay to avoid the related bias, especially in more recent incident years (Lewis, 2015; Huang, 2013; Clegg, 2002).

2.1.5. Quality of cancer registry data – Population

The quality of incidence rates does not depend only on the quality of the numerator of rates (number of cases). Moreover, it also depends on the quality - mainly the completeness - of the denominator (population). Usually, registries are uninvolved in the estimation of resident populations. They receive such data from the official Institutions in the area which enumerate the resident people based on census or ad hoc estimates.

Therefore, CRs have to check also the completeness of population data, including consistency among ages and sexes, in comparison with previous years.

2.1.6. Quality of cancer registry data – Survival

Until now, the description of registry activity, data handling and quality evaluation has focused on incidence. However, the date of incidence is also the starting point for measuring the time of survival. Therefore, the checks for incidence have to be performed before any survival studies.

Further, for survival there are other evaluations specifically related to the internal consistency between survival-related variables: e.g. vital status/autopsy, autopsy/basis of diagnosis, and autopsy/survival/dates of incidence and follow-up (Martos, 2014). The endpoint of survival is the time of follow-up at which the life status of any patient has to be defined (dead, alive, or lost).

What affects the date of incidence also influences survival, lengthening or shortening it according to different definitions. Possible registration flaws (Wilkinson, 2009) have been claimed to explain the reported lower survival rates in the UK than in other European Countries (Berrino, 2007). In particular, Beral and Peto, in their editorial in BMJ, attributed such results to a considerable proportion of cancer cases initially known in UK from death certificates for which the traceback failed to find information on the actual date of diagnosis confusing it with the date of recurrence, that is clearly closer to date of death (Beral, 2010).

Moreover, a selective incompleteness in the collection of some long-term survival patients has been proposed as another possible explanation for the supposed inferiority of the UK healthcare system (Beral, 2010). However, a simulation performed within the National Cancer Registry of England and Wales, supported the truthfulness of low survival in the UK, estimating that errors would have been implausibly extensive to produce the evidenced differences between UK and Sweden (Woods, 2011).

Furthermore, specific checks are necessary to evaluate the quality of the status of life definition, and they are dependent on the availability of certificates of death and the type (active or passive) of the search for the information on the status of life. Subgroups of people may request specific evaluation, e.g., cases known from death certificate only, cases lost to follow-up, patients recorded as alive and aged over 100 years or long-term survivors affected by lethal cancers (Rossi, 2015; Allemani, 2017).

The following Paper (1) proposes a new application of a mathematical law for the evaluation of the quality of cancer registry data.

2.2. Paper 1

The paper has been published in Front Public Health. 2016; 4: 225 with the following title:

"Using the Benford's Law as a First Step to Assess the Quality of the Cancer Registry Data".

Authors: Emanuele Crocetti, Giorgia Randi
Affiliation: European Commission, DG Joint Research Centre (JRC), Institute for Health and Consumer Protection, Public Health Policy Support Unit, Ispra (VA), Italy
Author Contributions: EC conceived the idea of the study, planned and designed it, and prepared the first draft. GR made substantial contributions to the statistical analysis and critically revised the paper. Both authors edited and approved the final version of the manuscript. Both authors are accountable for all aspects of the work.

Abstract

Background: Benford's law states that the distribution of the first digit different from 0 [first significant digit (FSD)] in many collections of numbers is not uniform. The aim of this study is to evaluate whether population-based cancer incidence rates follow Benford's law, and if this can be used in their data quality check process.

Methods: We sampled 43 population-based cancer registry populations (CRPs) from the Cancer Incidence in 5 Continents-volume X (CI5-X). The distribution of cancer incidence rate FSD was evaluated overall, by sex, and by CRP. Several statistics, including Pearson's coefficient of correlation and distance measures, were applied to check the adherence to Benford's law.

Results: In the whole dataset (146,590 incidence rates) and for each sex (70,722 male and 75,868 female incidence rates), the FSD distributions were Benford-like. The correlation coefficient between observed and expected FSD distributions was extremely high (0.999), and the distance measures low. Considering single CRP (from 933 to 7,222 incidence rates), the results were in agreement with Benford's law, and only a few CRPs showed possible discrepancies from it.

Conclusion: This study demonstrated for the first time that cancer incidence rates follow Benford's law. This characteristic can be used as a new, simple, and objective tool in data quality evaluation. The data analyzed had been already checked for publication in CI5-X. Therefore, their quality was expected to be good. In fact, only for a few CRPs several statistics were consistent with possible violations.

Introduction

The Benford's law (Benford, 1938), originally identified by Newcomb (Newcomb, 1882), states that in many numerical series the distribution of the first significant digits (FSDs) (the first non-zero digit on the left side of a number) is not uniform. In fact, for numbers which adhere to this law, the probability of 1 to be the FSD is 30.1%, and this probability steadily decreases for the following digits up to 9, which is the least common leading digit (4.6% of the cases). A distribution abides by Benford's law if the frequency [F(x)] of the FSD, $x \in \{1, \ldots, 9\}$, follows the logarithmic relation, $F(x) = \log_{10} 1 + 1/x$ (Benford, 1938). The law of "anomalous numbers" applies also to the frequency of digits in other positions (Benford, 1938).

Not all the numbers abide by Benford's law, but for those which do, violations raise concerns. For example, in accounting and auditing, also at governmental level, Benford's law has been widely used to detect possible frauds (Tödter, 2009; Rauch, 2011; Slijepcevic, 2014).

Population-based cancer registries produce a great amount of numbers (the cancer incidence rates). The evaluation of their quality is rather complex, involving different aspects, and it is mainly based on the knowledge of the clinical, diagnostic, and therapeutic pathways of patients and on the process of data collection and registration (Bray&Parkin 2009, Parkin&Bray 2009).

The most renowned publication on cancer incidence is Cancer Incidence in 5 Continents (CI5) (Forman 2014). The cancer registries submitting their data to CI5 have to pass a formal quality evaluation before being accepted. The data quality assessment implies checking the internal coherence, consistency, completeness, and comparability with the final decision taken by a group of experts in the field.

The aim of this study is to evaluate if cancer incidence rates adhere to Benford's law to use this mathematical characteristic as a further objective tool for their quality evaluation.

Materials and methods

The website of the CI5 volume X (CI5-X) (Forman, 2013) provides the data of the 290 population-based cancer registries included in the publication, detailed by all the 424 cancer registry populations (CRPs), as each cancer registry can provide information not only for the whole population but also for different racial and/or ethnic subgroups within the same population.

The CI5-X data include aggregated information for 244 combinations of cancer site and morphological group, specified for 19 age groups (5-year age groups from 0–4 to 85+, plus unknown age) and for the two sexes.

We drew a pseudorandom sample of 10% of the available CRPs, stratified by continent (considering South and North America separately), setting a random number seed to make the sampling reproducible.

Overall, 43 CRPs (from 40 cancer registries) were sampled and included in the analysis: 1 from Africa (Malawi, Blantyre), 3 from Central and South America (Argentina, Tierra del Fuego; Brazil, San Paolo and Ecuador, Quito), 18 from USA (Virginia, Asian and Pacific Islanders; Nebraska, Black; Ohio; Vermont; Montana; Michigan; Georgia; Indiana, White; Missouri, White; NPCR-National program of cancer registries – including 42 States; Colorado, Asian and Pacific Islanders; Arkansas, Black; Alabama, White; Arkansas, White; California, Asian and Pacific Islanders; Connecticut, Black; Virginia, Black; and California), 7 from Asia (India, Karunagappally; Singapore, Malay; Turkey, Edirne; Israel, Jews; Japan, Hiroshima Prefecture; Japan, Fukui Prefecture; and Israel), 11 from Europe (France, Isère; Germany, North Rhine – Westphalia; France, Hérault; UK, England; Estonia; Switzerland, St Gall-Appenzell; Bulgaria; Malta; Ukraine; Spain, Navarra; and Italy, Sondrio), and finally 2 from Oceania (New Zealand; Other and USA, and Hawaii).

The cancer data corresponding to age group 19 (age unknown) were excluded from the analysis.

After the exclusion of those combinations of cancer morphology and site with no cases, 146,590 combinations were included in the analysis.

Crude incidence rates were computed for each sex, age group, and topography and morphology combination dividing the number of cases by the corresponding population and expressed per 100,000 inhabitants.

The FSD distribution for crude incidence rates was then calculated for all the CRPs together, by sex, and by CRP. Moreover, a sensitivity analysis has been performed randomly excluding half of the most important cancer sites (prostate, lung, breast, and colon–rectum).

For checking the adherence of observed FSD distributions to Benford's one, we used different methods.

Since Benford's distribution has mean greater than median and is positively skewed (Wallace, 2002), these figures have been evaluated for cancer incidence rates.

Theoretical and observed distributions were plotted for graphical comparison.

Following literature, we did not use those tests (e.g., χ2 and Kuiper's statistic) that are extremely sensitive in rejecting the null hypothesis (being a distribution Benford-like) for large samples (Rauch, 2011; Gollbeck, 2015; Judge, 2009; Kienle, 2015). To test the goodness of fit, we used the following tests:

r: the Pearson correlation. This is commonly used to measure how closely a distribution follows Benford's law (Gollbeck, 2015; Judge, 2009). The closer the coefficient "r" is to +1 the higher the correlation between Benford's law and the observed FSD distribution.

- χ^2-/n: χ^2 divided by the sample size (Rauch, 2011; Leemis, 2000).

– m: the maximum distance in absolute terms between expected and observed frequencies for each of the nine digits (1–9). The statistics may vary between 0 (no differences between the two distributions) to + (maximum difference) and the corresponding formula is $m = max_i = 1, 2, \ldots, 9\{|b_i-e_i|\}$ (Judge 2009), where b_i is the frequency expected by Benford and e_i is the observed frequency for each digit i.

- d*: the normalised Euclidean distance between the two distributions divided by the maximum possible distance, which would occur when the FSD was 9 for all the numbers. The corresponding formula is:

- $d^* = \sqrt{\sum_{i=1}^{9}(b_i - e_i)^2} / \sqrt{\sum_{i=1}^{8}(b_i)^2 + (1 - e_9)^2}$

where b_i is the frequency expected by Benford and e_i is the observed frequency for each digit i. The statistic may vary between 0 (no differences) to 1 (maximum difference) (Judge 2009).

Z statistic: the average of the Z values for each comparison between the nine observed and theoretical digits distributions (Slijepcevic, 2014):

$$Z = \frac{1}{9}\sum_{i=1}^{9}\sqrt{n}\left[\frac{|b_i - e_i| - \frac{1}{2n}}{\sqrt{b_i}(1 - b_i)}\right]$$

where i = 1, . . ., 9 is a fixed digit, b_i is the frequency expected by Benford, and e_i is the observed frequency for each digit i. The cut-off value for statistical significance, with alpha = 0.05 and one side tail, is 1.64.

For providing an inter-CRP comparison, the mean, the median, and the 10th or the 90th (the one including the most extreme values) percentile of each statistic were computed.

A summation index has been computed for rating the CRPs according to the statistics' results. Each CRP received one point for each statistic in the 10th or 90th percentile (whichever represents the worst values). The summation index could vary from 0 (no statistic beyond the threshold) up to 5 (all statistics beyond the threshold). The probability for each statistic to be in the most extreme decile was 0.1 (approximately 4/43) assuming independence between statistics, considering that the summation index follows a binomial distribution (pr = 0.1, n = 5) the random probability for a CRP to have the summation index equal to 0 is 0.59, to 1 is 0.33, to 2 is 0.07, to 3 is 0.008, to 4 is 0.0005, and to 5 is 0.00001. The analysis has been performed with Stata v. 12,

using specific commands for extracting the sample ("sample" and "seed") and for computing observed and Benford FSD distributions ("digdis").

Results

When considering all cancer incidence rates together (146,590 observations), the distribution of the FSDs appeared to be positively skewed (0.84), with the mean (3.38) greater than the median (3.0). These values were close to those of the theoretical Benford's distribution (skewness 0.8, mean 3.44, and median 3.0), as were the ratios between 1st vs. 9th (observed 6.6 vs. Benford 6.6), and between 1st vs. 2nd (1.8 vs. 1.7) FSD.

These results suggest that the FSD distribution of cancer incidence rates might adhere to Benford's pattern. In fact, when the observed FSD distribution was graphically compared to the theoretical one, as shown in Figure 1, they were almost overlapping.

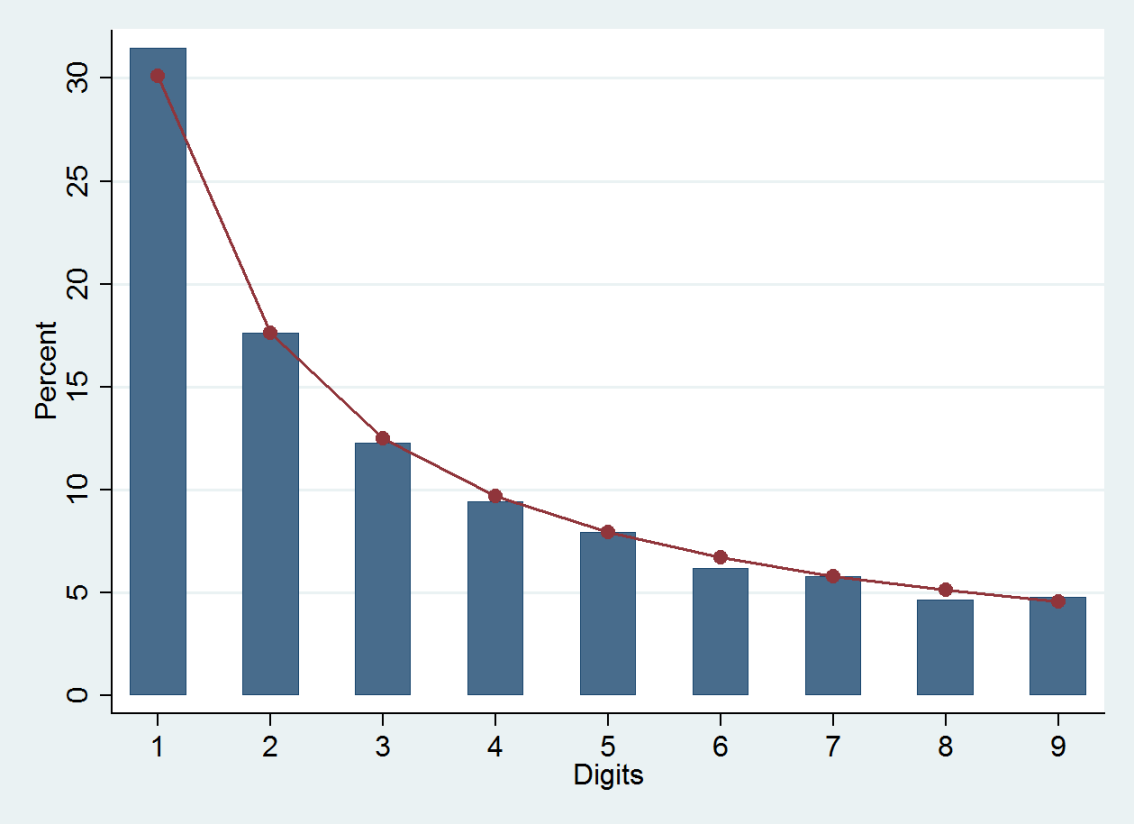


Figure 1: Benford (line) and observed (columns) distributions of first digits for all crude cancer incidence rates.

Pearson's correlation coefficient, r, showed an almost perfect direct correlation between the observed FSD distribution and the expected one (0.999); moreover, all the measures of the distance between the distributions were very low (m = 0.014 and d* = 0.015), and the average Z was below the significance level. Finally, the $\chi 2$ test, weighted on the number of observations ($\chi 2/n$), was also very low (0.002).

The analysis has been repeated by sex and the same results were confirmed (results added in Appendix 2 [Supplementary information for Paper 1]; data not shown in the original paper). Also, after the exclusion of half of the rates for the major cancer sites the overall results confirmed the adherence of the FSD distribution to Benford's law (r = 0.999, m = 0.014, d* = 0.016, $\chi 2/n$ = 0.002).

When single CRPs were evaluated, each FSD distribution was positively skewed, and the mean was greater than the median (results added in Appendix 2 [Supplementary information for Paper 1]; data not shown in the original paper).

In Figure 2, the FSD distribution of all cancer incidence rates and the Benford distribution were compared for each of the 43 analysed CRPs. The shapes of all distributions generally resembled Benford's, with a decreasing percentage of FSD from 1 to the 9. However, a few possible differences were shown.

The Pearson correlation coefficients were very high for the majority of the CRPs (median = 0.97); however, some values were relatively low (0.85 representing the 10th percentile). Also, the other measures of distance were generally low (median: m = 0.05, d* = 0.07), but still the corresponding 90th percentiles reached rather higher values (0.10 and 0.12, for m and d*, respectively). For the ratio between the $\chi 2$ and the number of rates, the 90th percentile was almost 3 times the median (90th percentile = 0.14 and 50th percentile = 0.05), and, finally, for the average Z, the value of the 90th percentile corresponded to the value of statistical significance (1.64).
Although the majority of the CRPs reported statistics showing an agreement with Benford's law, for a few of them the values seemed to indicate possible discrepancies.

For 35 CRPs, the summation index was 0, for 4 CRPs was 1, and for 2 CRPs was 2. Only one CRP (Argentina, Tierra del Fuego) reported a summation index of 3 (r = 0.839; d* = 0.125; χ/n = 0.146) and another one (USA, Virginia, Black) had all the five statistics in the worst classes (r = 0.82; m = 0.13; d* = 0.147; χ/n = 0.182; Z = 1.88). The probability for the two latter results to happen by chance is very low. Therefore, for such CRPs, a possible violation of Benford's law should be *considered.

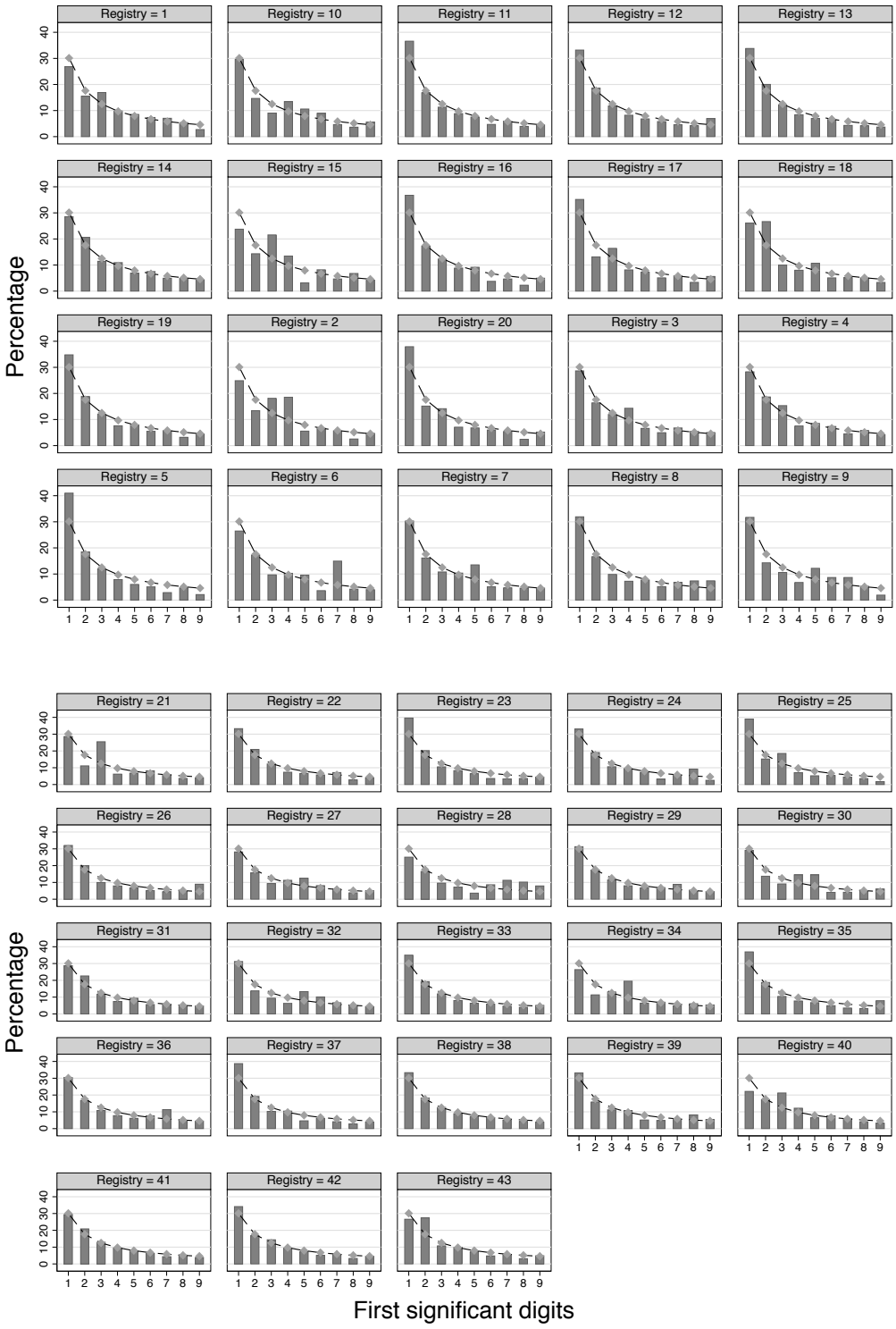


Figure 2: Theoretical (line) and observed distributions (columns) of first digits for all the analysed incidence rates, by registries and populations. 1: Malawi, Blantyre; 2: Argentina, Tierra del Fuego; 3: Brazil, San Paolo; 4: Ecuador, Quito; 5: USA, Virginia, Asian and Pacific Islanders; 6: USA, Nebraska, Black; 7: USA, Ohio; 8: USA, Vermont; 9: USA, Montana; 10: USA, Michigan; 11: USA, Georgia;12: USA, Indiana, White; 13: USA, Missouri, White; 14: USA, NPCR- National program of cancer registries (including42States); 15: USA, Colorado, Asian and Pacific Islanders; 16: USA, Arkansas,

Black; 17: USA, Alabama, White; 18: USA, Arkansas, White;19 :USA, California, Asian and Pacific Islanders; 20: USA, Connecticut, Black; 21: USA, Virginia, Black;22: USA, California; 23: India, Karunagappallyp; 24: Singapore, Malay; 25: Turkey, Edirne;26: Israel, Jews; 27: Japan, Hiroshima Prefecture; 28: Japan, Fukui Prefecture; 29: Israel; 30: France, Isère; 31: Germany, North Rhine–Westphalia; 32: France, Hérault;33: UK, England; 34: Estonia; 35: Switzerland, St. Gall-Appenzell; 36: Bulgaria; 37: Malta; 38: Ukraine; 39: Spain, Navarra; 40: Italy, Sondrio; 41: Germany, Brandemburg; 42: New Zealand: Other; 43: USA, Hawaii.

Discussion

In the present study, a considerable and heterogeneous sample of CRPs included in CI5-X was analyzed to evaluate, for the first time to our knowledge, if the FSD distribution of cancer incidence rates abided by Benford's law.

The results showed a substantial adherence of FSD distribution of cancer incidence rates to Benford's law.

This was not surprising. In fact, FSD distribution of cancer incidence rates had a priori some characteristics for being Benford prone. Indeed, they are the second generation distribution, being the result of the division of the number of cases diagnosed in a time span by the corresponding resident population, they comprise a large range of numbers covering several orders of magnitude (from units to thousands per 100,000 people, according to different ages and cancer types), and they are not influenced by human thought (Hill, 1995; Durtschi, 2004).

We verified that cancer incidence rates respect the quantitative measures suggested by Wallace (Wallace, 2002) to assess whether a distribution may be expected to respect Benford's law. In fact, the mean of their observed FSD is greater than the median, and their distribution has a positive skewness.

In the present study using graphic visualization, correlation coefficient, and some distance statistics, we observed that FSD distribution of cancer incidence rates abide by Benford's law when analyzed overall, by sexes, excluding half of the rates for the major cancer sites (female breast, colon–rectum, and lung and prostate cancers) and generally by CRP.

We have analyzed data which had been already examined for their quality and proved good for publication in CI5-X (Forman 2014). Therefore, no major problems in data quality were expected. Our results showed that for almost all the CRPs the FSD distribution substantially adhere to Benford's law. When the 43 CRPs were analyzed individually, the plot of their FSD distribution seemed to be in agreement with Benford's law. It must be mentioned that, due to sampling, two CRPs were sub-groups of the same registry (USA, Arkansas Black and White; USA Virginia, Asian and Pacific Islanders and Black) and another two included a subgroup and the whole population of the same registry (Israel and Israel, Jews and USA California and USA California, Asian and Pacific Islanders). No large difference within those cancer registries has been shown. Therefore, quality of cancer registry data and related activity (in terms of data availability, data collection, etc.) seemed unrelated to racial/ethnic subgroups at least in the analyzed registries.

The cancer registry data quality evaluation is not a perfect process, and some residual heterogeneity could exist also in CRPs included in CI5-X. In fact, in the introduction of CI5-X, it is stated that in the registry specific pages for some CRPs "an asterisk preceding the registry title indicates that special considerations (which may include underregistration) must be taken into account in interpreting the published rates or indicators of quality. . ." (Forman, 2014). Overall, asterisks were reported for 114/424 CRPs in CI5- X (26.9%), and in 11/43 (25.6%) in our sample. One of the two CRPs which had three or more statistics with the worst values for Benford's compliance had the asterisks (50%), in comparison with the others in the sample (10/31, 24.3%).

We highlighted that, although the majority of CRPs seemed to comply with Benford's law, at least two of them showed possible violation. Random fluctuations could have driven the observed results (Leemis, 2000) even though probability was low, but the coherence across the different applied statistics made inconsistency with the Benford's distribution more probable for these CRPs.

According to our experience, based on the analyzed dataset that has been already checked for data quality and accepted for publication (CI5-X), cancer registries

showing the poorest results had r value below 0.9 and m, $\chi 2/n$, and d* values higher than 0.10; presumably in a wild situation greater values are expected.

Adherence to Benford's law has been widely used not only to detect fraudulent data in business and administration (Tödter, 2009) but also to test data irregularities in scientific research (Hein, 2012). Frauds in cancer incidence data are not expected. However, non-adherence to the law may be a clue for further evaluation. The distance from the expected distribution may be the consequence of selections or incompleteness of the data collection, of rounding of small rates (Diaconis, 1979), of errors in data recoding or in data transfer.

The meaning of Benford's violation is a red flag showing an unusual behavior requesting further data examination (Leemis, 2000). Once a violation is suspected, a CRP, which owns more data than those we analyzed, should try to find out clues for the possible origin of the problem. Our suggestion is to look for the Benford pattern for incident cases based on different (combinations of) sources of information (pathology reports, hospitalization, death certificate, etc.) to detect any source-specific pattern. Moreover, the stability over time of the data flow for each information source and cancer site should be evaluated.

References: The original references of the paper have been included in the general list of the thesis.

2.2.1. Comment

The current quality control in the data registration of cancers is based on checking many of the indexes that have been described in chapter 2 (Quality of cancer data registry) and summarised in Table 2.

Table 2: Quality dimensions and parameters usually checked during the process for CR data quality evaluation.

	Comparability	Validity	Completeness	Timeliness
Classification & coding	X			
Incidence definition	X			
Multiple primary	X			
Incidental diagnosis	X			
DCO		X	X	
DCI		X		
Histology verification		X	X	
Missing information		X		
Internal coherence		X		
Time to report data				X
Historic data method			X	
Homogeneity in the area			X	
Steadiness in age curves			X	
Incidence among children			X	
M:I			X	
n. sources / notifications			X	
Independent case ascertainment			X	
Capture-recapture method			X	
DCN/M:I method			X	
The flow method			X	

These traditional indexes follow the individual steps of the clinical course of each patient, e.g., diagnosis, microscopic confirmation, death, and they are based on epidemiological measures (numbers or rates) calculated for each patient group (e.g., by age, sex, cancer type).

The quality evaluation of the cancer registry's data implies a comprehensive examination of these indexes.

The evaluation of these indexes is not homogeneous worldwide.

For example, in North America, the North American Association of Central Cancer Registries (NAACCR) sets different quality levels for the data (Gold and Silver) and gives each one of them clear cut-offs for a series of measures regarding validity, completeness and timeliness (https://www.naaccr.org/certification-criteria/).

On the contrary, an Editorial Board, which includes members of IARC and IACR selects out of the many submissions received, which data will be published in the CI5. As regards CI5-X, 80 out of the 370 cancer registries that had submitted data, did not pass the quality control (Forman, 2014). Furthermore, among the accepted data, some of them are marked with an asterisk. The asterisk "indicates that some circumstances may require special considerations (e.g. incomplete registration) when interpreting the published rates or indicators of quality" (Forman, 2014).

The rules applied for CI5 are inexplicit.

When clear cut-offs for the indexes are missing the quality evaluation process cannot be reproduced, and the role of experts cannot be contested. Moreover, the whole quality evaluation process and the role of single the indexes in determining quality remains unclear.

For example, in Table 3, the quality indexes for 'All sites except non-melanoma skin cancer', in men and women, made available from CI5C-XI (http://ci5.iarc.fr/CI5-XI/Pages/Indices_sel.aspx) are presented. The percentages indicate: microscopic verification (MV%), cases known by a death certificate (DCO%) and cases coded in the topography as 'Other & Unspecified' (O&U%). For each index and each continent, the extreme values (the highest and the lowest) are shown, and there is also a symbol to indicate missing information (^). Moreover, the table includes the number of registries/populations for which there is a warning (asterisk).

Table 3: Some indices of data quality for 'All sites except non-melanoma skin', CI5-XI (C00-96 excl. C44) based on original data from Cancer incidence in 5 continents XI, are available at http://ci5.iarc.fr/CI5-XI/PDF/INDICES/21.pdf. Distribution for each continent (C&S= Central&South) of CR/populations with asterisks, range of proportion of microscopic verification (MV), death certificates only cases (DCO), and unspecified cancer site (O&U), in men and women. ^= missing value.

Continent	*/	MV %		DCO % (min-max)		O&U	
	CR	(min-max)				(min-max)	
		Men	Women	Men	Women	Men	Women
Africa	6 /7	53.9-97.8	54.5-98.7	^1.4-13.3	^0.8-9.6	2.2-5.6	1.7-3.1
America C&S	10/31	69.3-95.8	71.0-95.9	^0.9-14.9	^0.5-12.1	0.9-7.7	1.4-6.7
America North	41/171	77.1-97.7	76.1-97.5	^0.4-8.6	^0.2-7.9	1.2-4.3	0.7-3.8
Asia	31/96	44.2-96.9	58.2-98.3	^0.0-11.8	^0.0-11.9	0.1-23.4	0.1-13.6
Europe	23/128	73.0-93.9	75.7-97.2	^0.1-13.0	^0.1-14.5	0.7-4.7	0.7-4.4
Oceania	9/32	73.0-94.8	83.5-97.8	^0.1-14.7	^0.0-10.1	1.2-2.8	1.5-3.9

First of all, it is worthwhile mentioning that the definitions of data quality for IARC/IACR and of NAACCR are not coherent. Several of the CRs accepted for CI5-XI do not reach the Silver qualification standard required by NAACCR (that requests DCO<5%).

Moreover, each of the indexes presented in Table 3 shows an incredibly broad range of values. For example, MV% varies among men by 20-50 percentage points, and the same is true for women.

As regards DCO%, the size of variation is up to 100 times (from 0.1% to 14.7%). Finally, for the 'Other & Unspecified sites, there is also a variation of 100 times (from 0.1% to 23.2).

These values belong to the CR which have passed the CI5C-XI quality evaluation.

The heterogeneity of the values of some of the quality indexes used in CI5 in CRs data's evaluation shows that extreme values can be compatible with high-quality data.

The purpose of Benford's law is, through a preliminary check, to address the following traditional process. In case of violation, the evaluation of traditional indexes has to be stricter than in the case of abidance.
Benford's law applies a mathematical law which is not influenced by the examiner's level of knowledge regarding the process of cancer registration.
It does not require the involvement of experts.

Furthermore, this law may be tested and applicated in other phases of registration, e.g., to check the digital sources of information or the population data.

The very nature of this law is at the heart of the evaluation process, helping the reviewer divide presumably poor quality CRs from good quality ones.

The reason I analysed CRs already accepted for publication is because I did not have access to 'raw' data.
I proposed using the N-B law to Freddy Bray and Eva Steliarova from IACR. However, IARC's quality check methods will not be incorporating this method for the time being. Big International projects, such as CI5, should include data from different countries and continents. This, though, implicates a variability in quality due to variability of resources, access to data, legal barriers involved in accessing documentation, etc. For such studies, other reasons not only those related to the strict quality standards must to be taken into consideration. In fact, in the introduction of CI5-X, this point has been raised "… the editors also have to accept that some datasets from registries in less-developed countries may be less complete than those from registries in more developed ones because of problems of underdiagnosis due to local medical and economic situations (as opposed to under registration) and/or problems with the enumeration of the population. But these datasets are nevertheless of great interest because they describe cancer incidence in populations for which no other information is available and which maintain unique cultural habits that could provide valuable clues to cancer aetiology" (Forman, 2014).

However, I hope to have the opportunity to test Benford's law on the data collected by ENCR-JRC (https://www.encr.eu/sites/default/files/ENCR-JRC2015_Protocol.pdf), and reviewed with the traditional methods. The application of Benford's law would make it possible to evaluate its effectiveness in a real setting.

In the meantime, N-B law has been applied in digital gene expression datasets (Karthik, 2016), in 'soiled' numbers in known scientific fraudulent articles (Hüllermann, 2017), and also for evaluating the reasons for changes in large datasets homogeneity over time (Lee, 2019).

2.3. Original material: an example of the application of Benford's law to skin melanoma in the German registries' data published in CI5-XI.

The compliance of the distribution of incidence rates to Benford's law can be helpful for a preliminary check of registry dataset quality. It is critical to use the most considerable number of incidence rates (age, site, morphology, sex, etc.) to retain a considerable amount of values for checking the correspondence between expected and observed distribution of the first significant digits (FSD).

If we select a subset of data, for example, one cancer site or one incidence period, we may manage too few numbers to test abidance to Benford's law reliably.

However, considering the linking theme of this thesis is skin melanoma, Benford's law has been applied equally to a sample of skin melanoma cases.
Data came from those published in CI5-XI and made freely available (http://ci5.iarc.fr/CI5-XI/Pages/download.aspx). An extended data set, including 241 categories, based on a combination of ICD-10 three- or four-character site codes and ICD-O-3 morphological groups, was chosen.
Among these groups, all the codes available for skin melanoma (Melanoma of skin: C43; Head: C43.0-4; Trunk: C43.5; Upper limb: C43.6; Lower limb: C43.7; Other and unspecified: C43.8-9) were selected.

As a sample dataset, I chose the nine available German registries: Bavaria, Bremen, Hamburg, Lower Saxony, Munich, North Rhine-Westphalia, Rhineland-Palatinate, Saarland, Schleswig-Holstein. All of them contributed to the publication with incidence data for the period 2008-2012. Age- and sex-specific incidence rates for the period 2008-2012 were available. In theory, the maximum number of expected incidence rates from each registry (if there was at least one case for each combination) was: 1 (period 2008-2012) X 2 (sexes) X 18 (age-classes) X 6 (site and subsites skin melanoma), adding up to 216.

The same tests described and used in the previous paper (Paper 1) checked the correspondence between the first significant digit distribution of incidence rates and Benford' s law, in this series.

Results

Considering the nine German registries collectively (1,652 observations), the distribution of the FSDs of melanoma incidence rates appeared to be positively skewed (0.66), with the mean (3.65) higher than the median (3.0). These values were close to those of the theoretical Benford's distribution (skewness 0.8, mean 3.44, and median 3.0.) The same was mostly valid for the ratios between 1st vs 9th (observed 5.2 vs Benford 6.6) and between 1st vs 2nd (2.1 vs 1.7) digits. The graphic distribution of observed FSDs for incidence rates for the nine German registries reproduces the theoretical Benford's distribution quite well. Figure 3.

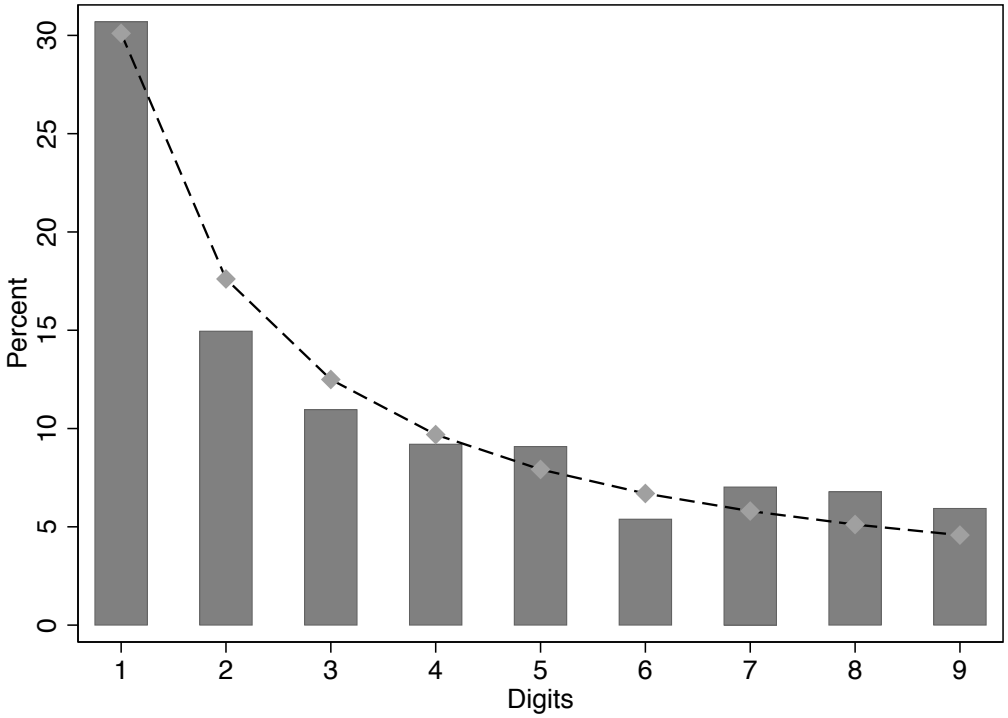


Figure 3: Benford (line) and observed (columns) distributions of the first digit for skin melanoma incidence rates for the nine German registries included in CI5-XI. Data retrieved from the extended dataset available at http://ci5iarc.fr/CI5-XI/Pages/download.aspx.

Pearson's correlation coefficient, r, was extremely close to 1 (0.98) demonstrating a strong correlation between the observed FSD distribution and the expected one. Moreover, all the measures of the distance between the distributions (theoretical and observed) were minimal (m=0.027 and d*=0.042), and the average Z was below the significance level for alpha=0.05 and one side tail (1.64, z=0.49). Finally, the χ^2 test, weighted on the number of observations (χ^2 /n), was low (0.022) as well.

The FSDs distributions confirmed that melanoma incidence rates broadly follow Benford's curve shape, Figure 4, for each German registry, with data sets varying from 157 to 211 observations.

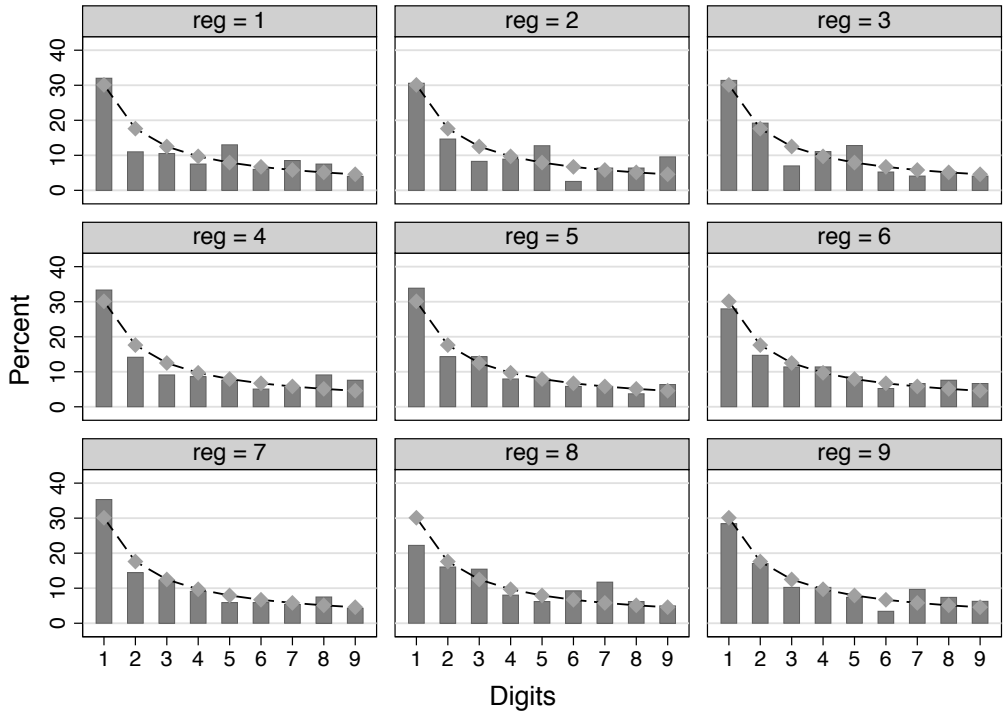


Figure 4: Benford (line) and observed (columns) distributions of the first digit for skin melanoma incidence rates for each of the nine German registries included in CI5-XI. Data retrieved from the extended dataset available at http://ci5.iarc.fr/CI5-XI/Pages/download.aspx.

In Table 4 the results of the applied tests are presented.

Results show a relatively high overall correspondence between Benford's and observed distributions of the first significant digits.

Table 4: Measures of correspondence between Benford and observed distributions of the first digit in the incidence rates for skin melanoma for the nine German registries included in CI5-XI. Data retrieved from the extended dataset available at http://ci5.iarc.fr/CI5-XI/Pages/download.aspx.

CR	r	m	d*	X^2/n	zeta
1	0.91	0.066	0.094	0.09	0.25
2	0.91	0.05	0.09	0.13	0.19
3	0.95	0.055	0.078	0.067	0.13
4	0.95	0.039	0.077	0.075	0.19
5	0.98	0.037	0.059	0.029	0.09
6	0.98	0.029	0.053	0.036	0.12
7	0.98	0.052	0.067	0.032	0.10
8	0.90	0.078	0.11	0.11	0.22
9	0.96	0.039	0.063	0.064	0.11

r = Pearson's correlation; m = maximum distance in absolute terms between expected and observed frequencies for each of the nine-digit ; d* = normalised Euclidean distance between the two distributions divided by the maximum possible distance; X^2/n = Chi-squared divided by the number of observation; zeta = average Z value for each comparison between the nine observed distributions of FSD and the theoretical one.

The evaluation of melanoma incidence rates FSD distribution should be tested against Benford's law at the beginning of the registries' quality of data evaluation process as the aim is to detect notable distortion which may suggest a selective loss of data.

In this example, the analysed dataset was already checked for data quality and accepted for publication in CI5-XI. Therefore, no critical violations to Benford's law were expected.

In case a violation of Benford's law is suspected, the traditional tools must be implemented very carefully, including the evaluation of the completeness of the sources of information.

Finally, although for a specific cancer site the number of observations (incidence rates) may be relatively low, also incidence rates for skin melanoma in the 9 German registries published in CI5-XI revealed an FSD distribution which abides by Benford's law, both altogether and individually.

3. Publication of registry data

Once the data of a registry are of good quality, epidemiological indexes are computed. Registry data are published in various forms to fulfil diverse purposes.

As regards incidence, the number of new cases is relevant to quantify the burden of diagnosis and first treatment needed in the area. The same information comes from the crude rate for units of the population at risk, usually per 100,000 people per year. Furthermore, incidence rates may be specific for sub-groups of the case series: for cancer type (topography and morphology), for age-group, for sex, etc.

Finally, age-standardisation (a method for performing reliable comparisons between two or more registries, or in the same registry over time), allows, with a direct or indirect approach, the consideration and adjustment for the age-group composition of the population under examination (Boyle, 1991).

Each measure may be complemented by the estimate of the error (the standard error), and/or confidence intervals.

Table 5 is an example of the traditional way of presentation of incidence data. For selected registries from CI5-XI and sex, there is the number of cases (all cancer sites except non-melanoma skin), the age-adjusted incidence rate (ASR based on the World standard population), and the corresponding standard error (s.e.).
Besides, this is the traditional approach applied for prevalence and survival, with their specific indexes.

Table 5: All sites except non-melanoma skin (C00-96 excl. C44). Distribution of the number of cases, age-adjusted incidence rate (ASR on the World standard population) and standard error (s.e.) for selected registries and sexes, 2008-2012. Based on primary data from Cancer incidence in 5 continents XI, available at http://ci5.iarc.fr/CI5-XI/PDF/INDICES/21.pdf. *= missing value.

	Women			Men		
registry	cases	ASR	s.e.	cases	ASR	s.e.
Belgium	6187	15.3	0.21	4345	10.5	0.17
Cyprus	142	4.6	0.40	156	5.3	0.44
Czech Republic	4766	10.7	0.17	5278	12.7	0.18
Ireland	2342	14.6	0.32	1835	12.1	0.29
Italy, Romagna	680	12.6	0.56	750	13.7	0.56
Malta	140	9.4	0.86	111	6.9	0.70
The Netherlands	12950	20.6	0.19	11070	17.2	0.17
Portugal, Azores	38	5,2	0,91	43	6.7	1.05
Ukraine	9435	4,7	0.05	6573	4.5	0.06

The following Paper (2) presents an innovative method which complements the traditional ways of giving incidence data and provides a measure of heterogeneity in sub-areas.

3.1. Paper 2

The paper was published in Eur J Cancer Prev. 2017; 26: 442-446 with the following title:

"Variability of cancer risk within an area: time to complement the incidence rate.

Authors Crocetti E, Giusti F, Martos C, Randi G, Dyba T, Bettio M.
Affiliation: European Commission, DG Joint Research Centre (JRC), Institute for Health and Consumer Protection, Public Health Policy Support Unit, Ispra (VA), Italy
Author Contributions: I (EC) declare to have conceived the idea of the study, planned and designed it, performed the analysis and drafted the first draft. The other Authors revised critically the paper and approved the final version of the manuscript

Abstract

The aim of this study was to show that age-adjusted cancer incidence rates for an area may not be representative of the incidence in subareas. We propose a simple measure to show the amount of geographical variability. European age- standardised incidence rates (ASRs) for 'all sites excluding non melanoma skin cancer', for men, in 2014, for Nordic countries as a whole, for each country (Denmark, Faroe Islands, Finland, Greenland, Iceland, Sweden and Norway) and for their regions, were retrieved from the Nordcan with corresponding standard errors (SEs). We compared the ASR for Nordic countries versus single country and single country versus specific regions. The overlapping of 95% confidence intervals was used for ASRs comparisons. As a measure of variability, we computed the range between the highest and the lowest ASR within an area and the ratio between this range and the ASR of the overall area, r/R = (range/ASR) × 100. The 95% confidence interval of the ASR for Nordic countries as a whole did not overlap those of the majority of the single countries; in fact, the r/R – which provides a clue for the amount of underlying geographical variability – was rather large (57.1%). Within countries, the variability was negligible in Iceland (r/R = 9.6%), whereas the highest value was found in Sweden (37.1%). The ASR does not provide any information on underlying geographical variability. Therefore, its interpretation could be misleading. When data for subareas are available, the r/R,

which is simple to compute and to understand, should be added to the ASR for providing more truthful information.

Introduction

As a standard practice worldwide, population-based cancer registries (CRs) express the occurrence of cancer in a defined population in a certain period as the ratio between the newly diagnosed cancers and the at-risk resident population. This ratio is called the crude incidence rate (Boyle and Parkin 1991). Cancer incidence increases with the ageing of the population. Therefore, incidence rates are strongly dependent on the age structure of the underlying population. Consequently, rates are computed using a standard age structure as a reference (age-standardised rate, ASR) to enable reliable comparisons across time and countries (Boyle and Parkin 1991). Crude rates and ASRs are the standard indicators reported by all CRs independent of the size of the population at risk. These statistics are usually complemented by a measure of precision, the standard error (SE) of the rate and/or the 95% confidence intervals (CIs).

An incidence rate expresses the summary probability of developing cancer in the area covered by the CR. It provides no clues on the homogeneity or the heterogeneity of incidence rates across subareas.

To gain more insight into this topic and to explore the possible variability in ASRs among subareas of CRs, we analysed the data of the Association of the Nordic Cancer Registries, which makes data available in the Nordcan project (Engholm 2016).

Methods

We retrieved from the Nordcan the European ASRs for 'all sites excluding non melanoma skin cancer', for men, in 2014.

ASRs are available for three geographical layers as presented in Table 6 (n. 1 in the original paper): (a) Nordic countries as a whole; (b) single country: Denmark, Faroe Islands, Finland, Greenland, Iceland, Sweden, Norway; and (c) regions: five in Denmark: North Jutland, Central Jutland, Southern Denmark, The Capital and Zealand region; five in Finland: Helsinki, Kuopio, Oulu, Tampere and Turku region; two in

Iceland: Reykjavik-Reykjanes and Outside The Capital; six in Sweden: Northern, Stockholm–Gotland, Southern, South-Eastern, Uppsala-Örebro and Western region; and four in Norway: Central, Northern, South-Eastern and Western region, Figure 5 (not present in the original paper) (from http://www-dep.iarc.fr/NORDCAN).

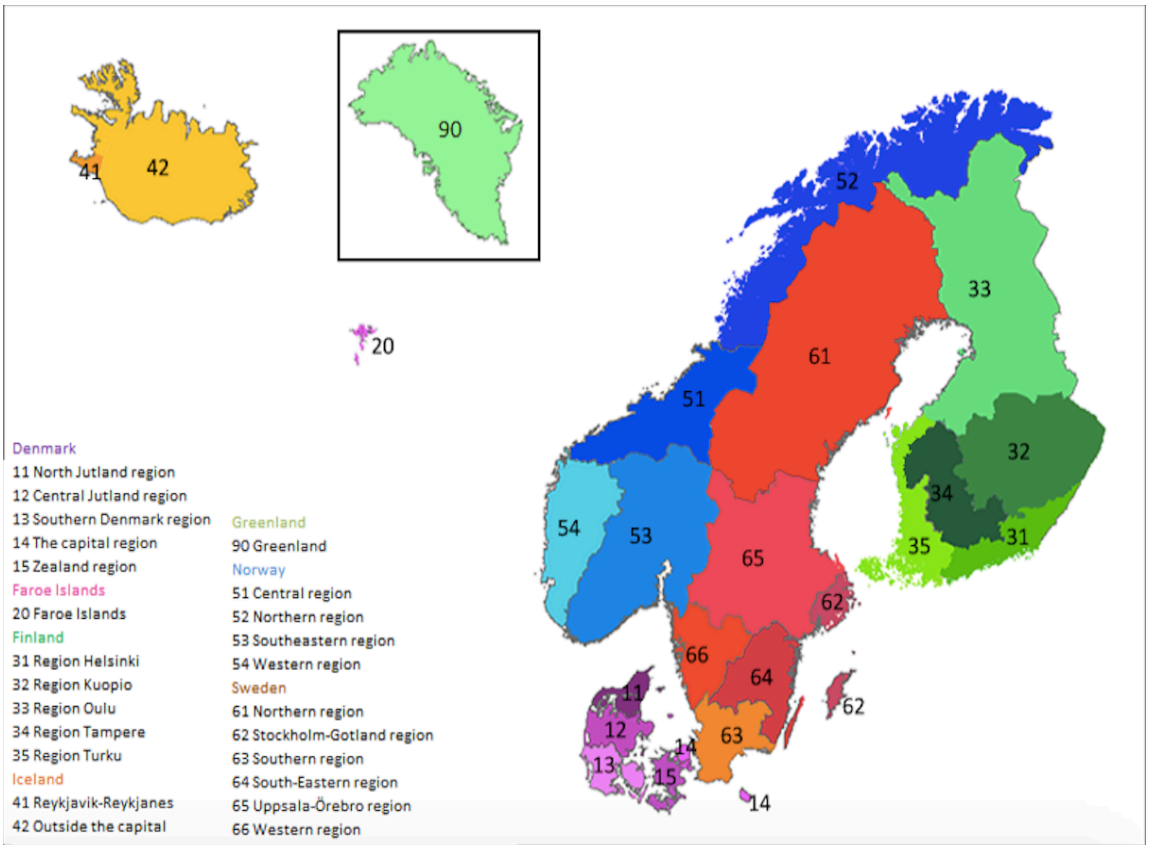


Figure 5: Countries and regions belonging to the Nordcan project (http://www-dep.iarc.fr/NORDCAN).

The overall male resident population in Nordic countries in 2014 was 13 075 123. Residents in single countries and regions are shown in Table 6 (n.1 in the original numeration).

ASRs express the number of new cases diagnosed among 100 000 men in 2014 according to the observed age- specific rates and the age-groups of the European standard population.

We also retrieved from Nordcan the SE of the ASRs and we computed the 95% CIs according to the method of the binomial approximation (Boyle and Parkin, 1991) (Table 6 [n.1 in the original numeration]).

We evaluated whether two rates were different inspecting the overlap between specific 95% CI (Schenker and Gentleman, 2001). The precision of the age-specific rates that concur in the calculation of ASR increases when the number of cases in this group increases. This applies to each age group and thus to ASR as the whole entity. The overall numbers observed yearly in the analysed series (Table 6 [n.1 in the original numeration]) were, with the exception of Faroe Islands and Greenland, in the order of several hundreds or even thousands. SEs of the ASR are greater when the numbers on which they are based are small.

We compared the ASR at each geographical level with the level underneath: Nordic countries versus Denmark, Faroe Islands, Finland, Greenland, Iceland, Sweden, Norway and single country versus specific regions.

Moreover, we computed the absolute difference (range) between the highest and the lowest ASR within a nested layer (the range between countries for Nordic countries and between regions for a specific country). Then we calculated the percent ratio between this range and the ASR of the level above interpreted as a summary value of the subareas (for Nordic countries or a single country, respectively), r/R = (range/ASR) × 100.

The r/R provides a measure of the variability across the available ASRs of the nested level for which the ASR represents the summary measure. The smaller the r/R (minimum 0%), the lower the variability across subarea ASRs.

Results

In Figure 6 (n.1 in the original numeration), the ASRs for Nordic countries as a whole, for single countries and for country-specific regions, are shown with the corresponding 95% CI. The ASRs for countries appear to be scattered in the picture.

In fact, the ASR for Nordic countries (453.1) is compatible with the Greenland's one only (384.1) because of the wide range of variability of the latter, because of the small number of cases on which it is based and the resulting imprecision in its computation (wide 95% CI). In contrast, Denmark (504.4) and Norway (509.9) showed greater ASRs that the Nordic countries one and Faroe Islands (251.0), Finland (404.5), Iceland (387.2) and Sweden (428.5) have lower values than the supranational summary ASR.

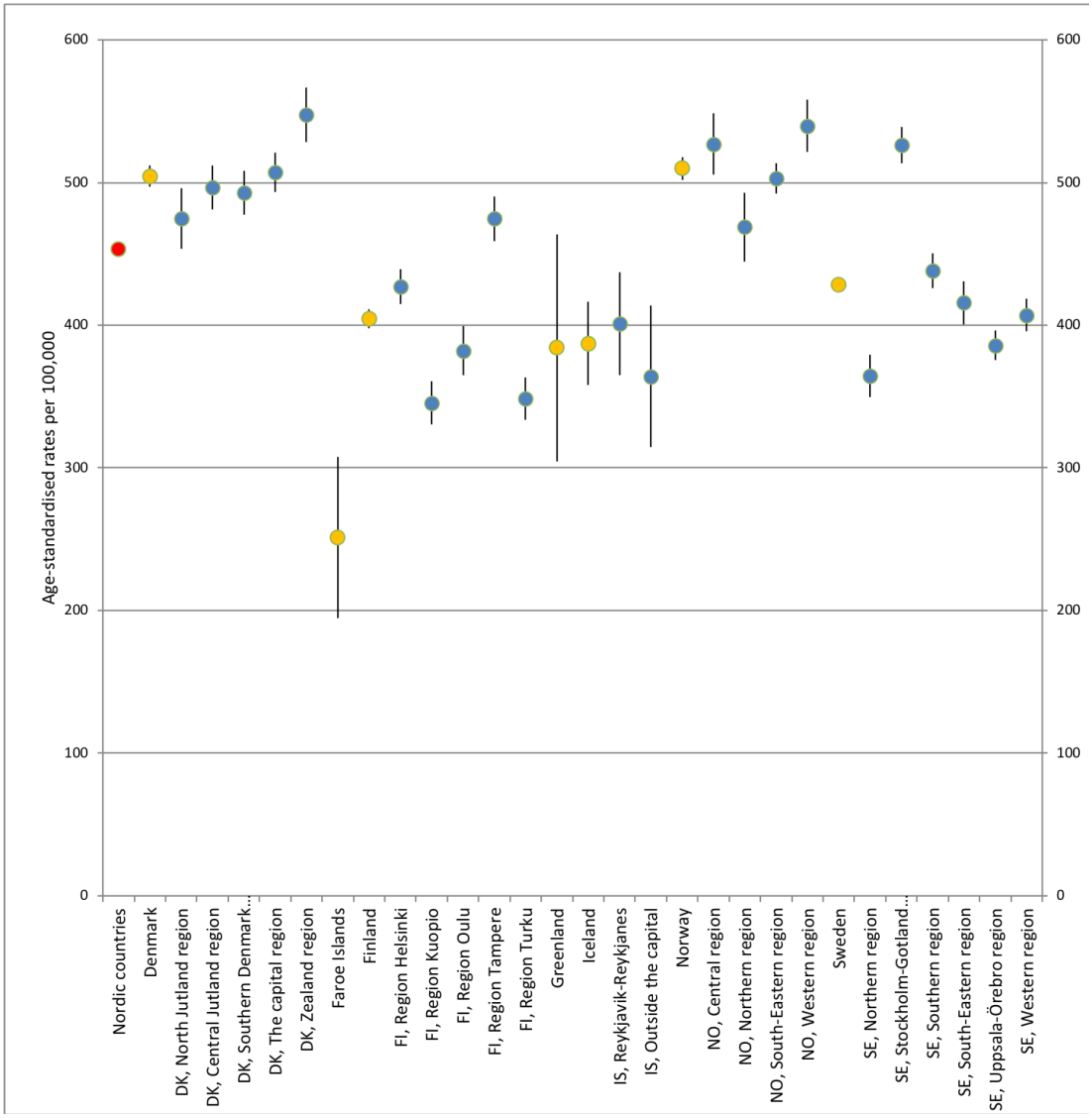


Figure 6 (n.1 in the original numeration): European age-adjusted incidence rates for 'all sites excluding nonmelanoma skin cancer', for men, in 2014, for Nordic countries, single countries and regions. From Nordcan (http://www-dep.iarc.fr/NORDCAN).

In Table 6 (n.1 in the original numeration), for each area (Nordic countries, country and region), the number of cancer cases for 'all sites excluding non melanoma skin cancer', for men, in 2014 and the resident population are reported together with the ASR, the SE and the 95% CI.
The inverse relationship between number of observed cases and the SE is evident. In fact, the SE is only 1.6 (cases per 100 000 men in 2014) for Nordic countries (on the basis of 79 441 analysed cases), whereas it is 40.6 for Greenland (99 cases).

In Table 6 (n.1 in the original numeration), the r/R is also reported for geographical level 1 (Nordic countries vs. countries) and 2 (single countries vs. regions).

When the ASR of Nordic countries is evaluated together with the r/R, the value of r/R = 57.1% provides a clear hint of a huge inter-countries variability in ASRs, clearly shown in Fig. 6. ([n.1 in the original numeration) In fact, this r/R means that the range between the lowest and the highest country-specific ASR is almost 60% of the Nordic country ASR.
Also, within single countries, the overall ASR may not represent the regional ASRs and the amount of internal variability (Fig. 6 [n.1 in the original numeration]) is well described by r/R (Table 6 [n.1in the original numeration]).

The smallest r/R value (9.6%) was observed in Iceland, where the small numbers of observed cases led to a non-negligible uncertainty in the regional estimates whose wide 95% CI overlapped the national one. A minor amount of variability (r/R = 13.9%) was present in Norway, where the Northern region (ASR = 468.8) had a lower value and the Western region (539.8) had a higher ASR than the summary one. Almost the same r/R was present in Denmark (14.4%), where North Jutland (474.8) showed an ASR lower than the national value and Zealand (547.5) showed a greater one. Finland showed a greater inter-regional variability (31.9%), with Kuopio (345.5) and Turku (348.5) below and Helsinki (427.2) and Tampere (474.6) above the national mean. Finally, the slightly higher internal variability was found in Sweden (RR = 37.8%) where three regions, Northern (364.5), Uppsala-Örebo (385.8) and Western (407.1), were below the national ASR and Stockholm–Gotland (526.3) higher than the country one.

Table 6 (n.1 in the original numeration): Country layers, number of incident cases of 'all sites excluding non-melanoma skin cancer', in men, in 2014, resident population, European age-standardised incidence rates (ASR), standard error (SE) of the ASR, lower (LCI) and upper (UCI) 95% confidence intervals and R/R (Range/Rate : Range is the absolute difference in ASRs between the greatest and lowest ASR for areas in the lower layer; Rate is the ASR) [The smaller the r/R the lower the variability across subareas ASRs. range/Rate% (r/R): range is the absolute difference in ASRs between the greatest and the lowest ASR of subareas in the lower layer; rate is the ASR. ASR, age-standardised incidence rate; LCI, lower 95% confidence interval; UCI, upper 95% confidence interval.] Data from Nordcan (http://www-dep.iarc.fr/NORDCAN).

Layer	Area	n. cases	resident population	ASR (E)	SE	LCI	LCS	R/R
1	Nordic countries	79,441	13,075,123	453.1	1.6	449.9	456.3	0.46
2	Denmark	19,031	2,799,895	504.4	3.7	497.1	511.7	0.14
3	North Jutland	2,004	292,697	474.8	10.9	453.4	496.2	
3	Central Jutland	4,165	639,192	496.5	7.8	481.2	511.8	
3	Southern	4,230	600,667	493.0	7.8	477.7	508.3	
3	The capital	5,245	860,818	507.1	7.1	493.2	521.0	
3	Zealand	3,387	406,521	547.5	9.7	528.4	566.6	
2	Faroe Islands	78	25,039	251.0	28.8	194.6	307.4	
2	Finland	15,142	2,686,119	404.5	3.4	397.9	411.1	0.32
3	Helsinki	4,849	922,582	427.2	6.2	415.0	439.4	
3	Kuopio	2,194	403,920	345.5	7.7	330.5	360.5	
3	Oulu	1,970	372,534	382.2	8.8	364.9	399.5	
3	Tampere	3,811	545,749	474.6	7.9	459.1	490.1	
3	Turku	2,303	441,350	348.5	7.5	333.7	363.3	
2	Greenland	99	29,742	384.1	40.6	304.5	463.7	
2	Iceland	694	164,257	387.2	14.8	358.1	416.3	0.10
3	Reykjavik	481	115,443	401.2	18.5	365.0	437.4	
3	Outside	213	48,818	364.1	25.4	314.3	413.9	
2	Norway	15,865	2,581,421	509.9	4.1	501.9	517.9	0.14
3	Central	2,362	357,476	526.9	11.0	505.3	548.5	
3	Northern	1,489	242,918	468.8	12.4	444.6	493.0	
3	South-Eastern	8,685	1,433,445	502.9	5.5	492.2	513.6	
3	Western	3,329	547,582	539.8	9.4	521.3	558.3	
2	Sweden	28,709	4,843,303	428.5	2.6	423.4	433.6	0.38
3	Northern	2,521	444,391	364.5	7.6	349.6	379.4	
3	Stockholm -G.	6,735	1,111,680	526.3	6.5	513.6	539.0	
3	Southern	5,375	872,866	438.1	6.2	425.9	450.3	

3	South-Eastern	3,130	510,943	415.8	7.7	400.6	431.0	
3	Uppsala-Örebro	5,852	1,002,193	385.8	5.3	375.5	396.1	
3	Western	5,096	901,258	407.1	5.9	395.6	418.6	

Conclusion

This epidemiological exercise underlines that ASRs, which clearly provide the level of cancer incidence in a specific area and time for geographical and time comparisons, do not provide any information on possible internal variability. In fact, the SE, which usually accompanies ASR, refers only to the precision of the estimate and does not reflect the possible heterogeneity in cancer incidence in the area.

Therefore, the ASR of a CR, although correct from the computational point of view, and informative for geographical and time comparisons, could represent the incidence level only in some subareas or even in none.

If a CR also provides ASR for subareas, r/R is not necessary because the information on possible geographical heterogeneity is available. In contrast, if a CR only publishes a summary ASR, as happens for many CRs in Cancer incidence in five continents (Ferlay 2014), which is the most well-known and authoritative publication in the field, r/R is invaluable to have a clear impression of the variability behind the ASR.

When incidence data are available for different geographical layers, it is possible to add to the ASR a summary measure about the underlying variability. The Nordic countries dataset provided the invaluable chance of evaluating three sub-geographical levels: supranational, national and regional.

We propose to compute the range between the highest and the lowest underlying ASRs to divide it by the ASR (r/R) and to express the result as a percentage.

The inde
x r/R has been chosen among other more formal statistics (e.g. extreme quotient) (Gumbel and Keeney, 1950) because it only relies on ASRs and provides a direct measure of the effect of internal heterogeneity (range between maximum and minimum ASR in subareas) on the overall summary ASR.

In our example, on the basis of long-standing high- quality Nordic countries incidence data (Ferlay, 2014), the r/R for the Nordic countries was quite high (57.1%), suggesting that the national ASRs could vary notably. In fact, the overall ASR for Nordic countries did not correspond with any of the national ASRs, with the exception of Greenland's one (Fig. 6 [n.1 in the original numeration]).

Also at a national level, when regional estimates are available, it is possible to add to the national ASR the r/R based on regional ASRs to express how well the national ASR represents the regional ones. In the dataset analysed, we showed that country ASR may reflect more (Iceland, Denmark and Norway) or less accurately (Finland and Sweden) the incidence of cancer in the different regions within a country.

The comparison between ASRs using the 95% CI overlap is simple and intuitive (Schenker and Gentleman 2001) and showed major differences in ASRs between and within areas.

This study was based only on one incidence year. To check the reliability of r/R, we repeated the exercise also for the year 2012. The r/R in 2012 were similar to that in 2014 (results added in Appendix 2 [Supplementary information for Paper 2]; data not shown in the original paper) for almost all the countries, with the exception of Sweden, for which r/R showed in 2012 a smaller heterogeneity (r/R = 16.6%) than in 2014 (37.8%). The reason for this strong change was the change in the incidence ASR in the Stockholm–Gotland region from 2012 (420.5 cases/100 000) to 2014 (526.3). This change was the effect of a study on prostate cancer carried out in the county between 2012 and 2014 (Grönberg ,2015). The ASR for all causes except skin and prostate cancer were 268.3 and 266.7, respectively. This example confirms that r/R reflects the true variability within an area.

Heterogeneity was identified among countries (areas between around 25 000 and 4 800 000 resident men) and among regions of several hundred thousand inhabitants, except for Iceland, where the population is smaller than in any of the other countries with regional information available.

It is possible to identify slight differences in cancer incidence between two geographical areas if the number of cases (population) is large (i.e., thousands). Then, the ASRs are precise and the 95% CI is narrow. Thus, it is easier to detect a slight difference between two large (populated) regions than between two small ones. For example, between the ASRs of Kuopio and Oulu (highly populated), there is the same difference as that between the two Icelandic regions (poorly populated), but only the first two do not have overlapping CIs.

In general, the unavailability of a unique population-unit for subareas (countries, regions, provinces, counties, etc.) makes comparisons across areas difficult.
With the increase in the number of subareas, the variability among them is expected to increase and consequently the r/R. The aim of r/R is exactly to offer summary and straightforward information on possible outliers. In case r/R is small ($\sim$ < 10–15%) it is immediately clear that all the ASR for each of the subareas are concentrated in a quite narrow range and if it is large (>30%) it underlines that at least one of them is rather different from the overall ASR.

The r/R is a measure intended as a macro indicator of major heterogeneity among quite large subareas (e.g. regions in a country). For small areas and cluster analysis, other methods have to be chosen (Colonna and Sauleau, 2013).

CRs should start to provide also general information on internal cancer incidence geographical variability in addition to standardised incidence rates. This would make the information more complete and clearer for readers, avoiding misinterpretations. When incidence for subareas is available, r/R, which is very simple to compute, could be presented together with the general ASR as a first attempt to raise the issue.
The interpretation of incidence ASR requires the combined reading of ASR, SE and r/R: the ASR shows the level of incidence, the SE shows the precision of the ASR and r/R shows the amount of internal geographic variability. The r/R will be smaller if the ASR for subareas are quite similar to each other (more or less precisely estimated) or

greater if they are rather different. This is the original and useful contribution provided by the r/R.

References The original references of the paper have been included in the general list of the thesis.

3.1.1. Comment

The previous paper (Paper 2) addresses the topic of informativeness and usefulness.

The proposed innovative index provides, when necessary, a measure of the variability in cancer incidence in addition to an average ASR.

It is straightforward to calculate and therefore does not require further variables.
It is also straightforward to interpret.

The paper perfectly addresses the objectives of this thesis, illustrating how it is possible to exploit the potentiality of CR, enhancing the amount of information provided, improving the readability of a theoretically complex issue, and providing the stakeholders with a more comprehensive message.

This method has been implemented and presented in recent scientific meetings on cancer registries (Contiero 2018, Martos 2018, Rashid 2018).

The heterogeneity within an area in cancer incidence may also concern survival or prevalence. Therefore, this approach, tested with incidence, could also be exploited for other epidemiological indexes.

3.2. Original material: an example of the application of the new index of internal variability (r/R) to skin melanoma incidence

Since the core theme of this thesis is skin melanoma, the r/R method has also been applied to a skin melanoma case-series. The aim is to complement a summary age-adjusted rate (ASR) with this index of internal variability to provide a straightforward clue of variability of incidence in macro sub-areas.
The computation of the method implies that age-adjusted rates for sub-areas are known but not published.

I utilised data from Nordcan (http://www-dep.iarc.fr/NORDCAN/english/frame.asp), the same source of information used in the original paper, which provides information at three geographical levels: supranational, national and regional (Engholm, 2016). I selected ASR (European standard population) and standard error for skin melanomas incident in 2015 among men for All Nordic Countries, for each Country, and Regions, when available. I computed the r/R index to the Nordic Countries to provide a measure of inter-Countries variability, and to each Country for which Regional information was available to evaluate Regional variability.

Table 7 summarises results.

Variability is exceptionally high for Nordic Countries as a whole (r/R=1.10). The value means the underlying difference between the Countries with the highest and lowest ASR exceeds the ASR for Nordic Countries; such results are primarily due to the low rate in Greenland, 1.9, which is far from that of all other Countries.

However, age-standardised incidence for skin melanoma among men varies remarkably also among various regions of each country, in particular in Sweden where the internal variability, according to the index, is about 70% of the Swedish ASR, and in Finland, for which the r/R value is 0.55. These results confirm the usefulness and informativeness of the r/R index, which expresses supra-national or national ASR data representativeness for national and regional sub-areas incidence.

Geographical heterogeneity in Nordic countries was also documented in 1978-1982 and 1983-1987, with the highest incidence in Norway and the lowest in Finland and Iceland, and with regional high-risk areas in the south of Norway and the south of Sweden (Aase, 1994). The present results are in agreement with this previous paper (Aase, 1994). The Authors related the geographical difference mainly to variations in estimated local levels of UV radiation. Therefore, the Authors warned the readers about the effect of the on-going stratospheric ozone depletion. However, also difference in social level and the associated patterns of behaviour, and genetically determined susceptibility to exposure to UV were considered (Aase, 1994). The in-deep analysis of melanoma trends in Norway suggested that the documented increase in melanoma incidence could have been a result of more active outdoor recreation – and therefore UV exposure – from the end of the last century (Aase, 1996). The regional variability documented in Sweden, with more elevated rates in Western Sweden and in particular in the city of Gothenburg, was also related to the high average duration of the sunshine in the area, but also high sun exposure of citizens on holidays abroad (Claeson, 2012).

The difference in melanoma incidence among the leading five Nordic countries has also been documented by Hèry and colleagues who discovered broader differences in the risk than in prognosis of melanoma. However, the reasons for such heterogeneity were not clear (Hèry, 2010). As regards the relevant increase observed in Iceland among women from the late-1980s, and among men it was attributed at least in part to the very high prevalence of sunbed use, as well as considerable increase in travel abroad to southern areas during the past four decades and partly because of increased diagnostic activity due to a nationwide cancer prevention programme which came into action in 1990 (Hèry, 2010). A survey on tanning bed use put the participant Scandinavian countries at the first ranks among 30 European countries for people younger than 20 years. To be specific, the prevalence was 33.3 % in Norwegian adolescents, 23.9% in Denmark and 23.5 % in Sweden (Suppa, 2019). This result raises concerns. However, educational actions are on-going. For example, in Denmark, a successful campaign against the use of sunbeds has been carried out among adolescents (Køster, 2011) with beneficial effects that have increased from 2007 to

2015 (Køster, 2018). Such effects have pushed for more stringent legislation and/or enforcement of the existing regulations in those countries, especially for young individuals. In Denmark, efforts have been developed in recent years to reduce sunbed use among adolescents. An anti-sunbed campaign raised awareness of the sunbed-related risks among Danish adolescents, and indeed, sunbed use decreased substantially; until now, it has remained considerable (>30%) (Køster, 2011; Køster, 2018).

Moreover, an educational intervention in Danish schools produced a significant reduction in sunbed use but failed to change pupils' intentions and attitudes towards artificial tanning (Køster, 2011). Importantly, all three Scandinavian countries displayed elevated rates of long-term sunbed users (>10 years) in our analysis.
A survey was carried out by the Danish Cancer Society (https://www.ancr.nu/dyn/resources/File/file/7/4247/1412940269/total_document_survey_optimeret.pdf) in 2010 to evaluate the comparability of data among the Nordic cancer registries.

As regards melanoma screening, there were no programs in Denmark, Finland, Iceland and Norway. In Sweden, in the region of Lund, a particular "out-patient"-project, in which people had their "dots" checked, was organised in 1990. Moreover, still in Lund, random screening was offered by Cancer society at, e.g., beaches during the summer (https://www.ancr.nu/dyn/resources/File/file/7/4247/1412940269/total_document_survey_optimeret.pdf).

In conclusion, some of the Nordic countries have taken part in the Euromelanoma initiatives, Sweden since 2000, Denmark since 2011 and Finland since 2015. These actions include public awareness campaign, early detection and treatment, and public screenings during an annual 'Euromelanoma Screening Day'(https://www.euromelanoma.org/intl).

Table 7: Skin melanoma. Data from Nordcan (http://www-dep.iarc.fr/NORDCAN): Country layers, European age-standardised incidence rates (ASR), standard error (SE) of the ASR, and r/R (range/ASR: range is the absolute difference in between the greatest and lowest ASR for areas in the lower layer; ASR is the one of the upper layer).

Layer	Area	ASR(E)	SE	r/R
1	Nordic countries	27.8	0.42	1.10
2	Denmark	29.3	0.93	0.44
3	North Jutland	27.9	2.8	
3	Central Jutland	23.1	1.7	
3	Southern	28.9	1.97	
3	The capital	35.9	1.9	
3	Zealand	27.9	2.3	
2	Faroe Islands	4.8	2.8	
2	Finland	22.7	0.82	0.55
3	Helsinki	25.6	1.5	
3	Kuopio	18.9	1.9	
3	Oulu	14.7	1.8	
3	Tampere	27.2	2.0	
3	Turku	20.4	1.9	
2	Greenland	1.9	1.9	
2	Iceland	10.7	2.49	0.35
3	Reykjavik	12.0	3.2	
3	Outside	8.3	3.9	
2	Norway	32.6	1.04	0.36
3	Central	28.8	2.62	
3	Northern	23.3	2.81	
3	South-Eastern	35.1	1.44	
3	Western	33.2	2.33	
2	Sweden	28.0	0.68	0.70
3	Northern	14.1	1.59	
3	Stockholm-G.	26.4	1.43	
3	Southern	32.0	1.73	
3	South-Eastern	33.8	2.32	
3	Uppsala-Örebro	27.3	1.48	
3	Western	31.0	1.67	

4. Data from the registry: beyond topography and morphology

As shown in the previous table, the way registries publish data has not sensibly changed since the beginning of their history. The topography is still the primary way in which CRs present numbers and rates. Therefore, such information may raise limited interest among clinicians who diagnose and treat cancers for which new prognostic and predictive factors are available every day.

The interest of clinicians in registry data has to be encouraged.

The process of production of high-quality data is demanding in terms of time and resources. The inclusion of any extra variable in a registry opens a new front for validity, completeness and comparability evaluation and reduces timeliness. As a consequence, most registries plan their primary activity to meet the demands and deadlines of the big international projects (e.g., CI5, Concord, Eurocare).

Therefore, basic registries' information is mainly provided to those who ask exactly for them, for purposes based on their known availability. There is a high degree of recursion within the world of registries.

Further, big projects involving registries from different countries and even continents must have a global point of view with the need to find out the minimum data set, which allows the majority of registries to participate.
In case extra variables are needed, as started to happen with the High-Resolution studies of the Eurocare project, specific funds or just analysis on samples or both should be planned.

Besides, the data format has been kept simple and stable over time, also for making trend evaluations possible. Furthermore, big International projects have also marked the time; for example, CI5 publishes new data with a minimum delay of 5 years from the clinical diagnosis. This interval is partly due to the amount of work necessary in

CI5, but also to the timeliness of registry data production. This delay causes another critical recursive phenomenon: there is no pressure to shorten timeliness.

In conclusion, registry data may be of interest for those who want to compare incidence, prevalence, or survival in the world but may not be for the clinicians who work in the publishing registry's area.

The following Paper (3) presents a fruitful collaboration between registries and dermatologists in the analysis and interpretation of skin melanoma.

4.1. Paper 3

The paper has been published in Melanoma Res. 2010; 20: 422-6 with the following title:

"The thickness of melanomas has decreased in central Italy, but only for thin melanomas, while thick melanomas are as thick as in the past".

Authors Crocetti E[a], Caldarella A[a], Chiarugi A[b], Nardini P[b], Zappa M[a].
Affiliation: a) Clinical and Descriptive Epidemiology Unit and b) Melanoma Early Diagnosis Service, Institute for Cancer Study and Prevention ISPO, Florence, Italy
Author Contributions: I (EC) declare to have conceived the idea of the study, planned and designed it, performed the analysis and drafted the first draft. The other Authors revised critically the paper and approved the final version of the manuscript

Abstract

The objective of this study was to evaluate the time trend of melanoma thickness in a population-based case series. All invasive (n = 2862) and *in situ* (n = 605) cutaneous melanoma incident cases diagnosed in 1985–2004 were retrieved from the Tuscany Cancer Registry, central Italy.

Standardized (European population) incidence rates were computed for four periods: 1985–1989, 1990–1994, 1995–1999, 2000–2004, and for Breslow thickness classes (r 1, 1.01–2.00, > 2 mm). The annual percent change (APC) of the standardized rates was computed. Thickness was evaluated on the basis of sex, age, morphology type, site and period of time. Median thickness was evaluated by means of a nonparametric K-sample test. The incidence rate of melanoma rose significantly for both invasive (APC = + 5.1%) and *in situ* lesions (APC = + 11.1). The sex distribution of patients with invasive melanoma did not change over time (mean male/female ratio 0.95). The mean age at diagnosis did not change (57.2 years; SD = 17.2 years). From 1985–1989 to 2000–2004 the median value of thickness decreased from 1.68 to 0.8 mm ($P < 0.001$). Within the Breslow categories the median value of thickness decreased significantly for thin melanomas

Introduction

The incidence of melanoma has notably increased over the last decades in all western countries (Linos 2009, De Vries 2004) and also in Italy (Crocetti 2004). A major cause for this increase has been early diagnosis. In addition, most of the diagnosed lesions are now *in situ* (Criscione 2009) or thin melanomas (Garbe 2009, Baumert 2009, Lipsker 1999), especially of the superficial spreading melanoma type (SSM) (Linos 2009, Warycha 2008); therefore, their prognosis is very good.

In contrast, the incidence of thick melanomas has not decreased (Criscione 2009, Lipsker 1999, Richardson 2008, Crocetti 2003, Tejera-Vaquerizo 2008) or increased (Linos 2009).

Thickness is the most relevant prognostic factor. Moreover, a great proportion of thick melanomas is of the nodular type (Tejera-Vaquerizo 2008, Galler 2009, Crocetti 2006). Against the background of stable or even rising rates for thick melanomas, melanoma mortality has not significantly decreased (Linos 2009).

In recent years, health-care professionals have focused on how to cope with thick deadly melanomas, but the problem is still unsolved (Murray 2005).

The aim of this paper is to evaluate the incidence trend of cutaneous melanomas with special reference to thickness in an Italian population-based case series

Materials and methods

We retrieved all the invasive (n = 2862) and *in situ* (n = 605) cutaneous melanoma incident cases diagnosed in 1985–2004 from the archives of the Tuscany Cancer Registry (RTT). RTT is a population-based cancer registry active in the provinces of Florence and Prato (approximately 1161000 residents in the 2001 census), central Italy (Paci 2007).

We arranged Breslow thickness in three classes, thin (<= 1 mm), intermediate (1.01–2.00), and thick (> 2 mm), as indicated by Balch *et alii* (Balch 2009). As for thick tumours, we also analysed thick (>= 3 mm or >= 4 mm) lesions.

We used median to measure thickness because of the skewness of its distribution.

Thickness categories were evaluated according to the following variables:

(1) Sex

(2) Age (0–49, 50 + years)

(3) Morphological type (the International Classification of Diseases for Oncology morphology codes):
SSM(ICDO-M = 8743), nodular melanoma (NM; ICDO-M = 8741), lentigo maligna melanoma (LMM, ICDO-M = 8742), not otherwise specified (NOS, ICDO-M = 8720), 'other types' (ICDO-M = 8722, 8730, 8740, 8744, 8770, 8771, 8772).

(4) Site: head and neck, trunk, upper limb, lower limb, others and unspecified

(5) Period of time: 1985–1989, 1990–1994, 1995–1999,2000–2004.

Statistical methods

Incidence rates were age-standardised through the direct method using the European standard population.

We computed the annual percent change (APC) of the standardised rates using the weighted least squares method, on the basis of single incidence year. We performed the above-mentioned computations by means of the SEER*Stat 6.3.6 software (www.seer.cancer.gov/seerstat). We compared proportions by means of the X^2 test, or the Fisher's exact test, when we expected less than five observations.

As regards the median, we used the Stata command 'median' that performs a nonparametric K-sample test that evaluates the null hypothesis that those samples were drawn from populations with the same median (www. stata.com). In the case of two samples, the X^2 statistic test is calculated with and without a continuity correction.

Results

From 1985 to 2004, 3467 patients residing in the RTT area had a diagnosis of malignant melanoma (2862 invasive and 605 *in situ*). The standardised incidence rate of invasive malignant melanoma rose from 6.4 per 100 000 in 1985–1989 to 13.6 in 2000–2004, with a mean annual pace of + 5.0% [95% confidence intervals (CIs), + 4.0/ + 8.2], Table 8. The growth of incidence was statistically significant for both men (APC = +5.3, 95% CI, +4.2/+6.5) and women (APC =+4.9, 95% CI, +3.5/+6.3).

In the analysed period the sex ratio was stable, with a slight predominance of women. The median age at diagnosis did not change over time, being 58.1 years (range 21.2–100.6 years).

Table 8 (n.1 in the original numeration): Tuscany Cancer Registry: invasive melanoma, absolute numbers, proportion of females, standardised (European population) incidence rates, annual percent change (APC) of standardised rates. *In situ* melanoma, absolute numbers, proportion of females, standardised (European population) incidence rates. Annual percent change of standardised rates (APC) are computed on single years of diagnosis. Probability (P) for APC to be equal to 0 or for proportions that each sample has the same proportion of observations. LM, lentigo melanoma; N.o.s., not otherwise specified; NM, nodular melanoma; SSM, superficial spreading melanoma.

	1985-1989	1990-1994	1995-1999	2000-2004	n/p
Invasive melanoma (n)	442	565	835	1020	2862
% females	54.5	54.0	55.3	50.2	P=0.13
Incidence rate	6.4	8.0	11.4	13.6	P<0.01 for APC
Breslow thickness					
>0 to <= 1mm, n (%)	92 (20.8)	182 (32.2)	375 (44.9)	476 (46.7)	1125
Incidence rate	1.4	2.7	5.5	6.7	P<0.001 for APC
1.01-2.00 mm n (%)	76 (17.8)	93 (16.5)	112 (13.4)	129 (12.8)	410
Incidence rate	1.1	1.3	1.5	1.9	P=0.006 for APC
>2mm n (%)	117 (26.5)	134 (23.7)	153 (18.3)	200 (19.6)	604
Incidence rate	1.6	1.8	1.8	2.2	P=0.016 for APC
Unknown, n (%)	157 (35.5)	156 (27.6)	194 (23.4)	215 (21.0)	723
Incidence rate	2.2	2.2	2.5	2.8	P=0.07 for APC
In situ melanoma N	55	72	182	296	605
% females	61.8	63.9	52.2	51.4	P=0.15
Incidence rate	0.8	1.0	2.7	4.0	P<0.01 for APC

Morphology type					
SSM	2.9	4.5	7.3	9.0	P<0.01 for APC
NM	0.8	0.8	0.8	0.8	P=0.96 for APC
LM	0.2	0.2	0.2	0.2	P=0.14 for APC
Other	0.3	0.3	0.4	0.4	P=0.15 for APC
N.o.s.	2.2	2.2	2.7	3.2	P=0.006 for APC

With regard to invasive melanomas, a statistically significant growing incidence was detected for thin (APC = +9.5; 95% CI, +7.1/+11.9), for intermediate (APC = + 3.1; 95% CI, +1.1/+5.2) and for thick melanomas (APC = + 2.1; 95% CI, +0.4/+3.7), and it was almost statistically significant for those melanomas without the information on Breslow's thickness (APC = +1.9; 95% CI, –0.3/+4.1; Table 8 (n.1 in the original numeration).

Six hundred and five *in situ* melanomas were diagnosed with a statistically significant growing trend (APC = +11.1; 95% CI, +8.1/+14.3); standardised incidence rates for *in situ* melanomas rose from 0.8 per 100000 in 1985–1989 to 4.0 in 2000–2004. The incidence trend for *in situ* melanomas was statistically significant for both men (APC = +12.7; 95% CI, +8.7/+16.9) and women (APC = + 9.4; 95% CI, +6.1/+12.8). *In situ* invasive melanomas did not show any statistically significant change, either in sex ratio (percentage of women 54.1) or in the median age at diagnosis (57.8 years).

The mean age at diagnosis for *in situ* melanomas was lower than for invasive melanomas (55.5 vs. 57.2 years; P = 0.02). However, among invasive melanomas (<= 1 mm) the mean age at diagnosis (52.7 years) was lower than for *in situ* melanomas (P=0.001), whereas the mean age of 60.2 years for melanomas (>1 mm) was higher than for both *in situ* melanomas (P<0.001) and those with a thickness of 1mm or less (P=0.001).

According to types of morphology, the increasing incidence trend was statistically significant for SSMs (n= 1655, APC= +7.0; 95% CI, +5.2/+8.8) and for the group 'other types' (n=155, APC= +6.7%; 95% CI, +2.6/+11.0), whereas for NMs (n=238), LMM (n=79) and for NOS melanomas (n=775), the trends did not reach any statistical significance, Table 8 (n. 1 in the original numeration).

From 1985–1989 to 2000–2004 the median value of thickness, for invasive melanomas, decreased from 1.68 to 0.8 mm (P<0.001). The median thickness decreased statistically significantly for both men (from 2.1 to 0.8 mm) and women (from 1.3 to 0.75 mm); Fig. 7 (n.1 in the original numeration).

The decrease over time in median thickness was not present in all Breslow categories. In fact, the median value of thickness decreased statistically significantly for thin melanomas (<=1 mm) but not for intermediate (1.01–2.00mm) or thick melanomas (>2 mm) Table 9 (n.2 in the original numeration).

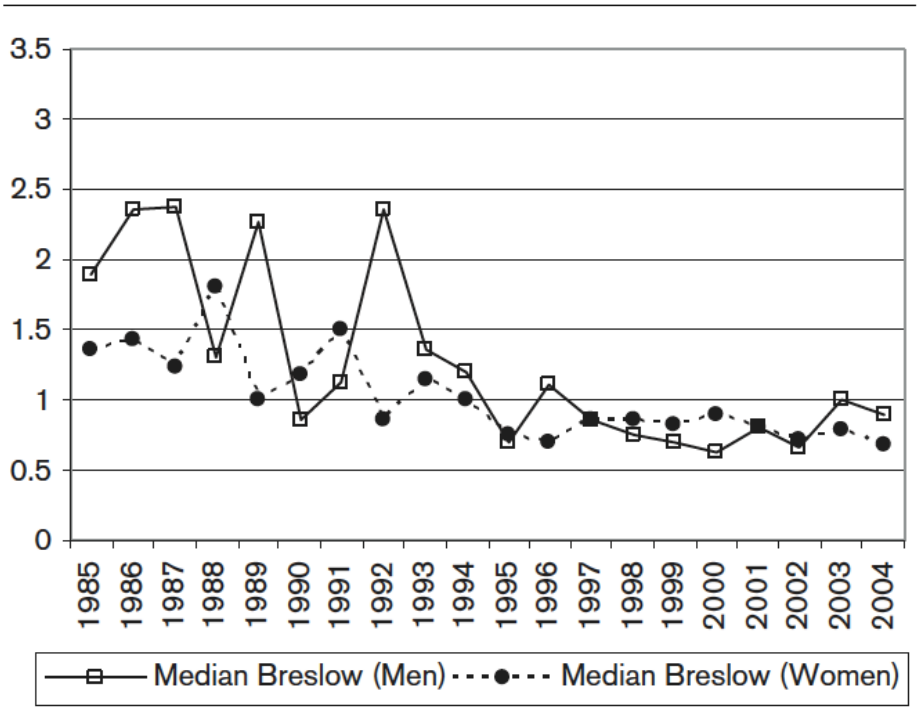


Figure 7 (n.1 in the original numeration): Tuscany Cancer Registry. Invasive melanoma: median Breslow thickness for men and women by calendar years.

Median thickness did not decrease among thick melanomas in other Breslow categories either (>=3 mm, >=4mm; (results added in Appendix 2 [Supplementary information for Paper 3]; 'data not shown' in the original paper).

The decrease in median thickness was statistically significant only for SSM (from 1.20 to 0.68 mm) and for other types (from 2.75 to 1.10), whereas it did not reach statistical significance for the other morphological types of melanoma (Table 9 [n. 2 in the original numeration]). In the most recent period (2000–2004), SSM represented the largest proportion of melanomas (62.6%; 639/1020), followed by melanomas NOS (17.5%; n=179), the group 'other' (10.7%; n=109), nodular melanomas, (6.2%; n = 62) and LMMs (2.9%; n = 30). Among the 208 thick melanomas diagnosed in 2000–2004, 45.7% were SSM and 24.5% of nodular type.

Table 9 (n.2 in the original numeration): Tuscany cancer registry. Invasive melanoma: median thickness by period of diagnosis for males and females, for Breslow's thickness categories, for morphology type (LM, lentigo melanoma; N.o.s., not otherwise specified; NM, nodular melanoma; SSM, superficial spreading melanoma), and site. P shows the probability that the medians in different groups are medians of samples drawn from the same population.

	1985-1989	1990-1994	1995-1999	2000-2004	p
Overall	1.68	1.2	0.8	0.8	<0.001
Males	2.1	1.3	0.8	0.8	<0.001
Females	1.3	1.15	0.8	0.75	<0.001
Breslow thickness					
<=1mm	0.66)	0.57	0.55	0.50	0.001
1.01-2.00	1.50	1.45	1.40	1.30	0.239
>2mm	3.4	3.5	3.7	3.8	0.418
Morphology type					
SSM	1.20	0.85	0.70	0.68	<0.001
NM	3.33	3.39	3.30	4.00	0.517
LM	1.80	0.81	0.67	0.79	0.398
Other	2.75	3.20	1.10	1.10	0.003
N.o.s.	2.25	2.35	1.79	1.80	0.675

With regard to age, the increase in incidence was present and statistically significant for both younger (0–49 years) men (APC = +5.0, 95% CI, +2.7;+7.2) and women (APC = +6.2, 95% CI, +4.0;+ 8.5) and for older (50 + years) men (APC = +5.6, 95% CI, +4.1;+7.0) and women (APC = +3.5, 95% CI, +2.2;+4.8).

With regard to skin sub-sites, by age group and sex (Table 10 [n. 3 in the original numeration]), melanomas of the head and neck did not show any statistically significant change either in incidence or in Breslow thickness in both sexes and age groups. The incidence of melanoma of the trunk increased significantly in both sexes and age groups, and the median thickness showed a statistically significant change towards a decrease, with the exception of younger women. The incidence of melanomas of the upper limbs increased in all age groups and both sexes, whereas their median thickness decreased statistically significantly only among younger subjects. With regard to melanomas of the lower limbs, there was a significant increase in incidence and a statistically significant decrease in median thickness among women. The group of melanomas of NOS sites increased over time, whereas their median thickness was stable.

Table 10 (n.3 in the original numeration): Melanoma, number of cases according to sex, age (0–49, 50 + years and all ages), annual percent change (APC) of standardised incidence rates 1985–2004 with corresponding 95% Confidence Intervals (95% CI), median thickness during 2000–2004, probability (P) that the median thicknesses for the periods 1985–1989, 1990–1994, 1995–1999 and 2000–2004 belong to populations with the same medians according to a non-parametric k-sample test.
N.o.s., not otherwise specified.

	Males			Females		
	<=49 years	50+ years	All ages	<=49 years	50+ years	All ages
HEAD&NECK						
Number	32	132	164	28	125	153
APC 85-04	-1.3	+2.6	+0.4	-1.9	+0.3	-0.3
95% CI	-5.8+1.4	-1.6+4.9	-2.6+3.5	-5.3+1.7	-3.5+4.3	-3.8+3.4
Thickness 00-04	0.55	2.7	2.3	0.7	1.19	0.9
p thickness change 85-04	0.10	0.64	0.54	0.20	0.17	
TRUNK						

Number	179	472	651	200	184	384
APC 85-04	+4.7	+4.8	+4.8	+6.9	+4.5	+6.1
95% CI	+1.3+8.2	+2.7+2.1	+2.9+6.7	+3.1+10.8	+1.1+8.1	+2.8+9.5
Thickness 00-04	0.86	0.7	0.8	0.7	0.7	0.7
p thickness change 85-04	0.005	<0.001	<0.001	0.69	0.01	0.02
UPPER LIMB						
Number	54	102	156	79	128	207
APC 85-04	+4.7	+7.5	+7.6	+5.7	+5.1	+5.3
95% CI	+0.9+8.3	+4.0+11.2	+4.7+10.5	+1.2+10.4	+1.8+8.7	+2.7+8.0
Thickness 00-04	0.54	0.85	0.65	0.8	0.9	0.88
p thickness change 85-04	0.054	0.20	0.03	0.07	0.60	0.12
LOWER LIMB						
Number	71	132	203	204	423	627
APC 85-04	+8.3	+4.4	+5.3	+4.3	+1.8	+3.0
95% CI	+3.8+13.0	+0.6+8.3	+2.5+8.1	+1.2+7.5	0+3.6	+1.1+4.9
Thickness 00-04	0.8	1.8	1.1	0.71	1.3	1.1
p thickness change 85-04	0.10	0.40	0.09	0.06	0.04	0.06
N.O.S.						
Number	58	110	168	56	93	149
APC 85-04	+7.4	+8.9	+7.9	+12.2	+11.1	+14.2
95% CI	+2.0+13.0	+3.7+14.3	+3.9+12.1	+3.8+21.8	+5.4+17.2	+7.6+21.1
Thickness 00-04	0.53	0.5	0.5	0.52	0.72	0.52
p thickness change 85-04	0.51	0.23	0.16	0.47	0.68	0.56

Discussion

The increasing incidence of invasive melanomas observed from 1985 to 2004 in central Italy was mainly supported by increasingly thinner lesions, especially of the SSM type. Therefore, the overall median thickness of melanomas has decreased over time, being in recent years 0.8 mm. This result was supported by the decrease in the median thickness of thinner melanomas (<= 1 mm; 0.5 mm during 2000–2004).

However, the shift in thickness observed for thin melanomas does not affect thick melanomas, which in Italy are now as thick as they were in the past. In fact, the median thickness of intermediate (1.01–2.00 mm) and thick melanomas (> 2 mm) did not decrease over the analysed period.

Moreover, although the proportion of intermediate and thick melanomas has reduced over time, their incidence rates show a statistically significant increasing trend (Linos 2009). In addition, their absolute number has increased (Lipsker 1999, Murray 2005). In agreement with the literature, we observed a strong increasing trend for *in situ* melanomas in central Italy (Criscione 2009), and almost half of the invasive melanomas are now thin. Such lesions have a very good prognosis, in so much as the UK recently proposed a change in guideline recommendations for those < 0.5 thick, suggesting less frequent follow-up (Einwachter-Thompson 2008). The mean age at diagnosis was lower for <= 1 mm melanomas than from *in situ* melanomas. This would indicate that *in situ* melanomas are not a precursor lesion of melanoma but have a different pathway than invasive melanomas.

As observed in other reports, we did not detect any decrease in the thickness of NMs at diagnosis (Lipsker 1999). We documented that in this population-based series a significant percentage (24.5%) of thick melanomas are NMs, reaching around 30% when melanomas NOS are excluded (Blach 2009).

There was an increase in incidence for all skin sites, with the exception of head and neck, for which the rates were stable over time. The trends for skin sites were similar for younger (0–49 years) and older (50+ years) patients of both sexes. A statistically significant change (decrease) in median thickness was present for melanomas of the trunk in men and older women, for upper limbs only in younger patients, and for lower limbs only in women.

The epidemiology of melanoma is further divided into different groups of lesions (Lipsker 2007). A first group includes those melanomas easily detectable by enhanced early diagnosis, increasingly thinner, mainly of a SSM type, with a very good prognosis

and presumably with some amount of over-diagnosis (Welch 2005). There is a second group of melanomas, more aggressive, thick at diagnosis, and often of a nodular type (Betti 2005). The relationship between these two types of lesions is still unknown. Therefore, it is crucial to further investigate the biological history of thick melanomas References The original references of the paper have been included in the general list of the thesis.

4.1.1. Comment

The previous paper (Paper 3) highlights a crucial step to make the activity of a CR more fruitful and useful to the healthcare system and the entire community: including the collaboration with clinicians.
This collaboration represents a win-win strategy, being beneficial to both clinicians and CRs, as well as for all the other stakeholders (e.g. patients).

The main advantage for the CR is having contact with the Real World (Aguiar, 2018; Crocetti, 2019). However, it is necessary to collect all the specific oncological variables that clinicians use in order to define the treatment and the prognosis of patients (Costa, 2019).
For their part, the CRs offer a population point of view. Therefore, the advantage to physicians is having access to theoretically complete and unselected datasets for a particular type of patient population (for example, by age, stage, therapy) (Verdasca, 2018).
In addition, CRs usually have more experience in data analysis and statistical methods which could be useful to clinicians.

CRs need to take advantage of this type of collaboration with the numerous professionals involved in cancer diagnosis and treatment. They should follow the exemplary contributions of many experts from different fields: screening experts (Vicentini, 2019), molecular and cellular biochemists (He, 2019), radiologists (Lehnich, 2019), pathologists (Roscher, 2018), surgeons (Bergvall, 2019), oncologists (Zijlstra, 2019), radiotherapists (Mullins, 2019), and other specialists.

Collaboration implies reciprocal knowledge and if fruitful, the appreciation of specific abilities. The use of CR's data by clinicians increases the visibility of CR in the scientific community and encourages the development of further collaborations.

Collaborations are indispensable for the improvement of CRs and the fulfilment of their potentiality. Working together usually results in gathering more clinical variables for producing epidemiological indexes customised to the needs of clinicians.

If the CRs are able to increase how informativeness of their data, this in turn will improve their usefulness as well as generate more interest among doctors and the other parties involved.

5. Time trend and projections

Incidence trends are essential when they are based both on frequency of cases (the real burden for the health system) and on age-standardised rates (the real risk of developing cancer) (Boyle, 1991), both of which are presented in Paper 3 to show the growth of the burden and the trend of incidence for skin melanoma.

In its most straightforward representation cancer trends are a plot of age-adjusted incidence rates for the following years or periods.

According to the numbers on which rates are computed, fluctuations due to random variability may make interpretation difficult. Therefore, observed data may be included in a model and presented as an estimated line. This computation is the base for log-linear models which express the slope of the time trend of cancer incidence as the beta coefficient of the time variables (Estève, 1994). In this way, it is possible to offer a summary measure of the trend that is easily interpretable as the annual average percent change of age-adjusted incidence rates (APC, as used in paper 3) with a statistical evaluation of its difference from a flat trend.

However, a limit for this approach may be the over-simplification of complex tendencies (e.g., sequences of increases and decreases) which, on the contrary, may be relevant to detect.

The purpose of the so-called 'joinpoint analysis' is to identify, within a defined period, those points in time when the tendency changes between periods having different slopes (Kim, 2000). Join-point analysis is widely used by registries also for the availability of the free software provided by the US NCI within the SEER programme web-site (https://surveillance.cancer.gov/joinpoint/).
Independently from the methods described until now, it is noteworthy to remember that a reasonable interpretation of age-standardised rates is possible only when the trends by age-groups for successive cohorts of birth or periods are parallel on a log scale (Estève, 1994).

This nudges us that time trends are not as straightforward to interpret as a unique line in a graph could enable one imagine.

IARC also presents cancer incidence trends. The extensive series of data published in the volumes of Cancer Incidence in 5 Continents are available in CI5plus (http://ci5.iarc.fr/CI5plus/Pages/online.aspx). In CI5plus data may be analysed as: time trends by period, time trends by age, time trend by birth cohort. This alternative way of analysing trends clarifies that time trends represent the combined effects of numerous variables: age, period and cohort (Clayton & Schiffers, 1987a,b).

• Age (cancer incidence typically increases with age). It may express an accumulation of carcinogenic effects and exposure or a weakening of the immune system or both. The risk of epithelial cancers, which comprise 90% of all cancers worldwide, increases approximately as the fifth power of age (Armitage & Doll, 1954).

• Period (when there are changes in risk in all individuals at the same point in time, regardless of age). It is 'cross-sectional' as it involves all ages and cohorts within a given period, e.g., a change in detection practice.

• Cohort (when there are changes in risk in successive generations of birth cohort). It is longitudinal.

Therefore a specific age-period-cohort analysis which includes all these components is necessary for an appropriate interpretation of time trends in general (Clayton&Schiffers, 1987a,b) and also for cancer (Rosenberg, 2011).

A broader use of such models in the interpretation of cancer trends was limited by the so-called 'drift' or the identifiability problem. Due to the relationship among the three terms (the calendar year is the sum of age and year of birth) the components of this model are collinear, and therefore the model cannot be uniquely determined (Clayton&Schiffers, 1987a,b).

Different methodological options may be implemented to overcome the identifiability problem: exclude one of the three components of trend, define constraints, compute estimations of second-order differences (Clayton&Schiffers, 1987a,b) or use the principle components approach (Tu, 2011). The parametrisation of the age-period-cohort model is not straightforward either for the factors (Holford, 2006; Carstensen, 2007) or if we choose to smooth with parametric functions, with splines, fractional polynomial or restricted cubic splines (Cartstensen, 2007).

Coping with identifiability is not easy for registrars. However, an efficient way to define constraints is based on the epidemiology/biology of the disease. As a consequence, the more constraints are epidemiology-based, the more results are interpretable by non-statisticians (and the model is simple.) For example, Vaccarella et al., having observed that the incidence of cervical cancer was broadly constant after the age of 45, solved the identifiability problem by constraining the incidence rate to be the same at age 45-69 and 65-69 years (Vaccarella, 2013). In another example seeking to analyse thyroid cancer trends in Italy, the cohort effect function was set to 1 on average with 0 slope (Dal Maso, 2011).

Age-period cohort models allow to evaluate the different components underlining cancer time trends. The availability of methods for providing more accurate time trend estimates increases over time, together with their complexity.

Moreover, changes in incidence are not usually unexpected or sudden (like those due to the introduction of screening programmes). Therefore, time trends based on high-quality data support reliable forecasting of the future short-term burden.

In the following Paper (4) the age-period-cohort model is applied for trend analysis, and time projections are additionally provided.

5.1. Paper 4

The paper has been published in Eur J Cancer Prev. 2007;16: 50-4 with the following title:

"Melanoma incidence in central Italy will go on increasing also in the next future: A registry-based, age–period–cohort analysis".

Authors: Emanuele Crocetti[a], Paola Carli[b], Guido Miccinesi[a]
Affiliation: a) Clinical and Descriptive Epidemiology Unit, CSPO, Florence, Italy, b) Dermatology Department, University of Florence, Italy.
Author Contributions: I (EC) declare to have conceived the idea of the study, planned and designed it, performed the analysis and drafted the first draft. GB gave a substantial contribution to the statistical analysis. The other Authors revised critically the paper and approved the final version of the manuscript

Abstract

The aim of the study was to evaluate malignant melanoma incident trend in central Italy by means of an age–period–cohort approach. A total of 1977 malignant melanoma (15–84 years) incidents in the area of the Tuscany Cancer Registry between 1987 and 2001 were analysed. Poisson regression has been used to estimate age, cohort and period effect. A nonlinear regression model was used to estimate the expected number of new cases in the period 2002–2006. Incidence rates increased in all age, period and cohort groups. The model that best fitted the data included age and 'drift'. The linear effect ('drift') showed, in each age group, an increase of the risk of malignant melanoma diagnosis of about 36.6% every 5 years of period or cohort. For the period 2002–2006, 1112 new cases were predicted with a standardised rate (age 15–84 years) of 19.2 × 100.000. In the Tuscany Cancer Registry area, no clues for malignant melanoma incidence rates levelling off were documented. Growing rates and number of malignant melanoma are expected in the next future.

Introduction

Malignant melanoma (MM) incidence has increased over the last decades in almost all Western countries (Lens and Dawes, 2004). In central Italy also incidence trends showed a significant increasing trend in both men and women (Crocetti and Carli, 2003). More recently, levelling of or even decreasing trends in incidence have been identified among young participants in Denmark, in the Netherlands, in Switzerland and in the United Kingdom (De Vries et al., 2003). A decrease in incidence among young cohorts predicts a future overall decline in rates as soon as these participants will contribute to older age groups.

The aim of the present paper was to explore incidence trends for cutaneous MM in central Italy focusing on the age, period and cohort effect.

Materials and methods

Incidence data were retrieved from the Tuscany Cancer Registry (RTT), a population-based cancer registry active in the provinces of Florence and Prato (about 1160000 residents in the 2001 census), central Italy, since 1985. The description of the criteria for collection, and registration followed by the Registry, has been presented elsewhere (Paci 2002).

During 1987–2001, 2071 incident cutaneous melanomas were registered in the RTT; in the present analysis, we selected age range 15–84 years and 1977 incident cases were included.

Number of cases and person-years were aggregated in 5- year age groups (from 15–19 to 80–84 years), 5-year periods (1987–1991, 1992–1997, 1998–2001). Five-year cohorts of birth were computed according to age and period. In the figure, cohorts are labelled according to the central year of the cohort of birth.

The incidence rates were standardized using direct method with the European standard population as the reference.

Poisson regression has been used to estimate age, cohort and period effect. We may consider cohort effects as influences that affect rates in a specified generation or birth cohort throughout life, whereas period effects affect rates equally across all age

groups at a specified period (Clayton and Schifflers, 1987a). When there is a regular temporal trend, which cannot be ascribed to either period or cohort influences, it is called 'drift'. Only when we observe irregular or sudden changes, we can ascribe the observed temporal trend to either period or cohort influences. The goodness-of-fit of different models was assessed by the deviance. The closer the deviance with the degree of freedom, the better the fit of the model. Differences between deviances allowed us to compare nested models (e.g. age alone vs. age + cohort) (Clayton and Schifflers, 1987b).

The expected number of new cases for the years 2002– 2006 was estimated according to a nonlinear regression model, proposed by Dyba et al., (Dyba 1997), including the effect of age and age-specific temporal trends. Friendly STATA macros for the application of such method for short-term prediction are available on the Web site of the European Network of Cancer Registries (http:// www.encr.com.fr/stata-macros.htm). This model assumes that the absolute change in incidence rate over time for a given age group is often larger when the baseline rate is larger. The age-specific absolute change of incidence is proportional to the corresponding age-specific baseline rate, whereas for a given time period the relative change in incidence is the same for all age groups. Owing to these characteristics this model preserves in the period of prediction the age pattern of incidence existing in the data, and the age-specific predicted rates are substantially more precise than those for a linear model (Dyba, 1997).

Several methods have been used for making predictions of the future cancer burden; the one used in the present analysis showed good performance when compared with other methods (Moller, 2003).

We used the annual sex-specific and 5-year age-specific population for the years 2002–2006 based on the official data and future estimates of the Regional Office of the National Institute of Statistics (years 2002–2003, 2008) and the estimates based on the Waring method for 2004– 2006 (Shryock, 1976).

Total number of expected cases and prediction intervals are presented. Predictions of incidence for each age group are also presented.

Results

Table 11 (n.1 in the original numeration), shows age-specific and age-standardised melanoma incidence rates from the Tuscany Cancer Registry area for three periods (1987–1991, 1992–1996 and 1997–2001).

Table 11 (n.1 in the original numeration): Malignant cutaneous melanoma. Number of incident cases, age-specific (males and females) and age- standardised (European population) incidence rates in the Tuscany Cancer Registry, according to time period. [a] Standardized rates are computed in the population 15–84 years.

		Period	
	1987-1991	1992-1996	1997-2001
Age (years)			
15-19	0.8	1	2.4
20-24	2	5.7	6.1
25-29	3.9	7	10.2
30-34	5.7	7.6	13
35-39	5.1	8.4	12.9
40-44	8.9	11.1	15.2
45-49	9.5	13.1	18.4
50-54	10.3	16.3	23.1
55-59	14.5	17.9	26
60-64	14.8	19.1	19.4
65-69	16.2	18	23.5
70-74	16.2	19.4	25.1
75-79	15	25.9	31.1
80-84	14.2	19.7	28.9
n. of cases	449	636	892
Standardized rates[a]	8.3	11.6	15.9

Incidence rates increased in all age groups. The age- specific incident rates increased in participants born in more recent years (Figure 8 [n. 1 in the original numeration]), and they also increased from the first to the third analysed period (1987–1991, 1992–1996, 1997–2001) (Fig. 9 [n.2 in the original numeration]).

The model approach used to disentangle age, period and cohort effect showed that the model that best fitted the data was the one including age and 'drift' (deviance of the 'age + cohort' model – deviance of the 'age + drift' model = 15.7; number of parameters of the 'age + cohort' model – number of parameters of the 'age-drift' model = 14; P-value of 15.7 on a X^2 distribution with 14 degrees of freedom = 0.392; for the age + period model vs. age + drift model; P-value = 0.808).

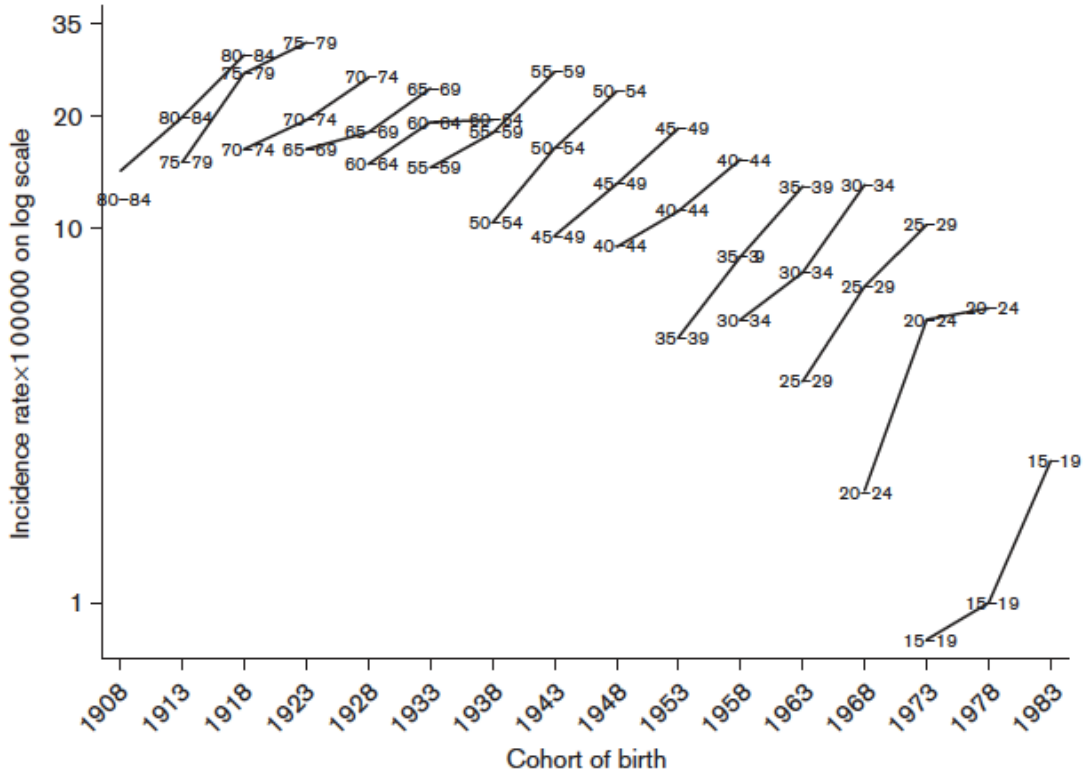


Figure 8 (n.1 in the original numeration): Tuscany Cancer Registry, malignant melanoma 1987–2001. Age-specific incident rates according to the cohort of birth, men+women.

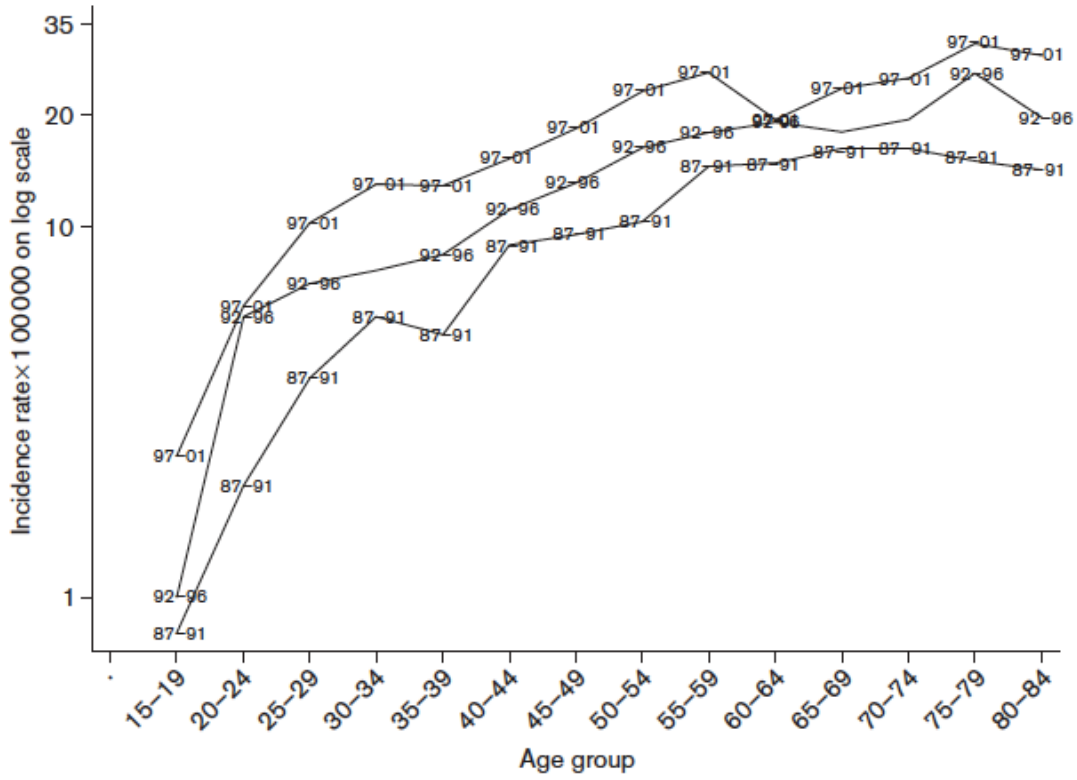


Figure 9 (n.2 in the original numeration): Tuscany Cancer Registry, malignant melanoma 1987–2001. Age-specific incident rates according to three different periods (1987–1991, 1992– 1996, 1997–2001), men+women.

The linear temporal effect ('drift') showed that in each age-group, there was an increase of the risk of MM diagnosis of about 36.6% every 5 years of period or cohort (mean annual increase 6.4%). The model including age and drift resulted the best also for explaining incidence data for men and women, when analysed separately (results added in Appendix 2 [Supplementary information for Paper 3]; 'data not shown' in the original paper).

The predicted standardised incidence rate (age 15–84 years) for the period 2002–2006 was 19.2; the predicted number of new cases was 1090 (95% prediction intervals 984–1196).

Figure 10 (n.3 in the original numeration), reports age-specific incidence predictions for 2002–2006, together with their prediction intervals.

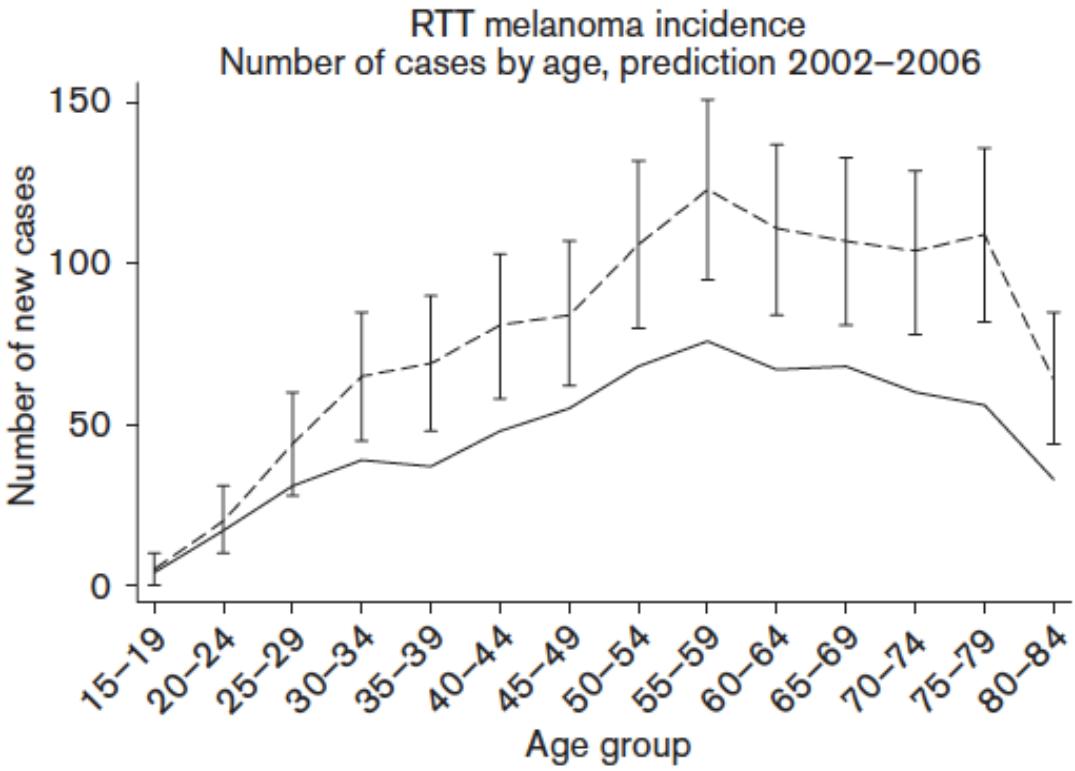


Figure 10 (n.3 in the original numeration): Melanoma of both sexes, Tuscany Cancer Registry area. Five-year mean number of observed incident cases in 1987–2001 (solid line) and incident cases predicted for 2002–2006 (dashed line) with approximate 95% prediction intervals; by age.

Discussion

The results of the present study indicate that MM incidence rates in the population covered by the Tuscany Cancer Registry, central Italy, are still increasing. This trend is present in all age groups, and also younger participants, both men and women, do not show any sign of levelling off.

The present figure contrasts with that recently found in some northern European populations. According to data extracted from the EUROCIM database 165 cancer registries, a deceleration in incidence trends occurred recently in Northern European countries among persons younger than 70 years; whereas in eastern and southern Europe incidence rates were still increasing (De Vries et al., 2003). Therefore, we are probably observing a shift between northern and southern European populations in terms of future scenarios about the melanoma 'epidemic'. Earlier detection and a

growing public awareness about risk associated with excessive sun exposure are the most plausible explanations for the deceleration found in northern Europe (De Vries et al., 2003).
Little is, however, known on the factors still holding up the increasing incidence trends in Mediterranean people.

In this study, models including the effects of age, period and birth cohort were used to adequately analyse the rising trends.

As known, cohort effect reflects exposures that affect rates in a specified cohort equally throughout life, whereas period effect affects rates equally across all age groups at a specific period (Clayton and Schifflers, 1987a). An age effect (increasing risk according to ageing) and a linear effect are causing the increasing incidence.
Owing to the linear dependence between the linear part of cohort and period (and age) however, the interpretation of regular incidence trend is not possible. Indeed, only when we observe irregular changes we must consider age-period or age-cohort models. On the contrary, this so- called 'drift' effect represents a situation equally well described by two models, age and period (linear) or age and cohort (linear) and cannot be ascribed to either period or cohort effect (Clayton and Schifflers, 1987a).
Although statistics does not allow to disentangle period from cohort effect, some suggestions may come from the epidemiological knowledge.

The major environmental risk factor for MM is the intermittent exposure to ultraviolet radiation (UV) (Elwood and Jopson, 1997), especially during childhood (Naldi 2000). Such exposure has become popular after the Second World War, and in Italy, particularly from the second-half of the '50s to the early '60s as a consequence of a rapid economic growth.

The present data are, however, not fully explained by the hypothesis of increasing risk for UV exposure during childhood only. In fact, the increase in incidence was evidenced in all birth cohorts, also in the oldest ones, in participants who during the 50ths–60ths were middle-aged. Our results are consistent with the hypothesis that UV

exposure increases the MM risk independently from the age of exposure (Elwood and Jopson, 1997). Recent data suggest that cutaneous melanomas may arise through two pathways, one associated with melanocyte proliferation and the other with chronic exposure to sunlight (Whiteman, 2003). Australian patients with head and neck melanomas – lentigo maligna melanoma excluded – compared with patients with melanomas of the trunk, were statistically significantly less likely to have more than 60 nevi [odds ratio (OR) = 0.34, 95% confidence interval (CI) = 0.15–0.79] but were statistically significantly more likely to have more than 20 solar keratoses (OR = 3.61, 95% CI = 1.42–9.17); moreover, they were more prone to a past history of excised solar skin lesions (OR = 1.87, 95% CI = 0.89–3.92) (Whiteman, 2003).This finding has been confirmed in Italian patients (Carli and Palli, 2003). This means that not only exposure in early life, generally intermittent, but also cumulative lifelong exposure may contribute to melanoma development. Recent data from Creta show that in the relatively dark-skinned population, sun exposure indices represent the most important risk markers for cutaneous melanoma, which contrast with data from fair-skinned Caucasian populations in which melanocytic naevi are the main risk factors (Lasithiotakis, 2004). Therefore, the change in lifestyle with increasing exposure to UV owing to the growing popularity of sunbathing and tanning seem to have affected all age-groups as a period effect presumably occurred during late '50s–early '60s.

On the other hand, more frequent excision of pigmented skin lesions may contribute to explain the increasing number of melanomas diagnosed overtime in the Tuscany Cancer Registry area by means of a period effect. In the Tuscany area, the awareness among population and the development of early diagnosis activity has increased over the last decades; a preventive campaign addressed both to family doctors (general practitioners) and the general population for the surveillance of pigmented skin lesions is active in the RTT area since the late 1980s (Carli, 2002). A Pigmented Lesion Clinic working in the Dermatology Department of University of Florence was also implemented for rapid referral of participants with self-detected or GP-detected suspicious lesion (Carli, 2002). A clue for the effect of early diagnosis was the growing rates of 'thin' lesions (<= 1 mm) that showed a mean annual increase of about 16.1%

from 1985 to 1997 (Crocetti and Carli, 2003). This situation probably explains the role, if any, of the period-effect in sustaining the incidence rates.

In conclusion, in the Tuscany Cancer Registry area MM rates will probably go on increasing in the next future. According to prediction model, the standardised incidence rate (age 15–84 years) for the period 2002–2006 will be 19.2 cases per 100000, approaching that observed in northern European populations. Although the increasing trend was explained by an age-drift model, the two major explaining factors – changes in lifestyle with increasing exposure to UV and increased early diagnosis seem to have acted more as period than as cohort effects.

References: The original references of the paper have been included in the general list of the thesis.

5.1.1. Comment

The previous paper (Paper 4) shows an example of how the CR's activity can be improved and its potentiality exploited.

The potentiality of the information included in CR's data is superior to that which is normally presented. For example, the data necessary for the analysis of skin melanoma cancer (MM) time trends using the method of age-period-cohort are already included in those used for traditional annual ASR plotting and provide further insights into trend interpretation.
Therefore, in this example, the limit to overcome is not the quality (looked at in Paper n. 1) or quantity (look at in Paper n.3) of the collected data, but it is the ability of CR's personnel to use the most appropriate statistical approach to better address a specific topic.

Moreover, high-quality data can also be used, as exemplified in the previous paper (Paper 4), to make reliable projections about future cancer burden. These types of estimates fulfil a dual objective. On one side, they fill the time gap between cancer diagnosis and the availability of cancer incidence rates. Furthermore, if projected in the near or far future, they help policymakers in their decision making regarding the distribution of the resources needed to satisfy a given population (Erdmann, 2013; Nguyen, 2019; Jung, 2019). Projections address specific age-groups, i.e. the elderly (Clèries, 2018; Pilleron, 2019), children (Ward, 2019), cancer sites (Araghi, 2018; Earnest, 2019; Lewis, 2018; Yu, 2019).
Projections have not only been applied to cancer incidence, but also to prevalent cases (Colonna, 2018; Guzzinati, 2018) and cancer survival (Gondos, 2009).

CRs need to use their data to the best of their potentiality, taking advantage of statistical methods.

Short-term projections should become the normal praxis for all CRs. Furthermore, reliable, high-quality data series' may be used as estimates of cancer burden (Bray, 2018) and on specific types of cancer and the stages of disease. (Crocetti, 2018).

A greater understanding of the data and knowing how to better use that which is available will make the CR more informative and more useful to the various stakeholders.

6. Survival estimates from cancer registry data

Cancer survival based on CR data represents a measure of the quality of the healthcare system. It refers to the average cancer patient in the population, provided all patients have the same access to a similar level of care (Donnelly, 2017). Survival is affected by both the time of diagnosis (e.g., availability of screening programmes), and the availability and accessibility of effective treatments.

For this reason, survival is among the essential measures provided by registries.

Survival has been particularly valued by the US National Cancer Institute, which in 1973 planned and financed the Surveillance, Epidemiology and End Results Program (SEER), which is still ongoing (Parkin, 2006).

Conceptually, survival represents the length of time between the onset of the disease (for registries the date of incidence) and the time an outcome (alive/dead) is verified. Death is part of human nature. Therefore, the interest in the investigation of cancer survival is to evaluate how cancer affects 'natural' survival (Parkin & Hakulinen 1991; Estève 1994).

Any survival analysis needs at least one specific extra variable: the status of life of the patients at a certain point in time since diagnosis, identified through high-quality follow-up of patients (Capocaccia, 2003).

Many alternative methods have been developed to measure survival in cancer patients. They measure the role of cancer in causing death (splitting those patients dying from cancer from those dying with cancer but from other unrelated causes). The fact that cancer patients also die for different reasons, just as cancer-free people, have made the computation and the comparison more complex (Parkin&Hakulinen, 1991; Estève, 1994).

Observed survival measures the real survival of patients who die from their cancer or other causes.

The aim of analysing survival in a group of cancer patients is to measure all-cause survival. The interest is evaluating whether this estimate is lower than what we would have observed if the same subjects were unaffected by the disease. If we consider two groups of cancer patients with diverse ages, survival will likely be lower for the more elderly ones due to the mortality for non-cancer causes. The same occurs when we compare survival in two populations of distinct areas, where death due to other reasons, may differ.

Therefore, observed survival is specific for the case-series on which it is computed and cannot be directly compared with others or even with the general population without cancer.

The probability of surviving is the result of two components: the effect of the disease under study and the overall impact of other conditions.

Net survival is the survival that would be observed if the risk of dying from all causes other than the one under study were eliminated (Estève, 1994).

Two methodological approaches allow us to estimate net survival: the cause-specific survival and the relative survival measurements.

In cause-specific survival, patients who die from cancer contribute to computation, while others do not (Parkin&Hakulinen, 1991).

The other approach, which takes into account the mortality from other causes, is relative survival. It is the ratio between observed survival and the survival expected in the age-sex-period-area- specific population. Its computation has changed with alternative suggested methods (Ederer I, II, Hakulinen) (Ederer 1961 and 1959; Hakulinen, 1982).

The effect of demographic characteristics, e.g., age, must also be considered and standardised to carry out reliable comparisons (Corazziari, 2004).

Not long ago, an alternative framework for the estimation of net survival has been proposed (Pohar Perme, 2006). The Pohar Perme method does not estimate relative survival (the marginal observed divided by the expected marginal survival.) It does, however, estimate net survival in a relative survival setting (rather than a cause-specific setting). Among the estimators in a relative survival setting, Pohar Perme represents the only unbiased estimator of net survival. It assumes the only possible cause of death is cancer, eliminating, in comparisons, the distortion which may arise from the chance of dying from other causes (Crocetti 2017b).

All these methods, and presumably also others, have been applied by registries to describe the experience of a group of cancer patients, by cancer topography, age, sex, time since diagnosis, etc.

The study of survival has been exploited by several big consortia which produce changes in survival over time and differences across countries (e.g., Concord, Eurocare).

The focus on the role in survival of cancer-specific variables has been underlined by the so-called high-resolution studies which compared results within and among European countries (Eurocare) and worldwide (Concord) (Sant, 2003; Coleman, 2008).

In such studies, the most frequently used approach is the Cox proportional model which takes into account the joint effect of numerous variables (Cox, 1972). This method is famous not only among registries but also among clinicians (Barraclough, 2011).

However, the Cox proportional model has a relative approach. For each analysed variable, the risk of dying is expressed in comparison with a reference level (risk put equal 1). Moreover, when Cox evaluates the independent effect of multiple variables, the overall risk is computable using the specific parameters of each variable, which requires some mathematical computation. The parameters of the Cox model and its

relative approach are unstraightforward to interpret for clinicians and especially for cancer patients.

For example, a multivariable Cox model needs some complex computations based on the coefficient of the model (exp (coefficient [e.g., sex] + coefficient [e.g., age] +coefficient [e.g., Breslow] + coefficient [e.g., Clark] + etc.)) in comparison with all the variables in the reference category.

A diverse approach might be to split cancer patients into subgroups in which people possess similar characteristics which affect the dependent variable of interest (e.g., survival). Belonging to a specific group: e.g., with a particular cancer, sex, age, stage, the same for all the people who exhibit these characteristics, corresponds to a specific probability of surviving at a certain time since diagnosis. There would not be coefficients to sum up, the approach would not be relative, but it would be a clear and straightforward answer to a vital question.

One of these methods is the classification and regression tree (CART) analysis, proposed by Breiman, as a decision tree methodology (Breiman, 1984 and 2001). The method, based on non-parametric statistics, if applied to a group of data, uses a binary recursive process to create a tree of subgroups. These subgroups are mutually exclusive and exhaustive. Each of them includes subgroups of patients (all with the same characteristics in each group) with different profiles for a given outcome. For example, more or less aggressive prostate cancer (Spurgeon, 2006), lower limb lymphedema (Spillane, 2010), melanoma histotype associated with CDKN2A genotype (Sargen, 2015), a serum biomarker panel which predicts nodal status in lung cancer patients (Borgia, 2009), and also differences in terms of survival, including melanoma (Garbe, 1995; Staut, 2010), just to mention some of the relatively few examples available in literature.

The topic of understandability of scientific information has to be considered for providing to cancer patients not only reliable but also useful information.

The following Paper (5) deals with the way cancer prognostic information are provided with alternative approaches: Cox model and CART analysis.

6.1. Paper 5

The paper has been published in Melanoma Res. 2006;15: 429-33 with the following title:

"Prognostic variables and prognostic groups for malignant melanoma The information from Cox and Classification And Regression Trees analysis: an Italian population-based study".

Authors: Crocetti E[a], Mangone L[b], Lo Scocco G[c], Carli P[d].

Affiliation: a) Tuscany Cancer Institute, Tuscany Cancer Registry, Clinical and Descriptive Epidemiology Unit, Centre for the Study and Cancer Prevention, b) Reggio Emilia Cancer Registry, Epidemiology Unit, Public Health Department, Reggio Emilia, c) Dermatology Unit, Hospital, Prato, Italy, d) Dermatology Department, University of Florence, Florence,

Author Contributions: I (EC) declare to have conceived the idea of the study, planned and designed it, performed the analysis and drafted the first draft. The other Authors revised critically the paper and approved the final version of the manuscript

Abstract

The common way to analyse the prognostic role of selected variables in cutaneous melanoma patients is by means of Cox proportional hazard model. The prognostic effect of the simultaneous presence of more than one independent variables in the same patients is, however, difficult to establish. This hampers the possibility of tailoring a survival expectance for a selected patient as well as to communicate it to the patient himself/herself. The objectives of the study were to compare information on cutaneous melanoma prognosis from multivariate Cox proportional hazard model and from Classification And Regression Trees analysis. Classification And Regression Trees analysis is an automatic method that splits data by means of a binary recursive process creating a 'tree' of groups with different profiles according to the analysed outcome, for example, the risk of death. This approach automatically produces data easy to be interpreted by clinicians. A total of 1403 invasive cutaneous melanoma, 1100 from the Tuscan Cancer Registry and 293 from the Reggio Emilia Cancer Registry,

Italy, were included. Cases were incident during 1996–2001 and followed up at the end of 2003. Cox proportional hazard model and Classification And Regression Trees analysis were applied to the following variables: age, sex, Breslow thickness, Clark level, Registry, sub-site and morphologic type. The Classification And Regression Trees analysis identified 10 categories with statistically different survival; this results were summarized into six classes of different risks based on Breslow thickness, age and sex. The best prognostic group (5-year observed survival, 98.1%) included those subjected with Breslow less than 0.94 mm and age 19–44 years. The same thickness but an older age (50–69 years) was associated with a statistically significant different prognosis (5-year observed survival, 92.8%). The Cox proportional hazard model found sex, age, Breslow thick- ness, Clark and morphologic type to have a significant independent prognostic value. In conclusion, compared with conventional approach based on Cox hazard model, Classification And Regression Trees analysis produces data closer to the clinical need of defining the prognostic profile of a specific patient. This may help the clinician both in the communication of risk and in the follow-up strategy.

Introduction

For malignant cutaneous melanoma (CM), the identification of prognostic factors has a crucial value for both dermatologists and patients. According to new evidence on effective prognostic factors, the CM staging system has changed during recent years (Balch 2000; Balch 2001; Black 2003; Black 2004).

Usually, more factors act together, for example, thickness, age, site, Clark's level of invasion, ulceration, sex, etc., (Balch, 2001), and therefore their cumulative effect should be evaluated by means of multivariate models. The use of multivariate Cox proportional hazard model for survival analysis is common in the scientific literature (Cox, 1972). For example, in the PubMed library, more than 900 papers matched a search for the terms 'Cox AND survival AND cancer', among those published during the year 2005 (overall 671163). By means of this statistical approach however, the clinical relevance of the simultaneous presence in a patient of more than one independent prognostic factor cannot be easily established. This is a pity, as the possibility for the clinician of tailoring a prognostic profile for each patient taking into

account the role of independent survival predictors simultaneously present would be clinically relevant both to plan follow-up and to communicate the personal risk to the patient him/herself.

During the last two decades, a tree-building technique has been slowly entering the literature. It is the so-called Classification and Regression Tree (CART) analysis (Breiman, 1984). This is an automatic method that splits data by means of a binary recursive process creating a 'tree' of groups with different profiles according to the analysed outcome, for example, the risk of death. Although this approach automatically produces data easy to be interpreted by clinicians, it does not seem to be very popular yet.

The aim of the present study was to compare malignant melanoma survival information from both Cox and CART approaches in a multicentric population-based Italian data set.

Materials and methods

In the present study, invasive CM incident in the period 1996–2001 in the area of the Tuscany Cancer Registry (RTT) and of the Reggio Emilia Cancer Registry (RECR) were included.

Both RTT and RECR are located in central Italy and are included in the Italian Network of Cancer Registries (AIRT, www.registri-tumori.it). Further details on their organisation and data management are available for RTT at http://www.cspoweb.it/rtt.asp and for RECR at http://www.ausl.re.it/Home/DocumentViewer.aspx?ID = 773&-TIPODOC = IAP.

Overall, 1403 malignant melanoma were analysed (1110 from RTT and 293 from RECR).

Follow-up has been carried out up to 31 December 2003 or death, whichever comes first. The mean follow-up time was 4.2 years, ranging from 0 to 8 years (median follow-up, 4.0 years).

The CART analysis was used to identify different prognostic groups. This is an automatic method that build up a 'tree' by means of a binary recursive partitioning of data. Full data are evaluated for all the possible binary splits. It means that the method evaluates the possibility of splitting each variable into two groups if there is a statistically significant difference according to the analysed outcome (e.g. risk of death). The process repeats building up a tree until all groups are unsplittable (Breiman, 1984).

The variables included in the CART analysis were as follows: sex (males, females), 5-year age classes, registry (RTT, RECR), Breslow's thickness (continuous), Clark level (2–5, missing), site (head and neck, trunk, upper arm, lower arm, not specified), morphology (superficial spreading – SSM, nodular, others, not otherwise specified). We chose not to analyse separately lentigo maligna melanoma owing to their small number, n= 31).
Information on lymph node involvement was not available.

CART uses the martingale residuals of a Cox model to calculate (approximate) w2 values for all the possible cut points on all the CART covariates. The significant value for cut point definition was 0.05. The minimum size of the group was defined as 30 participants.
Once CART had identified different groups, Kaplan and Meier survival curves were computed and compared by means of log-rank test (Mante 1966). Only statistically significant groups were considered for the definition of the final groups of patients with significantly different risks of dying.

Moreover, we analysed the prognostic role of the same variables also in a more conventional approach by means of the Cox conditional hazard model. Within each variable, the statistical significance of the difference between values has been evaluated; the original variable definition has been modified according to the comparison, in particular the site has been recoded as head and neck, other specified, not specified.

The variables that showed a significant effect in the univariable analysis were included in a multivariable Cox proportional hazard model to evaluate their independent effect. By means of a step-forward approach, the effect of each variable in improving the fit of the model was evaluated with the likelihood ratio test.
The product-limit estimate according to Kaplan and Meier was used for computing overall and variable- specific observed survival probabilities (Kaplan, 1958).
The analysis has been performed by Stata 8 (www.stata.com).

Results

The analysis was based on 1403 patients with newly diagnosed, incident CM; during the follow-up (mean time, 50.5 months) 343 of these patients died.

The overall observed survival was 93.1% at 1-year, 82.7% at 3-year and 74.7% at 5-year period.

In the CART analysis, 1154 patients with complete information for sex, age, Breslow's thickness, Clark's levels, registry, sub-site and morphology were included and all those variables were evaluated. The main relevant split involved Breslow's thickness identifying lesions thinner and thicker than 2.25 mm. The second split of the regression tree involved age classes. The third split was based on sex, and again on Breslow's thickness, as the fourth one (Fig. 11 [n.1 in the original numeration]).

The CART analysis identified 10 prognostic groups according to Breslow's thickness, sex and age. For each of these groups, Kaplan and Meier survival curves were computed and compared by means of log-rank test. When there was no statistical difference in survival between groups, they were summed. The result produced six groups of patients with significantly different risks of dying (Fig. 12 [n.2 in the original numeration]).

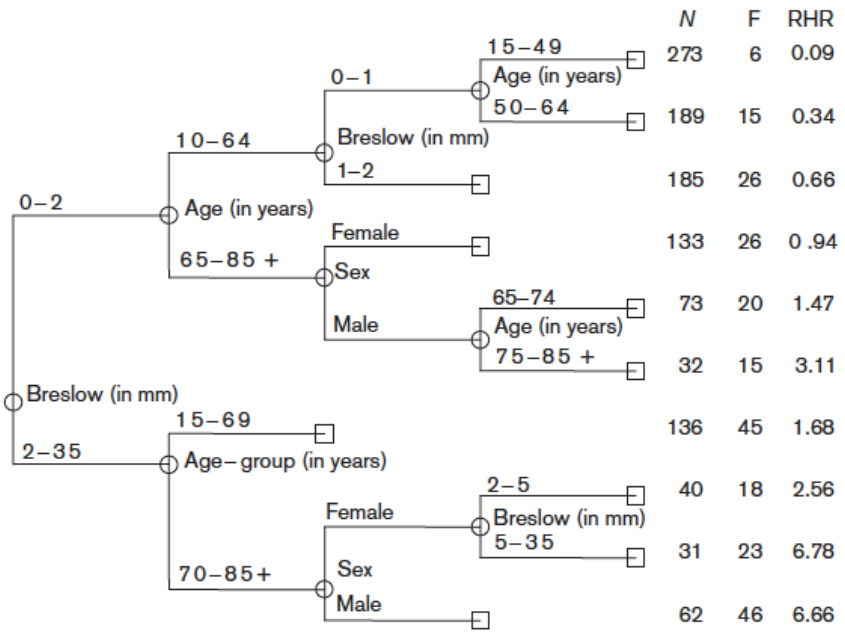


Figure 11 (n.1 in the original numeration): Classification And Regression Trees with the following variables: sex, registry, sub-site, age, Breslow, Clark and morphology. Split if (adjusted) P < 0.05. RHR, relative hazard ratio.

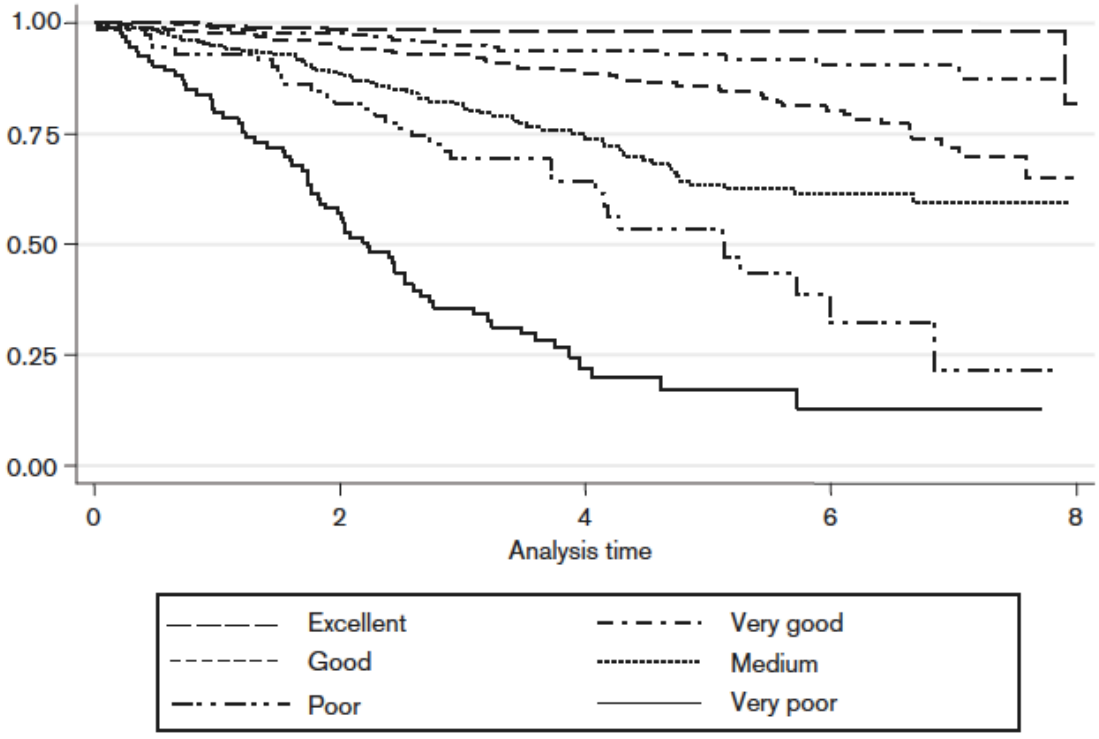


Figure 12 (n. 2 in the original numeration): Kaplan–Meier survival rates of the six prognostic groups defined by Classification And Regression Trees analysis (groups definition is given in Table 12).

In Table 12 (n.2 in the original numeration), the CART results are used to show a straightforward way to define these six prognostic groups according to Breslow, age class and sex. The best prognostic group (5-year observed survival, 98.1%) included male and female patients with Breslow's thickness less than 0.94 mm and aged 19–44 years. The same thickness but an older age (50–69 years) was associated with a statistically significant different prognosis (5-year observed survival 92.8%). The prognosis was good (85.9% at 5 years) for both men and women, with Breslow's thickness between 0.93 and 2.24 mm and age 10–64 years or for females with thickness less than 2.25 mm. The next group (5-year survival, 63.5%) included men aged 65–74 years with Breslow < 2.25 mm and both men and women aged 15–69 years with Breslow >= 2.25 mm. The prognosis was poor (5-year survival, 53.5%) for men with Breslow thickness < 2.25 mm and age 70 + years and for women with Breslow thickness 2.25–4.99 mm and age 70 + years. The group with the worst prognosis (5-year survival, 17.1%) included participants of 70 or more years of age, women with Breslow >= 5.00 mm, or men with Breslow >= 2.25 mm.

Table 12 (n.2 in the original numeration): Malignant melanoma, 1996–2001. Prognostic groups according to Classification And Regression Trees analysis. Each group is defined according to the characteristics of the included patients, for example, male and female patients with Breslow lower than 0.94 mm, 19–49 years old belong to the group with excellent prognosis.

		Prognostic factors			5-year observed survival
Group of prognosis	N.	Breslow mm	Age years	Sex	%
Excellent	273	<0.94	19-49	F/M	98.1
Very good	189	<0.94	50-64	F/M	92.8
Good	318	0.93-2.24 <2.25	10-64	F/M F	85.9
Medium	209	<2.25 >=2.25	65-74 15-69	M F/M	63.5
Poor	72	<2.25 2.25-4.99	75 70+	M F	53.5
Very poor	93	>=5.00 >=2.25	70+ 70+	F M	17.1

Data on traditional survival analysis based on univariate approach followed by multivariate Cox regression models are shown in Table 13 (n. 3 in the original numeration). The survival was significantly higher among women than among men, among younger patients (< 60 years) than among older ones, for melanoma with Breslow thickness less than 1 mm than for >= 1 mm, for smaller Clark's levels, for specified subsites (or not-specified ones) than for head and neck and for superficial spreading melanoma than for not specified ones. No statistically significant differences were evidenced between the two registries.

Discussion

The population-based series of CM analysed in this study resulted perfectly in line with that expected – according to the literature – in terms of overall 5-year survival (75%) and significance of major variables of risk of death. Indeed, the conventional Cox hazard analysis showed that Breslow thickness, sex, Clark level, morphology and age resulted in independent prognostic variables (Leiter, 2004; Garbe, 1995).

In this study, a CART analysis of survival was performed in order to easily investigate what happens when more than one independent prognostic variables are present in a patient. CART analysis enables clinicians in identifying subgroup of risk – characterised by the simultaneous presence of more than one variable – that differ significantly among them in terms of survival. Indeed, by means of this approach, the identification of different subgroups of risk is possible, with easier risk communication and, eventually, a more patient-tailored follow-up strategy.

Table 13 (n.3 in the original numeration): Malignant melanoma. Crude and adjusted risk of dying (hazard ratio - HR) for several prognostic variables with corresponding 95% confidence intervals (95% CI). The multivariate model includes sex, age, Breslow's thickness, Clark's level and morphology subtype. n.o.s., not otherwise modified.

Variable	N.	Crude HR	95%CI	Adjusted HR	95% CI
Sex					
Woman	762	1		1	
Men	644	1.40	1.14-1.73	1.30	1.05-1.61
Age (annual)					
<60	784	1		1	
60+	622	3.73	2.95-4.70	2.92	2.30-3.71
Registry					
Firenze	1113	1	0.73-1.24		
Reggio Emilia	293	0.96			
Breslow (mm)					
<1	631	1		1	
>=1	524	3.93	2.95-5.23	1.97	1.27-3.07
Missing	251	5.39	3.95-7.37	1.76	0.96-3.23
Clark					
2	396	1		1	
3	361	1.60	1.06-2.41	1.09	0.67-1.76
4	337	4.04	2.79-5.84	1.45	0.84-2.51
5	48	11.08	6.85-17.9	2.89	1.51-5.53
Missing	264	5.72	3.95-8.29	1.79	0.93-3.45
Site					
Head&neck	136	1		1	
Other specified	1085	0.43	0.32-0.57		
Other unspecified	135	0.91	0.63-1.32		
Morphology					
SSM	903	1		1	
Nodular	133	3.72	2.76-5.03	1.59	1.14-2.22
N.o.s.	225	4.14	3.21-5.33	2.23	1.61-3.08
Other	142	2.07	1.44-2.97	1.19	0.81-1.73

In the present study, the CART analysis constructed a pruned tree of 10 groups. Breslow thickness defined the first and the third split (Leiter, 2004; Garbe, 2002). Moreover, the cut-point automatically identified by the software in the third split (0.94 mm) almost overlapped the well-accepted limit of 1 mm used to separate thin from

thick melanomas. The original 10 groups were reduced, according to Kaplan and Meier comparisons, into six homogeneous groups with statistically significant different risks of dying. The attribution of a patient to the corresponding group seems very simple and it is based on the age, sex and Breslow. Using this method, it is possible to define the prognosis of a patient in a qualitative (from excellent to very poor) and quantitative (from 98.1 to 17.1% survival at 5 years) way. For example, facing a patient with a melanoma thinner than 0.94 mm, the prognosis significantly changes according to age with better 5-year survival (98.1%) if the patient is younger than 44 years than that expected if older than 50 (92.8%, $P < 0.01$). It is interesting that the effect of sex as independent prognostic factor seems to act for some subgroup, only when balanced with the effect of major predictors of risk as Breslow thickness and age. Indeed, the sex was not able to significantly modify the risk of dying facing the subgroups at better prognosis (defined in this study excellent and very good (5-year observed survival, more than 92%); on the contrary, sex was statistically relevant facing the other groups at less favourable prognosis.

The results from the present study may be influenced by the number of cases included and by the underline tumours characteristics; however, the three prognostic variables identified in the present study were the same as that in a study including 11 688 German patients (Leiter, 2004).

The use of the results from Cox analysis seems not so easy. In fact, Cox model produces for each variable coefficients (b) that show the relationship between the specific value of the variable and the outcome. The overall risk should take into account all the computed coefficients, according to the following formula: ex- p(coefficient [e.g. sex] + coefficient [e.g. age] + coefficient [e.g. Breslow] + coefficient [e.g. Clark] + etc.) in comparison with all the variables in the reference categories.

Although the CART approach has more closeness to the clinical reasoning process than regression models, it is not as common as expected. The main reason for the low dissemination of CART seems to have been just the low diffusion itself that had hampered both the growth of statistical interest and knowledge on it and therefore

the availability of specific procedures within more diffused software packages. Although it has been criticised (Marshall, 2001), its application has also been suggested (Lemon, 2003), and the method has been developed over time (Friedman, 2003), also to overcome the limits of traditional models (Xu, 2002). As regards the latter point, one of the advantages of CART analysis is just to be a nonparametric method that works with quantitative and qualitative data without any assumptions on their distributions (Lewis, 2000).

The present method has several appeals for clinicians: its automatic process, the easily interpretable results and the growing availability of software packages (Lewis, 2000). The number of papers that use this method to evaluate cancer prognosis is growing, and this will stimulate discussion and knowledge on this possibly very useful clinical tool.

In conclusion, the present study contributes to demonstrate the possible usefulness, in addition to traditional approach, of the CART analysis for defining the risk profile of a melanoma patient in a clinical set.

References: The original references of the paper have been included in the general list of the thesis.

6.1.1. Comment

The general idea behind my thesis is that CR's data has to be of good quality and as informative as possible. Only in this way, can it be useful to the largest number of possible stakeholders.
Cancer patients and clinicians are among the principle stakeholders. The previous paper (Paper 5) addresses the aspect of prognosis and compares two ways in which to communicate the outcome. Therefore, the topic is communication (how to give a prognosis) between oncologists and cancer patients.

Getting cancer patients to fully comprehend their prognosis is still an on-going problem (Cartwright, 2014). For example, 39% out of a sample of metastatic breast cancer patients were unaware of their real prognosis (Shin, 2016). In a study carried out in Japan, involving two samples of patients and radiation oncologists, more than half of the patients reported that they had not discussed, as they would have liked to, their life expectancy. In general, there was little agreement between patients and clinicians as to whether the prognosis had been discussed (Mackenzie, 2018). It is also complex and difficult for oncologists and patients to reach an agreement. In a recent paper, which addresses this topic concerning head and neck cancer patients, the Authors (Dronkers, 2018) prepared guidelines about sharing prognostic information. They suggested, among other things, 'to assure that the information given in the prognosis is as accurate and as specific as possible to the individual patient and to use the information in a 'digestible way ' (Dronkers, 2018).

Moreover, variables related to the patient may influence their understanding. For example, the capacity of breast cancer patients to understand their disease and treatment has been seen as highly correlated to them having received higher education, daily Internet access and being in paid employment. These factors gained the highest scores on disease prognosis knowledge (Berger, 2018). However, in general, patients' knowledge of their disease increases during the course of their treatment (Berger, 2018; Engqvist Boman, 2017).

The National Cancer Institute in the US, which has a website set up for patients, includes a section on the 'Questions to Ask Your Doctor about Your Diagnosis'. One of the questions listed is: "What are my chances of survival?" (https://www.cancer.gov/about-cancer/diagnosis-staging/questions).
The possessive adjective 'my' helps to clarify why a method like CART analysis would be appealing to both clinicians and patients. CART is capable of theoretically answering this question (giving a prognosis to those patients who share the same personal and cancer characteristics).

The use of the Cox model is widespread. However, the interpretation of the hazard ratio (HR) produced by such a model may not be straightforward to clinicians. HRs are commonly interpreted as *relative risks*. The reading may cause a misinterpretation of the clinical relevance of the treatment effect (Case, 2002; Trinquart, 2016). 'The mean limited survival time' (Pak, 2017; Trinquart, 2016) has been proposed as an alternative; it offers intuitive, clinically meaningful interpretation (Pak, 2017).

As regards to CART and nomograms, a comparison of the capability of predicting prostate cancer-related outcomes, revealed a better predictive performance of the nomogram. Moreover, the Authors stated that CART (and other tested methods) rely on methodologically sound and valid alternatives (Chun, 2007).
In the field of urology nomograms and CART have been considered the best aids in decision making out of those available (Shariat, 2009).
Nevertheless, the CART analysis has been applied to identify the predictive variables to be included in a nomogram (Makkouk, 2017).

One of the substantial criticisms raised against CART is the choice of the variables which relies on a mathematical algorithm which maximises the model's discriminant ability (Shariat, 2009).
However, this is not the application I propose.
On the contrary, once the prognostic (or predictive) variables are identified, the CART could be applied to split patients into homogenous groups.

After the publication of Paper 5 (in 2006), the use of CART analysis has been documented many times in scientific literature to address different clinical outcomes. Some examples are illustrated below.

CART analysis has been applied:

- to measure the prognostic effect of lymph node dissection (in particular more than ten lymph nodes) in women with corpus uterine cancer (Abu-Rustum, 2008);
- to stratify risk groups for pulmonary complications after lung resection, in resected lung cancer patients (Kim, 2012);
- to distinguish low-risk sub-groups in melanoma stage III patients for which adjuvant therapy may not be warranted (Egger, 2013);
- to evaluate which clinicopathologic factors (age, stage, tumour subtype, grade, myometrial invasion, total lymph nodes removed, and para-aortic lymph node) influenced overall survival (OS) in endometrial cancer (Barlin, 2013);
- to identify different prognostic groups with different molecular subtypes of ovarian cancer (Lu, 2016);
- to investigate the risk factors for bone-only metastasis in advanced breast cancer patients (Diessner, 2016);
- to find optimal cut-offs points for prognostic variables identified by Cox model in a cohort of myeloid leukaemia patients in blast phase (Jain, 2017);
- to identify three risk levels which could be incorporated into the TNM staging system for nasopharyngeal cancer patients (Yuan, 2017);
- to classify patients with surgically resectable micropapillary bladder cancer in three risk groups (low, high, highest) under continuous observation and disease-free (Fernández, 2017);
- to develop a prognostic transcriptome molecular staging model for oesophageal squamous cell carcinoma in China (Guo, 2018);
- to develop a polygenic approach (genes, their regulators, and DNA repair genes) to identify various polymorphic variants in determining lung cancer susceptibility (Bhardawaj, 2018);
- to identify the predictors of treatment-related (eribulin) toxicity in advanced soft tissue sarcoma of Japanese patients (Kobayashi, 2019);

- to develop an algorithm that has achieved high accuracy in being able to identify breast cancer recurrences using available administrative data in a universal health system in Canada, (Xu, 2019).

However, although there is much scientific literature based on the CART application, I was unable to find anyone who could compare whether the prognosis was understandable to the cancer patients, using the different statistical approaches.

Therefore, a formal study comparing the understandability of HRs (from Cox) and patient groupings, (from CART) should be planned and analysed with qualitative methods (Cartwright, 2014). Hence, if the CART application should prove more effective in informing patients than other approaches, it could open the road for its widespread clinical application.

7. Evaluation of cancer care

The previous papers (Paper 1-5) display some of the indexes produced by registries and their use.

Among the purposes of registries' data, there is also the evaluation of the quality of the diagnostic and therapeutic pathways and their agreement with clinical guidelines (Jensen&Storm, 1991). The sources of information of CRs represent many of the 'contacts' of patients with the health systems: e.g., screening participation, out-patient visit, biopsy, hospitalisation, surgery, bio-marker testing, prescriptions, chemotherapy, radiotherapy, palliative care. Moreover, the number of sources has increased over time, thanks to the digitalisation of the health systems.

Therefore, CRs can describe the patients care throughout the pathway from the date of diagnosis until death.

The population-based point of view of a registry allows it to measure real-world care and effectiveness, appropriateness, accessibility and equity in treatments (Spitale, 2017; Malin, 2006; Schneider, 2004).

Appropriateness remains a relevant topic due to the documented overuse of unnecessary health services and sometimes can be unsustainable for single patients and healthcare systems (Grilli, 2018).

Sustainability of cancer care is a general and recognised concern in the light of the simultaneous increase in the number of cases and cost of care. Therefore, the evaluation of quality of care can also be helpful to address such issue (Campbell, 2000).

Moreover, considering the long-time necessary to accomplish the full registration process (Zanetti, 2015), short term deliverables should also be considered.

Cancer-specific indexes focusing on diagnosis and/or treatment could represent the ideal target for such type of short-term deliverables.

Clinicians, policymakers and cancer patients could all be interested in a measure of the quality of diagnosis and care at the population level.

The analysis of the distribution of cancer-specific treatments in different area hospitals and services and the comparison between volumes and outcomes of activity can identify too low volumes of clinical or surgical care inconsistent with a similar offer in the same healthcare system with an elevated level of experience and capability (Amato, 2017). Proper management is even more challenging for rare cancers for which the need for, and effectiveness of specific networks are underlined (Gatta, 2017; Nicolai, 2018).

Registries have to collaborate with clinicians and policymakers in the evaluation and improvement of quality of care (Jochems, 2017; Spitale, 2017).
The quality of the whole clinical path should be evaluated, including, for example, palliative care in terminally ill patients (Ziegler, 2018).

Many registries are presently working on this topic (e.g., Spitale, 2017; Andreano, 2016; Andreano, 2018; Sacerdote, 2016; Vrijens, 2018; Vlayen, 2012; Stordeur, 2012) contributing to increasing knowledge on the quality and accessibility of the best possible care in the real-world setting.

The next paper (6) presents a pioneer experience - at least in Italy - which started assessing the feasibility of an evaluation of quality of care based on indexes computable based on registry and health records for the most frequent cancer sites (Caldarella, 2012).

The usefulness of CR's data manifests measuring and evaluating clinical care of melanoma at the population-level based on indicators. To measure and highlight, with high-quality indexes, points of weakness contributes to circumscribe the areas (services, procedures) that need to be improved to offer a better quality of care to patients. Therefore, the stakeholders of this activity of CRs are several: policymakers, clinicians, patients and citizens.

7.1. Paper 6

The paper has been published in Melanoma Res. 2013; 23: 283-9 with the following title:

"Indicators of the standard of care for melanoma: Tuscany data".

Authors Crocetti E[a], Caldarella A[a], Massi D[b], Sacchettini C[a], Amunni G[a,d], Borgognoni L[c].

Affiliation: a) Clinical and Descriptive Epidemiology Unit and b Melanoma Early Diagnosis Service, Institute for Cancer Study and Prevention ISPO, Florence, Italy, b) University of Florence, Florence, Italy, c) Plastic and Reconstructive Surgery Unit, Regional Melanoma Referral Centre, S.M. Annunziata Hospital, d) Tumour Institute of Tuscany (ITT), Florence, Italy

Author Contributions: I (EC) declare to have participated to conceive the idea of the study starting from the registry's point of view, I have planned and designed it, I performed the analysis and drafted the first draft. The other Authors revised critically the paper and approved the final version of the manuscript

Abstract

Formal indicators for the evaluation of the quality of melanoma care are needed. We identified 13 process indicators, which encompassed early diagnosis, pathology reporting and surgical treatment. We evaluated the adherence to these indicators using a population-based series on incident skin melanomas (only primary melanomas) for the year 2004 (687 cases) and for the first semester of 2008 (539 cases). We compared the indicators for the 2 years. The melanoma incidence increased between 2004 and 2008. There were statistically significant increases in the percentage of thin (<= 1 mm) melanomas (from 50.7 to 61.3%) and in the number of pathology reports that mentioned ulceration (from 61.4 to 84.6%) and margin statuses (from 76.8 to 84.3%). The percentage of patients staged by sentinel lymph node biopsy was stable (63%) and was higher for patients younger than 75 years of age (74%). The number of nodes almost invariably exceeded the proposed site-specific cutoff reference, and, in 2008, the number of nodes removed was always reported for

lymphadenectomy. From 2004 to 2008, surgical and pathological waiting times increased. Collection and analysis of these indicators would enable continuous evaluation of the quality of melanoma care in Tuscany and provide sources for a comparative study between Italy and abroad.

Introduction

The incidence of skin melanoma has greatly increased in Italy during recent years (Airtum, 2009), as in many other western countries. In Tuscany (central Italy), cancer care is coordinated by the Tumour Institute of Tuscany (ITT), a network-based Institute that organises and supervises public health services in the community with the aim of providing optimised cancer treatment. The ITT also conducts studies on tumours (http://www.ittumori.it). In particular, in 2007, it published clinical recommendations to standardise diagnostic and therapeutic pathways for skin melanoma for specialists and hospitals throughout the region. On the basis of the ITT's recommendations, the standard and effective treatment for cutaneous melanoma is now surgery-based (ITT, 2007). As most skin melanomas are diagnosed early (when, in most cases, the lesion is thin) (Crocetti, 2010), complete pathological examination and timely surgery represent the best option for the majority of patients. To reduce variability in cancer care procedures and improve the general quality of care, the ITT works on the basis of indicators. Specific indicators for breast, lung, colorectal and ovarian cancer care are available (Caldarella, 2012).

The role of indicators in evaluating clinical care is increasing (Mainz, 2003; Mainz, 2009). However, to our knowledge, in Italy, there are no formal indicators for care quality for melanoma. In the USA, the four measures available in the National Quality Measures Clearinghouse address only a few points in the clinical course (diagnosis, imaging and follow-up, http://www.qualitymeasures.ahrq.gov/), although further indicators, mainly for surgical treatment, have recently been evaluated (Bilimoria, 2009).

This paper aims at evaluating the clinical care of melanoma in Tuscany through formal indicators and gauges whether ITT recommendations have succeeded in instituting and improving standardised patterns of care

Materials and methods

Setting and study population

We first evaluated primary skin melanoma cases occurring in Tuscany between 2004 and 2008, excluding melanomas with unknown primary sites. We retrieved melanoma incidence data for 2004 from the archive of the Tuscan Regional Cancer Registry (RTRT), a population-based cancer registry that collects data from several information sources, including hospital discharge notes, pathology reports, and death certificates (Paci, 2007). The RTRT was, however, not active in the Tuscan province of Arezzo in 2004, accounting for 330123 out of 3566071 Tuscan inhabitants.

We identified melanoma cases occurring in 2008 from the regional pathology archive (http://web.rete.toscana.it/attinew/?LO=00000001a6b7c8d9000000030000003400 003f7a397d5a0f0000000000012b800000000000000000000000000000000000 00000000000000000000000000000000000000&MItypeObj=application/pdf), which includes all pathology reports, positive or uncertain, for cancers, as defined by SNOMED (vers. 3) morphology codes M-8**** and M-9****. We also collated the incidence of the term 'melanoma' in each pathology report from the first semester of 2008 (from January to June). We included individuals residing in Tuscany and restricted our interest to those with primary melanomas, making sure that the individuals had no pathology reports or hospital discharge notes, nor had they undergone any outpatient treatment for melanoma in previous years. When melanoma-related reports were not available (neither in the RTRT nor in the regional pathology archive system), as was the case with a few individuals, we searched through the reports in the regional pathology departments. In total, we analysed 687 melanomas from 2004 (161, 23.4% *in situ*) and 539 from the first semester of 2008 (143, 26.5% *in situ*). For each patient, we collected and detailed information on pathology characteristics, diagnostic methods, staging and type and dates of surgery.

THE LIST OF INDICATORS

Working groups within the ITT studied specific cancers in depth, in order to better define and exchange experiences on specific topics and to make better clinical recommendations. The melanoma working group is a multidisciplinary panel of individuals and institutions, each belonging to the regional cancer network, including experts in melanoma (dermatologists, pathologists, surgeons, oncologists and epidemiologists), academic institutions and community hospitals. In 2007, the ITT published a set of recommendations on the diagnosis and treatment of melanoma based on the investigations of the melanoma working group, whose components agreed upon a consensus procedure based on scientific literature and on their experience (ITT, 2007).

The ITT Melanoma Working Group (see Acknowledgements) identified a preliminary set of indicators by analysing scientific guidelines on melanoma diagnosis and treatment, literature reviews (Lilford, 2007; Garbe, 2010; Garbe, 2012; Marsden, 2010; AIOM, 2013; Coit, 2009; Dummer, 2011; Dummer, 2012; Morton, 2006; Aitken, 2006; Carli, 2003; Carli, 2001; Schiffner, 2003; Wagner, 2000; Lens, 2008; Garbe, 2007; Erickson, 2008; Blazer, 2007; Riker, 2006; Lavie, 2007; Eddy, 2003; Wright, 2011; Mocellin, 2011; Sladden, 2009; Lens, 2007; Haigh, 2003; Borgognoni, 2004; Amersi, 2007; Sondak, 2007; Cook, 2003; Cook, 2008; Chakera, 2009; Lens, 2002; Greene, 2002) and panel discussions. A necessary condition while selecting the indicator was that only data from the Cancer Registry could be used, so that a homogeneous analysis could be carried out and objective results valid for the whole of Tuscany could be obtained. After meetings, discussions and email exchanges, the Group selected 13 indicators. The chosen indicators measured the appropriateness of the clinical care procedure on the basis of ITT recommendations (ITT, 2007) and encompassed the diagnosis, pathology reporting and surgical treatment of melanoma care.

STATISTICAL ANALYSIS

We calculated percentages and medians for quality measures using the w2-test or median test to compare values between the years. The results were reported at a regional level but were also available at subregional levels for local use.

Results

Table 14 (n.1 in the original numeration), presents the selected quality indicators.

Table 15 (n.2 in the original numeration), compares the average regional values of each quality indicator for 2004 and 2008.

1. Percentage of incident cases <= 1 mm

The thickness of a melanoma is one of the most relevant prognostic factors. The ITT and other recommendations state that a pathology report must specify the thickness. The number of new melanoma cases increased from 2004 to 2008, the standardised incidences being 14.1 and 18.4 per 100000, respectively. Almost all the pathology reports recorded information on Breslow's thickness (97.5% in 2004, 513/526 and 98.0% in 2008, 388/396). The percentage of thin (<= 1 mm) melanomas increased over time from 50.7 to 61.3% (P = 0.001). Moreover, there was a significant percentage of *in situ* melanomas: 23.9% (161 out of 674 invasive and *in situ* melanomas) in 2004 and 26.9% (143/531) in 2008.

2. Percentage of pathology reports on incident invasive melanomas mentioning ulceration (present or absent)

Ulceration has been included in the staging process since 2002 (Greene, 2002). The ITT recommendations state that pathology reports must specify ulceration (ITT, 2007). The percentage of relevant pathology reports mentioning ulceration statuses (present or absent) increased over time from 61.4% (404/526) in 2004 to 84.6% (334/396) in 2008 (P<0.001). This increase concerned both thin (from53.1 to 82.5%, P < 0.001) and thick melanomas (from 70.4 to 88.0%, P < 0.001).

Ulceration was present in a decreasing percentage of reports over time – 34.4% in 2004 (111/323) and 25.7% in 2008 (86/335) (P = 0.005).

3. Percentage of pathology reports on excised incident invasive melanomas with margin statuses (positive or negative).

The margin status confirms the completeness of melanoma excision. The ITT recommendations state that pathology reports must specify the margin status (ITT

2007). The percentage of the margin status (whether positive or negative) mentioned in the pathology reports increased significantly, from 76.8% (404/526) in 2004 to 84.6% (334/396) in 2008 (P = 0.005). On stratifying for thin and thick melanomas, an almost similar statistically significant increase was observed (thin 79.1 vs. 85.4%, P = 0.064; thick 74.3 vs.82.7%, P = 0.052).

In 2008, 92% of the excisions had negative margins. Among the 24 cases with positive margins, 15 (62.5%) dealt with large melanomas of the face and acral sites, for which an incisional biopsy was planned.

Table 14 (n.1 in the original numeration): Quality indicators of skin melanoma care and areas of interest (SLB, sentinel lymph node biopsy).

	Area	Main indicators
1	Early diagnosis	Percentage of incident cases<=1mm
2	Pathology reporting	Percentage of pathology reports on incident invasive melanomas with ulceration mentioned (present/absent)
3	Pathology reporting	Percentage of pathology reports on excised incident invasive melanomas with the margin status (positive/negative)
4	Surgery	Percentage of patients staged by SLB
5	Surgery	Percentage of patients staged by SLB <75 years of age
6	Surgery-Pathology	Percentage of patients with positive sentinel nodes
7	Pathology reporting	Percentage of lymphadenectomies with the number of removed lymph nodes reported
8	Surgery-Pathology	Percentage of cervical lymphadenectomies with >15 nodes removed
9	Surgery-Pathology	Percentage of axillary lymphadenectomies with >10 nodes removed
10	Surgery-Pathology	Percentage of inguinal lymphadenectomies with >5 nodes removed
11	Time for pathology	Waiting time between the first surgery and the pathology report (time for pathology of diagnosis; days, median)
12	Time for surgery	Waiting time between the diagnostic pathology report and the second surgery (enlargement/SLB, time for surgery; days, median)
13	Time for overall surgical treatment in positive SLB	Waiting time between the first melanoma diagnosis and the final lymphadenectomy in patients with a positive SLB (days, median)

Table 15 (n. 2 in the original numeration): Average regional value of the quality indicators of skin melanoma care for 2004 and 2008, probability of equality according to testing (X^2, rank sum).

	Main indicators	Regional average 2004	Regional average 2008	p
1	Percentage of incident cases<=1mm	50.7	61.3	0.001
2	Percentage of pathology reports on incident invasive melanomas with ulceration mentioned (present/absent)	61.4	84.6	<0.001
3	Percentage of pathology reports on excised incident invasive melanomas with the margin status (positive/negative)	76.8	84.3	0.005
4	Percentage of patients staged by SLB	63.6	63.3	0.95
5	Percentage of patients staged by SLB <75 years of age	73.9	74.3	0.95
6	Percentage of patients with positive sentinel nodes	26.1	26.3	0.97
7	Percentage of lymphadenectomies with the number of removed lymph nodes reported	98.2	100	0.51
8	Percentage of cervical lymphadenectomies with >15 nodes removed	100	100	1
9	Percentage of axillary lymphadenectomies with >10 nodes removed	85.7	66.7	0.20
10	Percentage of inguinal lymphadenectomies with >5 nodes removed	100	100	1
11	Waiting time between the first surgery and the pathology report (time for pathology of diagnosis; days, median)	9	13	<0.001
12	Waiting time between the diagnostic pathology report and the second surgery (enlargement/SLB, time for surgery; days, median)	23	26	0.084
13	Waiting time between the first melanoma diagnosis and the final lymphadenectomy in patients with a positive SLB (days, median)	72	82.5	0.71

4. Percentage of patients staged by sentinel lymph node biopsy

Sentinel node biopsy (SLB) enables melanoma pathological staging. The sentinel node (SN) status is a major prognostic factor. The ITT recommendations state that SLB has to be performed when the thickness is at least 1 mm, in cases of ulceration or when

the Clark level is at least IV (ITT 2007). The percentages of SLB were similar in 2004 and 2008 – 63.6% (161/253) and 63.3% (95/150), respectively.

5. Percentage of cases with sentinel lymph node biopsy in patients < 75 years old

SLB is a technique under-utilised in elderly patients (Cormier, 2005). As it is more invasive than excision, it may sometimes not be indicated in older patients because of the presence of comorbidities. Moreover, most SLB trials did not include patients above 75 years of age (Morton, 2006; McMasters, 2004; Reintgen, 2004). Among the total number of patients who underwent SLB, the majority were less than 75 years of age – 73.9% in 2004 (139/188) and 74.3% in 2008 (75/101).

6. Percentage of patients with a positive sentinel node

The SN status is a major prognostic factor; therefore, node/s should be correctly removed and pathologically examined (ITT, 2007). Around one-quarter of SLBs were positive both in 2004 [26.1% (42/161)] and in 2008 [26.3% (25/95)].

7. Percentage of lymphadenectomies with the number of nodes reported

The ITT recommendations state that the pathological report must specify the number of nodes removed after lymphadenectomy (ITT, 2007). The number of nodes with metastases is a basic point for staging (Greene, 2002). The number of nodes was mentioned in almost all the relevant pathology reports in both years. The number of the nodes removed in cases of complete node dissection was also reported in almost all the cases, both in 2004, 98.2% (54/55) and in 2008, 100% (24/24).

Indicators 8, 9 and 10 concern the number of lymph nodes removed during lymphadenectomy. The number of lymph nodes removed by lymphadenectomy could be considered a measure of the quality of surgical care.

8. Percentage of cervical lymphadenectomies with >15 lymph nodes removed

In cervical lymphadenectomy, a cut-off value of 15 lymph nodes has been adopted (Bilimoria 2008). The number of nodes removed exceeded the cut-off value in all cases in both years. In 2008 the median number of excised nodes was 24.5.

9. Percentage of axillary lymphadenectomies with >10 lymph nodes removed

In axillary lymphadenectomy, a cut-off value of 10 lymph nodes has been adopted (Bilimoria, 2008). The number of nodes removed exceeded the cut-off value in 85.7% of cases in 2004 (24/28) and in 66.7% (6/9) cases in 2008. In 2008, the median number of excised nodes was 16.

10. Percentage of inguinal lymphadenectomies with >5 lymph nodes removed

In inguinal lymphadenectomy, a cut-off value of 5 lymph nodes has been adopted (Bilimoria, 2008). The number of nodes removed exceeded the cut-off value for all patients in both years. In 2008, the median number of excised nodes was 14.

Indicators 11, 12 and 13 concern the waiting time.

11. Waiting time between the first surgery and the corresponding pathology report

The time between the removal of a suspected skin lesion and the pathological diagnosis of melanoma may influence the timeliness of the necessary subsequent actions (enlargement, SLB, etc.). The median diagnostic time was 9 days in 2004 and 13 days in 2008 ($P < 0.001$). On analysing thin and thick melanomas separately, the median time increased for thin melanomas (from 9 to 14 days, $P < 0.001$) and was stable for thick ones (from 10 to 11, $P = 0.509$).

12. Waiting time between the diagnostic pathology report and second surgery (enlargement/sentinel node biopsy)

The ITT recommendations (ITT, 2007) suggest performing enlargement or SLB within 3 months from the biopsy. The median time was 23 days in 2004 and 26 days in 2008 ($P = 0.084$). There were no differences with regard to Breslow's thickness.

13. Waiting time between the first melanoma diagnosis and final lymphadenectomy in patients with positive sentinel node (time for overall treatment in sentinel node positive patients).

The total waiting time for the surgical treatment of a melanoma patient is the time from the first biopsy to the report of the lymphadenectomy (in case of a positive SN). It was reported to be 72 days in 2004 and 82.5 days in 2008 (P = 0.71).

Discussion

This study provides a comprehensive population-based evaluation of the diagnostic and surgical treatment of melanoma in the central Italian region of Tuscany. Our proposed indicators are based on local cancer registry data and the official regional pathology databases. The method of computation of the indicators enables us to compare different years or geographical/administrative areas when the same information systems are available. The population-based setting theoretically provides unbiased information and shows average measures for clinical care in the population. In Tuscany, a multilevel strategy for early diagnosis of skin melanoma has been in use since the 1990s. The strategy includes specific initiatives for raising awareness in the population, namely, distribution of information packs for skin self-examination, organising training courses for general practitioners (GP) and setting up skin units across the region. In addition, the multi-disciplinary ITT Melanoma Working Group has developed specific clinical recommendations to optimise the clinical course throughout the region (ITT, 2007).

A hypothetical and optimal clinical course for skin melanoma should start when an individual discovers a changing skin lesion, the GP confirms the need for further evaluation, the skin clinic dermatologist performs the excisional biopsy and the pathologist reports the diagnosis, specifying all the main prognostic factors that address the subsequent appropriate surgery. Therefore, the aim of the ITT Melanoma Working Group in identifying possible melanoma indicators was to find parameters useful in evaluating the diagnostic phase, the completeness of the pathological reporting and the surgical treatment of the melanoma. The indicators analysed in the present study were selected by the ITT Melanoma Working Group on the basis of data available in the Cancer Registry in order to have unbiased, homogeneous and comparable information. On the basis of the above requirements, the ITT Melanoma

Working Group analysed scientific guidelines on melanoma diagnosis and treatment and performed literature reviews (Garbe, 2010; Garbe, 2012; Marsden, 2010; AIOM, 2013; Coit, 2009; Dummer, 2011; Dummer, 2012; Morton, 2006; Aitken, 2006; Carli, 2003; Carli, 2001, Schiffner, 2003; Wagner, 2000; Lens, 2008; Garbe, 2007; Erickson, 2008; Blazer, 2007; Riker, 2006; Lavie, 2007; Eddy, 2003; Wright, 2011; Mocellin, 2011; Sladden, 2009; Lens, 2007; Haigh, 2003; Borgognoni, 2004; Amersi, 2007; Sondak, 2007; Cook, 2003; Cook, 2008; Chakera, 2009; Lens, 2002) and panel discussions with the purpose of identifying possible indicators of an early diagnosis and appropriate melanoma treatment. After discussions, meetings, reports and email exchanges, the Melanoma Working Group selected 13 indicators on the basis of the Cancer Registry data (Table 14 [n.1 in the original numeration]). All measures were process indicators that are among the most suitable management tools for measuring quality (Lilford, 2007).

During the analysed period, there was a strong increase in melanoma diagnosis, in both the standardised and crude rates. As they are based on the number of cases, crude rates show the real workload of a healthcare system in terms of the need for diagnosis and treatment. This is the rate that must be taken into account when 2 years are compared. The increased incidence was mainly driven by the growing percentage of thin melanomas (<= 1 mm), which in 2008 was 61.3% of all newly diagnosed invasive melanomas. This indicator attests to the efficacy of early diagnosis in the region, as a consequence of increased awareness in the population, the role of the GP and of the activity of the dermatologists in skin clinics (with the support of dermoscopy). In the area of the present study, a randomised clinical trial confirmed the usefulness of dermoscopy in improving melanoma diagnosis and reducing the number of pigmented skin lesions excised for diagnostic verification (Carli, 2004). The possible use of dermoscopy in the diagnosis of pigmented skin lesions is included in the ITT melanoma recommendations (ITT, 2007; Carli, 2001; Carli, 2004).

Mortality for melanoma is still increasing in Tuscany (Chellini, 2007). No effective treatments for advanced melanomas were available in Italy at the time of the study. Encouraging results from new molecular drugs have emerged only recently

(Eggermont,2010; Hodi 2010; Chapman, 2011). Therefore, early diagnosis and appropriate surgical treatment are the primary recourse for the healthcare system. However, an increase in diagnostic drift and instances of reclassification of previous severely atypical dysplastic nevi as melanoma has been found (Frangos, 2012). The increased diagnoses of thin but not deadly melanomas (Crocetti, 2010) have a positive significance, although the possible role of over-diagnosis of indolent pigmented lesions should also be considered (Welch, 2010).

Melanoma prognosis is strictly dependent on characteristics that may be evaluated only by a pathologist, and, for this reason, there are indicators that address the completeness of the pathology reports. The quantity of information available in the pathology reports has increased over time. In 2008, almost all reports included Breslow's thickness (98%), around 85% included the presence or absence of ulceration, 84% included the margin status for excised melanomas and all reports included the number of nodes in patients of lymphadenectomy. The latter results were as good as those recently documented in the USA (Bilimoria, 2009).

As regards the weaknesses of this study, for the patients diagnosed in 2008, we used only the information from pathology reports. Therefore, few cases may have been lost and the estimates may be slightly overestimated.

Moreover, not all the relevant prognostic factors were analysed; for example, the number of mitoses was not included. However, the analysed data refer to 2004 and 2008, and it was only in 2009 that AJCC melanoma staging (Blach, 2009) included the number of mitoses as an important prognostic factor. For this reason, recent melanoma guidelines, such as those in the UK (Marsden, 2010), include the mitotic count among the requirements for a pathological report, and this parameter will be included in the updated ITT recommendations for melanoma that are to be published. They will also be included on future indicator analyses.

About two-thirds of eligible patients had SLB. The percentage for patients younger than 75 years (74%) shows that this procedure was performed mainly on non elderly

patients. Presumably, the presence of comorbidity in patients older than 75 years influenced the decision of the surgeon for SLB. Once performed, an SLN biopsy showed approximately the same rate of positivity in 2004 as that in 2008 – about 26%. The mean regional figure of 26% in Tuscany is higher than the percentage of 16% reported in the Multicenter Selective Lymphadenectomy Trial I (Morton, 2006) and the 19.9% reported in an Italian Multicentre Study (Testori, 2009), indicating appropriate surgical and pathological approaches. Pathology reports always specify the number of lymph nodes removed during lymphadenectomy in the entire region. The cutoff values obtained by Bilimoira in a recent work on melanoma indicators in the USA (Bilimoria, 2009) were adopted to evaluate the quality of the surgical procedures in stage III patients. Moreover, the same cutoff values made it possible to compare the results between the two studies. The number of nodes removed from sites in the cervical and inguinal regions always exceeded the levels of 15 and 5 nodes, respectively. As regards the axillar region, two-thirds of patients had 10 or more nodes removed in the first semester of 2008. However, this number is based on just nine patients. The values registered in Tuscany are higher than those reported in Bilimoira's study (Bilimoria, 2009). Moreover, it has been reported in a previous American study that less than 50% of patients with positive SLB in the USA underwent complete node dissection (Bilimoria, 2008), whereas the percentage in Tuscany was 63%.

A crucial aspect for any public healthcare system, for both patients and healthcare policy makers, is the waiting time. Three indicators addressed waiting times in the diagnostic and surgical course of melanoma. The median time for receiving the report for a skin lesion suspected to be a melanoma was 13 days in 2008. The length of this period depended on the pathologist, his/her workload, and the availability of resources. It increased from 2004 by 44%; however, the crude incidence rate (number of melanomas) increased during the same period by about 33% and, therefore, the longer waiting time for melanoma diagnosis resulted mainly from the increased number of patients awaiting diagnosis. The waiting time after a pathology diagnosis depended on the surgeons' workload and resources. The waiting time, for wider excision or SLB, increased by 13% (median from 23 to 26 days) during the aforementioned period (during which there was an increase in melanoma incidence).

In particular, there was an increase not only in the number of thin melanomas but also in the number of cases with a thickness greater than 1 mm (an increase of 4%). The total surgery-related waiting time, commencing with the positive SLB report and ending with lymphadenectomy, increased by about 15% (from 72 days in 2004 to 82.5 days in 2008). The implementation of a modified EORTC protocol in Tuscany as the standard procedure for extensive pathological handling of SLB (Cook, 2001) presumably had an impact from 2004 to 2008, lengthening the time of reporting.

Conclusion

In this study, we analysed a set of 13 quality indicators for evaluation of the diagnosis, surgical treatment and pathological definition of melanoma. These indicators, evaluated using a population-based dataset in Tuscany, showed an increase in the number of thin melanomas (a mark of early diagnosis), an improvement in the completeness of the pathology reports, and good surgical management. Some of the latter improvements may be related to the development of ITT recommendations for melanomas for the regional healthcare system in 2007, a product of the regional multidisciplinary Melanoma Group. These as well as other indicators, such as the mitotic rate (Balch, 2009; Thompson, 2011), should be monitored to identify the potential improvements and shortcomings in melanoma care.

References: The original references of the paper have been included in the general list of the thesis.

7.1.1. Comment

The previous paper (Paper 6) exemplifies a practical application of the data from a cancer registry (CR), and more generally, the CR's staff's technical competence in managing data. The information from a CR, as well as the information selected from the pathological reports, have been used to compare some quality indexes in the diagnosis and treatment of MM in the Tuscan region.

MMs almost always have a microscopic verification. All quality indexes were measured in MMs with microscopic confirmation (e.g., cases with biopsy, surgery). The use of pathological reports was made possible by the availability of digitalised information and the expertise of the CR staff diligently data mining these reports (Crocetti, 2005). This last point made it possible to select skin melanoma, both *in situ* and invasive, Breslow's thickness, and information on the involvement of lymph nodes.

In this way, the CR was useful for the entire scientific community involved in MM care, offering a reliable and comparable population-based data set and the ability to analyse it.

The CR has entered the real-world as a useful and valid partner, collaborating with dermatologists, pathologists, surgeons, and policymakers, to measure the quality of care in the early stages of MM.

With the help of CRs, the diagnosis and course of cancer treatment could improve, and therefore modify the information given by patients during their contact with healthcare services. That same information would then be used by the CRs to produce still better, high quality and insightful information at the disposition of the scientific community, which furthermore, could verify the quality of MM care.

It is a virtuous circle that the CRs must exploit to show their concrete and constructive role in the healthcare system.

8. Discussion

8.1. Introduction

The field of research exemplified in the six papers presented in this thesis focusses on the process followed by registries in the transformation of raw data into information applicable in the real world.

This procedure identifies three crucial and inter-related areas: quality, informativeness, and usefulness of data.

The sequence of these six papers illustrates a possible fruitful path that registries have to consider to become a bringer of knowledge in cancer epidemiology and care. Although each issue represents a specific and independent topic, which may apply to various subjects, a correct and fruitful activity of registries must necessarily consider such aspects jointly. In Figure 13 the leitmotiv of the research presented in this thesis, which deals with cancer registration as a whole, from raw data to concreate messages to guarantee quality, informativeness and usefulness of data, is exemplified.

The present thesis summarises how an application of this research in the real-world context contributed to improving the role of population-based cancer registries as valuable tools for the development of knowledge on cancer. The topics covered by the six papers presented here concern not only cancer epidemiology but also public health organisation and clinical and translational research.

This thesis may be helpful to those who want to understand how powerful registries maybe if only their potential is exploited to serve all the possible stakeholders in the community. Registries collect and provide data on all malignant tumours arising in a specific population (Shanmugaratnam, 1991). These data can be useful in different contexts and for different stakeholders (e.g., clinicians, patients, public health operators, policymakers).

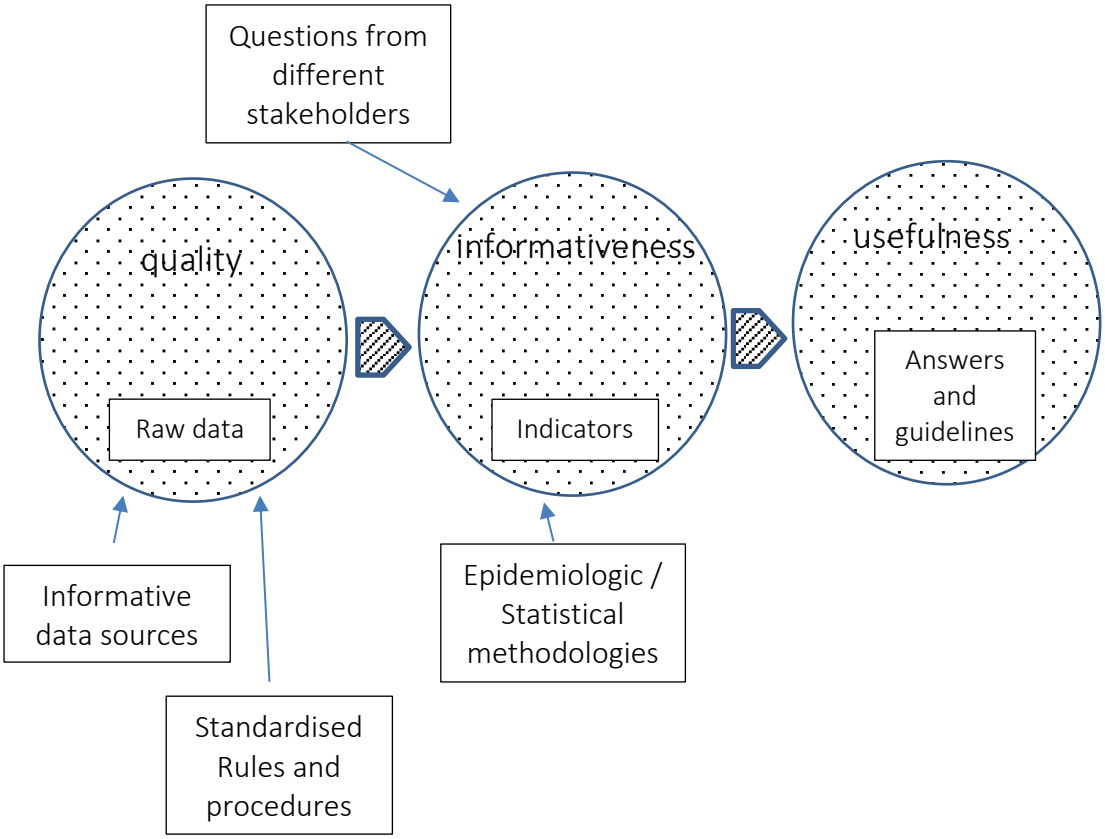


Figure 13: Exemplification of the cancer registry' data path presented in this thesis.

Moreover, it offers a structured set of effective methods and tools for those who work in this field and are willing to improve their registry's capability to be effectively used by different stakeholders.

Most of the papers presented here ground on skin melanoma. For two of them, which address generally applicable methods in cancer epidemiology, unpublished analyses have been added in which these methods have been applied explicitly to skin melanoma.

8.2. Quality of data

The evaluation of data quality is a prerequisite for any data and methodology which strives to provide scientific evidence. In the introduction, the available methods to

evaluate the different dimensions of registry data quality (validity, comparability, completeness and timeliness) have been presented. They have been used for many years and are widely appreciated.

In general terms, they are based on simple computations providing numbers. Based on such evaluations, registries may be attributed various degrees of quality, as exemplified by the North American Association of Central Cancer Registries using specific criteria to identify two positive certification levels: gold and silver (https://www.naaccr.org/certification-criteria/).

Regrettably, the evaluation of quality indicators is not always explicit and reproducible because of the lack of standard cut-offs. Notwithstanding, registrars attribute to a specific range of values of the analysed indexes the meaning of good quality when considered high enough, e.g., ..."93% were morphologically verified"... (Leinonen, 2017), or "The proportion of microscopically-verified cases increased from 73.6% in 2002 to 82.3% in 2012" (Ryzhov, 2018). Similarly, when they were low enough, e.g., "The proportion of cancers with uncertain or ill-defined primary site and the proportion of death-certificate-only registrations were both low at 1.9% and 2.6%" (Leinonen, 2017), or "...with death-certificate-only (DCO) proportions stable at around 0.1% and unknown stage recorded in 9.6% of male and 7.5% of female solid tumours..." (Ryzhov, 2018).

Although the quality evaluation is a complex process based on the joint interpretation of several parameters, the lack of unique and shared cut-offs for indexes may hamper the reproducibility and consequently, the interpretation of the literature.

For example, when comparing the data selection in a study on childhood leukaemia of the Concord-2 project (Bonaventure, 2017) with an International incidence of childhood cancer study (Steliarova, 2017), the criteria for case-series selection did not seem to match. In both situations, the thresholds adopted for acceptance for the numerous checks performed in the quality assessment process were unmade explicit (Crocetti, 2017b).

Therefore, although the quality of data is crucial for registries and many methods to evaluate it are available, overall agreed rules with specific cut-offs are lacking, and ad hoc interpretation cannot be ruled out.

Paper 1 represents one of the possible approaches to overcome the limitation caused by subjectivity in data quality estimation and proposes an entirely innovative method for quality assessment that could enrich the traditional toolkit by introducing an objective evaluation.

The method presents the application of a mathematical law, Newcomb-Benford law (N-B) (Newcomb, 1881; Benford, 1938) describing the frequency distribution (probability function) of the first digits (from 1 to 9, being 0 not considered) of real numbers. In the series of numbers which abide by N-B law, the frequency distribution has a robust positive skewness, with 1 being the first digit more than 30% of times.

Paper 1 evaluated for the first time whether cancer incidence rates abided by this law. To document it, a vast data set of 146,590 incidence rates was randomly sampled from the dataset used for CI5-X, and the 1-digit distribution was evaluated showing a positive skewness (0.84) with mean higher than the median (as expected by the theoretical N-B distribution) and with ratios between 1st and 9th digits and 1st and 2nd digits corresponding to the theory.
In the second part of paper 1, once incidence rates were considered among those numbers abiding by the law, N-B distribution was checked in each of the 43 registries which had been previously analysed jointly.

To evaluate the abidance to N-B law, I presented the plotting of observed and theoretical frequencies and a set of five tests. An internal evaluation compared those registries whose data were more or less coherent with N-B. Results of each test were ordered to identify the most extreme values (the decile with the most unsatisfactory results). One of the registries had all the results of the applied tests in the worst decile. This result had a very low probability of occurring by chance. However, this outcome

is relative to the internal comparison in the analysed data set. All the registries had passed the CI5-X data checks, which assessed their quality.

The value of this paper is to have documented that cancer incidence rates abide by Newcomb-Benford law. Consequently, the first digit distribution of cancer incidence rates can be compared with the theoretical N-B one.

N-B can be quickly evaluated as it is based on the comparison of the observed frequency of the first nine digits of incidence rates with the theoretical ones. Observing a graphical plot of the two distributions may provide sufficient hints about rough anomalies. A more formal evaluation is made possible by specific commands included in many commercial (e.g. Stata, SAS) and open source (R) statistical software.

Traditional quality indicators (e.g., mortality/incidence ratio, DCO or Other&Unspecified sites) may vary widely among registries of accepted quality (e.g., CI5-XI for 'All sites but non-melanoma skin cancer' http://ci5.iarc.fr/CI5-XI/Pages/Indices_sel.aspx). Therefore, a comprehensive evaluation of data may be useful for preliminary screening and to address further thorough analysis.

N-B check could represent the first step in the process of data quality assessment. A violation of such distribution may suggest questionable data requesting a comprehensive and strict evaluation of traditional quality indexes.

The additional value of the paper is its demonstration that in the methodology on registry quality evaluation, there are still opportunities for innovation and improvement. Moreover, the continuous development of digitalisation has provided a growing number of datasets that are increasingly used by registries. N-B can equally be applied to them, once abidance of data is confirmed, for a preliminary evaluation of their global correspondence to what expected and early detection of surprising results. Some examples were already present in the paper by Benford. Benford found the first digit of the sizes of 3,259 populations in USA abided this law. Therefore, registries, which use populations (sex-, age-, time- or area-specific) as denominators

for computing rates, can apply Benford's law to examine them (Benford, 1938). Moreover, newly published applications such as those which evaluate waiting lists (Pinilla, 2018) or death counts (Daniels, 2017) may be useful for registries. There is a growing interest in the application of N-B law in the scientific environment (Hüllermann, 2017; Karathik, 2016; Lee, 2019) and it has been suggested for many fields of medicine and community health (Pollach, 2016).

8.3. A new index to improve informativeness

Once the quality of a dataset is considered checked and confirmed, it may be used for analysis. Traditionally registries have described incidence in a defined population showing, in tables and figures, overall cases and rates with cross-tabulations for a few variables: e.g., cancer topography, age, sex or place of residence. The previous is the usual way to describe the frequency of cancer in a specific area (Jensen, 1991).
The interest in the comparison of incidence in different areas or different periods in the same area has also made it necessary to consider the possible effect of age on incidence and the different age distributions in the populations under comparison, leading to age-standardisation, with both direct and indirect methods (Boyle, 1991).

Therefore, the publication of frequencies and crude rates meets local needs, and age-standardised rates (usually computed with the direct method) allow for comparisons over time and/or with other registries.

This way of presenting incidence has remained almost the same in the last decades (Jensen, 1991). For example, the frequency of cases, with crude and age-adjusted rates (Segi, 1960) has been chosen to summarise and compare incidence for cancer types within and between registries participating in CI5-XI (http://ci5.iarc.fr/CI5-XI/Pages/summary_table_site_sel.aspx). This information is complete and unbiased, and it provides handy data allowing for reliable comparisons.

However, in descriptive international studies incidence in a vast country with national coverage may be summarised with the same index used for a small national or even regional registry.

Paper 2 presents an entirely new application which could enrich the traditional presentation of ASR by adding a measure of internal variability. It does not request new variables since it is based on available data.

It may be applied to those registries which have incidence information also for sub-areas (e.g., regions in a country, provinces in a region). When only global ASR is provided, the proposed index should be handy to show further information straightforwardly.

This index is computed as the ratio between the difference among the highest and lowest sub-area specific rates ASR (r), and the overall one (R), r/R. The index is expressed as a percentage. It shows how much the maximum difference in the sub-areas is (between the highest and the lowest ASRs) respect the overall ASR (in percentage points). The smaller the index (0% when incidence is the same in all the sub-areas), the smaller the heterogeneity of incidence in the sub-areas represented by the summary ASR. Conversely, the higher the index is, the more significant the heterogeneity.

In paper 2 this index has been tested in the data available from the Nordic countries (NORCAN). NORCAN provides data from three decreasing geographical dimensions: supra-national (all Nordic countries), national (individual Nordic countries), and regional. The overlapping of 95% confidence intervals has been used for comparisons between ASRs. The application of the proposed index has quantified in a comprehensible way how much an incidence ASR may be more or less representative of the incidence ASRs in sub-areas.

As an empirical assessment, when r/R value is <10-15% it may be considered representative of homogeneous sub-areas. In fact, at the most, ASR incidence in one

of the sub-areas is 10-15% higher or lower than the other most extreme value. On the contrary, when r/R is >30%, the dispersion of incidence across sub-areal ASRs might not be negligible (at least one of the ASRs is 15% smaller and another 15% larger than the average summary).
This simple index would help users understand how much a summary incidence ASR significantly represents the incidence in underlying sub-areas, showing a measure of the amount of variability in incidence in the area.

The index has also been applied to incidence rates in the USA in 2014 for all cancers (except skin epitheliomas) available at U.S. Cancer Statistics Working Group (www.cdc.gov/uscs) for the USA overall, for four geographical regions, for two or more sub-regional areas, for 51 States and five metropolitan areas. The 95% confidence intervals of sub-areal ASRs did not always overlap with that of the broader area of which they were part. In this setting, the r/R index also confirmed its use in providing a measure of the amount of underlining variability in incidence related to a summary ASR (Crocetti 2017a).

In the example presented in paper 2, the index was applied to countries of a somewhat different dimension and more homogeneous regions (except for Iceland). The variability in incidence may also be related to the dimension of the sub-areas considered and may increase when they are small. This topic may be a problem when the index is computed for more registries based on sub-areas of different dimensions. In this case, the value of the index may not only show the amount of heterogeneity but also the presence of different geographical granularity, with at least one of them with a value far from the summary ASR.

The index may be used as a measure of macro-area variability in incidence for a specific registry and different registries when sub-areas have almost the same magnitude in the size of the resident population (and level of incidence).

The addition of a measure of variability would considerably improve the use of the index.

The main assets of the method are simplicity of computation, the fact that it grounds on already available measures and its ease of interpretation. The simplicity of computation and interpretation have been preferred to more formal statistics (e.g., analysis of variance, extreme quotient) (Gumbel and Kerry, 1950; Fisher, 1921) to make such approach suitable for all the people working in registries or for those who use registry data and who may not have a background in statistics.

Such index has already received some appreciation. In fact, the method was presented in recent scientific meetings of cancer registries where it was used to compare internal variability in registries located in countries where a language derived from Latin is spoken (Contiero, 2018), in some registries participating in a European project (Martos, 2018) and to discuss the effect of the choice of different standard populations (Rashid, 2018).

The r/R index could complement ASR and standard error when incidence is presented. In this way three kinds of information would be provided: a comparable level of incidence in a specific area (ASR), the precision of such estimate (standard error) and a simple measure of heterogeneity in incidence in macro sub-areas (r/R index). If grounded on reliable data, this information is essential.

This index of heterogeneity could be extensively applied in those International collaboration initiatives, which include dozens or hundreds of registries. For example, in CI5 many countries have national coverage. Their incidence is presented as a national average only. In case also regional estimates would be available although not published, the computation of the r/R index would provide additional information on the homogeneity/heterogeneity of incidence within the country.

Moreover, the same approach used here for incidence could be exploited and applied to prevalence or survival (e.g. Eurocare, Concord) or other epidemiological metrics.

The proposals for new methods presented in papers 1 and 2, address quality evaluation and data visualisation. The next examples take a step forward and show how registry's data potentiality may be exploited to increase their informativeness, interest and usefulness.

8.4. Which variables to analyse?

The information that a registry traditionally collects is somewhat limited (e.g., cancer topography, morphology, sex, date of incidence, basis of diagnosis and age) and based on the requests of big International projects (for example CI5, Concord, Eurocare).

Adding any other variable implies a certain amount of extra time and resources as well as additional validity, completeness and comparability evaluations, which may request specific methods. Moreover, this implies an increased workload which does not improve overall timeliness.

Therefore, while the critical information provided by the registries fulfils big consortia's necessities, it is scarcely fascinating for other stakeholders (e.g., oncologists).
Further, big projects involving registries from different countries and even continents, need to find out the minimum data set, which allows the majority of registries to participate.
Besides, increase the workload by coding more or more detailed variables may jeopardise the compliance with big projects' deadlines.

It is a recursive process that limits the possible use of registries.

One of the bottlenecks of the registration procedure is the manual handling of traditional and new data. Registries should exploit the implementation of software for data-mining from sources for the preliminary harvest of information, involving personnel only at the second level of revision (Crocetti, 2005).

In conclusion, registries data specific for cancer topography, morphology, and a few other demographic variables are appealing for those who want to compare incidence or survival in the world but have raised limited interest among clinicians, including those working in the same area of the registry.

The risk for registries is to produce high-quality data with no value outside their self-referential circle.
The use of data should not be independent of the concept of quality of registry data applied until now. On the contrary, it should urge those who work in the registries' world to reconsider it.

If high-quality data are not used outside the limited world of registries these complete, valid and comparable datasets (on average timeliness is not an asset for registries) (Zanetti, 2015) are not recognised as valuable and appreciated by many others who should use such data.
Since registries use clinical data, clinicians should be expected to be partners in the analysis and interpretation of registry data. Regrettably, this is frequently not the case. In fact, in a comprehensive evaluation performed in 2009-2013, the majority of registries (65%-75%) declared they were not active in the range of cancer research areas but just provided incidence and survival data to others for analysis (Coebergh, 2015).

However, only if registries increase their capability to bridge the gap with the clinical world by providing data valuable for the clinical interpretation of epidemiological patterns, will they become an active part of their health system.
Unfortunately, in most cases, these surveillance systems are isolated and do not collaborate with all those for whom data would be available. Therefore, the collaboration with all the possible stakeholders is among the future challenges recommended to cancer registries (Zanetti, 2018).

Paper 3 provides an example of a fruitful collaboration between epidemiologists and clinicians (in this particular case, dermatologists). Such collaboration arose from the

mutual interest in understanding and integrating two professionalisms: the values of population-based reliable information and the relevance of some clinical variables for patients and clinicians.

In this experience, dermatologists extended their interests outside their own department/hospital, and epidemiologists entered into the real world.

In this study, which is the first on this specific topic published in Italy, the method of cancer registration and established clinical practices have interacted effectively.
The process has implied an exhaustive use of one of the official sources of information for registries: the pathology report. The diagnosis of skin melanoma, although suspected by visual inspection or enhanced by dermoscopy examination, requires a microscopic confirmation by the pathologist. Pathologists provide three essential sorts of information which steer the further therapy of skin melanoma: invasiveness, its extent in millimetres and morphology.

This information, included in the registries, has made epidemiological analysis more effective in the interpretation of skin melanoma clinical patterns. Although the increase over time of *in situ* and thin lesions may be related to an increased awareness of skin moles and a grown bioptic aggressiveness of dermatologists raising the concern of over-diagnosis, the concomitant increase of ASR of thick (and deadly) melanoma has confirmed a real epidemic in central Italy.

Some sub-site analysis of melanoma onset and type of morphology of melanoma and patients' age and sex have enriched the information provided to dermatologists and improved their interpretation of clinical patterns and identifying groups at risk.

The same approach has recently been applied to evaluate the trend of skin melanoma in Europe, based on the datasets of 18 registries (Sacchetto, 2018) confirming, after eight years from the publication of paper 3, the appropriateness of the variables and methods used.

However, much work is still necessary. A recent study aimed at evaluating stage-specific survival differences in Europe for seven selected cancers showed that between 15% to 22% of registries did not possess the information on stage (the range among European regions and cancer sites was 8-28%) (Minicozzi, 2018). In the same project stage evaluation in 15 solid cancers in Europe confirmed that stage was often incorrectly assigned or missing, especially among older people. Such results spoke out for the need for improving data collection and coding (Minicozzi, 2017).

Therefore, also, a quantitative aspect should belong to the quality of data.
High-quality data should imply a varied and rich offer of those variables relevant in defining and describing the diagnosis, treatment and prognosis of patients (Zanetti, 2018). More high-quality clinical variables are essential to root registries in the health system and to appoint them 'agents of change' as it has been recently hoped for their future (Zanetti, 2018).

8.5. The application of a wider variety of statistical methods could help the exploitation of data informativeness

The collection and analysis of data by registries represent a continuous process providing substantial and reliable information: clear identification of prevalent cases (challenging to be detected especially at the beginning of the registration activity), precise evaluation of the timeliness of reaching completeness (especially relevant for more recent years of incidence) and the identification of multiple primary cancers in the same person. Moreover, the availability of high-quality data for an extended period allows reliable time trend analysis.

Time trend evaluation in cancer incidence represents an invaluable addition to the information on incidence level; any difference over time or among groups of people may suggest possible inputs for analytic studies.

The analysis of changes in incidence levels over time represents a fundamental tool for community health. By monitoring such indicators one can evaluate the effects of

health programmes (e.g., awareness-raising campaigns on skin mole changes), the introduction of novel therapies or new diagnostic instruments (e.g., dermoscopy) as well as changes in lifestyles and environmental exposures (e.g., recreational sun exposure). Furthermore, time trends allow documenting the effectiveness of prevention measures at the population level.

In paper 4, we applied an age-period-cohort model which showed that the best model fitting the analysed melanoma incidence data included age and drift. Such drift (linear temporal effect) showed that in every 5-years or cohort analysed there was an increase of incidence by about 37% in each 5-year age-group. This paper was the first one in Italy to show that melanoma incidence was increasing in all the age groups in different periods and age-specific incidence rates increased in all the analysed cohorts, including the youngest.

Paper 4 presented the analysis of clinically relevant indicators in which registry data is not limited to a mere counting of cancer cases.

The role of CRs not only is to provide high-quality data but also to make this data utilized at best by as many stakeholders as possible. Reliable data must be used, involving all possible methodological approaches. In paper 4 the knowledge of melanoma clinics and epidemiology indicated that over the latter decades cultural change which caused an increase in recreational exposure to UV and the increase in early diagnosis has acted more as a period than as a cohort effect. Moreover, the availability of a high-quality series of data can constitute a basis for reliable projection for more recent periods. In fact, in the same paper, we used the high-quality data collected in the Tuscan registry from 1987 to 2001 to estimate the expected burden of newly diagnosed melanomas in 2002-2006. To be specific, we adopted the non-linear model proposed by Dyba et al., (Dyba, 1997), including the effect of age and age-specific temporal trends.

Also, the economic burden related to diagnosis and treatment of specific cancers can be calculated based on observed data and projected in the future to estimate the

amount of the necessary investment which has to be planned to fulfil resources adequate to the expected burden. For example, this evaluation has been performed in the US, applied to colorectal cancer. Moreover, costs had detailed for three different periods of the clinical care: diagnosis and first treatment continues phase of diseases and last year of life (Yabroff, 2007).

In conclusion, not only does trend analysis provide further insights on the crucial determinants of changes, but if based on high-quality data, it can be used to predict future burdens. Reliable future projections enable to monitor the effectiveness of preventive programmes and to determine future needs. For example, recently published estimates for the UK have drawn attention to the need to contrast smoking (women), alcohol, overweight and obesity, HPV and hepatitis infections to dampen the growth of those cancers which demonstrate an increasing tendency (Smittenaar, 2016). Similar conclusions came from Canada based on projections made available for tailoring resources and services' planning on the on-going demographic changes and to the increase of specific cancers. Besides, such projections suggest prevention as a useful tool to control the future burden (Xie, 2015).

The production of short-terms predictions should be extensively exploited by registries also to fill the timeliness gap between clinical events and availability of incidence estimates and move high-quality data series into recent years.

8.6. Communication of data contents

The knowledge in medicine evolves when the experience acquired on an individual patient extends to a similar group. This process also applies to survival for which population-based registries provide for clinicians, citizens and patients average measures at population-level for specific cancer, in specific sex or age-group or stage or even in more detailed aspects.

However, every day, any oncologist visits patients with specific genetic/phenotypic personal disease who want to know details about their prognosis.

If that patient was a man, middle-aged, ex-smoker and with regionally diffused cancer, the answer of the oncologist, based on the most updated information collected from high quality research published in peer-reviewed scientific journals, could be 4.3 times worse than for a women, younger than 30 years, who has never been a smoker and is affected by the same cancer but confined to the site of onset.

The previous example shows how the hazard risk is typically presented, applying the worldwide appreciated and adopted the Cox model (Cox, 1972).
The answer would be correct, evidence-based, but presumably meaningless for that patient who would leave the oncologist's office more confused than informed.
The output of a Cox model is reliable when the necessary assumptions are respected, but interpretation is complicated, even for health-care personnel.

Paper 5 dealt with this topic which links methodology and communication, comparing the prognostic effect of several variables with both Cox model and Classification and regression trees (CART) analysis in a high-quality registry series of incident skin melanoma.
CART possesses several advantages, starting from its being essentially nonparametric, and therefore requiring no assumptions about the underlying distribution of the data to analyse (Lemon, 2003 and 2012; Kraemer, 1992).

Moreover, in my opinion, the friendly outcome remains the most relevant asset, considering the need to comprehend and communicate in a comprehensible way information concerning the life expectancy of patients.

Patients are one of the most important and engaged stakeholder groups with whom registries have to collaborate. The relationship and communication between doctors and patients remarkably changed in the last decades. In fact, over time, the right of patients to be informed and the duty of doctors to inform them have been affirmed (Buckmann 1996). Patients and doctors have become responsible for shared accountability for trustworthy information (Schain, 1980).

Moreover, cancer patients' associations are advocating the right of patients to be involved in any decision which affects their lives (Souliotis, 2018). “No decision about me, without me” is the norm in many countries and hopefully all the others will follow suit.
Registries have to respond to patients' questions, and many collaborations are still going on to fulfil this task. For example, the network of Italian registries has provided a measure of time to cure for cancer patients (Dal Maso, 2014).

Paper 5 moved my research on making registry data more informative and useful for all possible stakeholders interested in survival by improving the way to measure and communicate the prognosis.

According to my knowledge, this type of analysis (CART) was the first-ever made in Italy on this topic. It involved two high-quality registries.
CART analysis proved able to identify homogenous groups of patients with the same levels of the variables analysed who had a defined observed survival. In this manner, patients can receive information on their specific average prognoses as absolute values that do not require complex computations.

Unfortunately, despite the greater straightforwardness of CART in comparison with Cox, CART is not widely used. A search for the terms "Classification and regression tree analysis" AND "Cancer" done on 17th May 2018 in Pub Med identified only 97 papers. Conversely, search for "Cox" AND "Cancer" highlighted 47,402 papers. Something hampers a more general introduction of CART analysis in the scientific community.

Statistically, by constructing subgroups directly on the covariates, the CART algorithm satisfies the objective of finding both independent prognostic factors and prognostic subgroups. Apart from some computational weakness (Marshall, 2001; Venkatasubramaniam, 2017), this methodology has several advantages, and its application suggested (Lemon, 2003). In contrast to traditional regression methods (e.g., Cox proportional hazard regression) which compute a prognostic index as a weighted average of the covariates, CART constructs groups based on logical

combinations of patient's characteristics. The results of the Cox proportional hazard model may be challenging to implement into clinical practice. Although the Cox model enables to predict the survival probability associated with a specific variable, it is always challenging to interpret or predict a patient's cumulative risk for a given set of prognostic factors. In daily medical practice, patients habitually have many prognostic factors, especially when there are interactions involved (Chang, 2012).

Hence, providing a useful and informative risk group definition for practical use is a tough task.

There are several examples in literature on the uses of CART analysis for risk assessment at the bedside (Abu-Rustum, 2008; Barlin, 2013; Diessner, 2016; Egger, 2013; Fernández, 2017; Hess, 1999; Kim, 2012; Kobayashi, 209; Jain, 2017; Lamborn, 2004; Langendijk, 2005; Lu, 2016; Xu, 2019; Yuan, 2017).
On the other hand, CART analysis raised criticisms for being inherently unstable (Rokach & Maimon, 2007; Protopopoff, 2009; Su, 2011). Minor changes in data can drastically alter a tree's appearance and its interpretation if the tree is unmanaged with caution. When a split changes, all subsequent splits change, too. Therefore, each optimal partition depends on the path already taken through the tree (Crichton, 1997).
Moreover, also the objections stated by Marshall (Marshall, 2001) are correct, but they apply exclusively in a blinded situation in which no previous scientific evidence is available.

In conclusion, results of paper 5 let suppose CART as a method capable of summarising evidence on the prognosis in a way which may be easily understood by both clinicians and patients. However, such an impression has to be tested and quantified in proper studies.
The need for a clear comprehension of cancer prognosis has also been addressed using nomograms. For example, SEER has developed a prognostic nomogram for prostate and colorectal cancer based on tumour characteristics, age, stage, gender and comorbidity (Feuer, 2014). The nomogram provides information on the chance of

dying of cancer and other causes, complementing the traditional survival information (Feuer, 2014). The use of nomograms goes in the direction of providing information on individualised prognosis and of helping clinicians to select the adequate treatment for a specific patient (Caulfield, 2018).
A more extensive use of CART in prognosis communication still requires further analysis which hopefully will be addressed in the future research on registries progression (Zanetti, 2018).
Moreover, also the information flow from registry to oncologists and patients has to be further evaluated to understand how, where, and with whom registry may best act to improve the comprehensibility of produced prognostic indexes.

The role of paper 5 was not to belittle Cox's model and its undisputed capability to handle uncertainty in survival estimates but to emphasize the still unmet need of an explicit, straightforward and comprehensible communication of epidemiological measures to not statisticians.

The scientific community cannot ignore this assignment.
Otherwise, high-quality data, correctly analysed and full of informativeness, would turn out to be useless when they cannot be adequately comprehended.

This issue belongs to the broader topic of communication in oncology. It is incredibly complex and implies ethical, cultural and legal issues (Gordon, 2003). The communication of risk - e.g., prognosis - is even more complicated, as it presupposes uncertainty (Spiegelhalter, 2011). However, recommendations are becoming available (Freeman, 2017).

Registries could engage patients and oncologists also in this topic by jointly planning and performing studies aimed at evaluating the comprehensibility of the prognostic message provided by different survival measurement methods. In therapies, patients' preferences are more and more carefully considered. Therefore, also the comprehensibility of prognostic statistics based on different tools should be formally compared and evaluated by users.

Approaching the last of the papers presented in this thesis, it is worth recalling that the activity of a registry starts from clinical (and administrative) data. The research presented here tries to link the production of data to the improvement and extension of their use, closing the circle with epidemiological data aimed at improving cancer care.

8.7. Cancer registry: a proactive part of the health system

In previous papers (Paper 1-5), the quality of the registry's data has been highlighted, as well as traditional and innovative tools to assess it. Quality is a prerequisite for any epidemiological evaluation, necessary to guarantee reliable inferential estimates.

The second discussed issue underlines that data not only have to be of excellent quality but they, also, have to bear a relevant, informative message.

Finally, in paper 6, also the last leading theme is presented jointly with the previous two: the usefulness of registry data.

Therefore, all three themes coexist in this paper, completing the circle.

Paper n. 6 deals with skin melanoma with recommendations for evaluating patterns of diagnosis and care. In Tuscan regional recommendations are evaluated here based on 13 quality indicators.

The paper involved not only the registry but also a large working group of the Tuscan regional oncological institute made up of all the experts from disparate professional fields in skin melanoma knowledge and care (dermatologists, pathologists, surgeons, oncologists and epidemiologists), academic institutions and community hospitals.

In paper 6 the registry entered the real world with full recognition and appreciation by clinicians and policymakers. It provided a fruitful contribution on evaluation of

population-based sources of information and in the production of objective and reproducible indicators of care.

The process started from questions on melanoma care, raised by clinicians and policymakers, aimed at evaluating the level of compliance of services to specific recommendations ('usefulness'). The registry defined and produced appropriate valuable indexes ('informativeness') based on population-based high-quality data ('quality').

Thanks to this multi-professional activity, expected and observed procedures and treatments for skin melanoma were compared at the regional level, and reliable information was provided to correct weaknesses.

Registries should become more actively involved in the evaluation of quality of care, and national and international registry associations should agree on standardised cancer-specific indexes, aligned with their periodic updates. The availability of common and comparable indexes would make useful comparisons possible.

9. Conclusion

The six scientific papers presented in this thesis by Prior Published Work aimed at exemplifying how CRs have contributed, and how their contribution could increase, in the field of cancer epidemiology and the use of data for a more appropriate and effective allocation of resources.

The CRs data represents not only the output of a complex process but also the result of technical and strategic decisions. All these steps are inter-related.

In the present thesis, I describe how my research has considered this process as a whole, a *continuum* in which any factor of novelty (new methods, new uses, new users) positively affects the following ones.

Registries activity represents a circular process that starts from clinical data and ends with epidemiological indexes useful to improve the related care.

The research activity reported in the six presented papers concerned the full process.

Different types of data feed registries. These data are checked, condensed, integrated, evaluated and at the end used for producing other data in the form of epidemiological indexes to apply by different users for various purposes.
The informativeness of such indexes, based on high-quality data, has to be continuously reconsidered and customised to stakeholders to improve their understandability and usefulness for users with different skills.
If registry data are underused, the money the community pays for supporting such activity is partly wasted. The larger the number of users, the higher the registry's usefulness for the community and the efficiency of resource investment.

High-quality data, if customised to the objectives and professions of different readers are more clearly understood, perhaps appreciated and hopefully used.

The process of quality evaluation has probably still rooms for improvement. The application of Newcomb-Benford law shows that the topic is still open to innovation, using approaches appropriate to the 'big data's epoch'.

However, the concept of quality should not be confined to data but expanded to marketability and use of the information. If high-quality data are not appealing and remain unused, they represent merely a waste of public resources typically limited.

Registries should not work merely for epidemiologists and registrars. This self-referential process distances even more registries from the real world.
Quality is just the starting point of a process that must continue until data are appropriately used as much as possible by all potential stakeholders.
Therefore, those who produce registries data and are aware of their value have to bridge the gap and engage stakeholders.

The summary information traditionally provided by registries (age-standardised incident rates) may be made more informative for all types of readers, including a measure of internal variability whenever this is possible. The r/R index represents an innovative and straightforward change which is objectively useful to improve informativeness.

Registries must move closer to clinicians and patients to gain their collaboration. For this reason, registries have to include as much clinical data as possible to make epidemiological indexes interesting for clinicians and useful for patients and policymakers as well. The on-going digitalisation of health systems databases and the development of methods for data-mining (Crocetti, 2005) make this possible. Registries have to consider the current paths made available by technology.

Collaboration with clinicians joins epidemiological and biostatistical expertise of registries with the real clinical world.

The experience of registries in data analysis should be exploited as much as possible using the potentiality of methodology, for example, to improve the timeliness of registries. Time trends exemplify how registries data could be employed to generate predictions, i.e., identify the future impact of the disease and the expected demand on health services. Such projections are quite relevant for policymakers when they plan health care strategies.

Patients should expect registries to provide relevant and reliable information. The example comparing Cox and CART analysis reinforces a point in between informativeness and usefulness. Registrars could involve both stakeholders (patients and oncologists) in the attempt of measuring the effective comprehensibility of the two methods and gather evidence for future decisions.
The scientific community should debate about the balance between statistical robustness and comprehensibility of statistics.

The duty to demonstrate the immense value of registry data concerns registrars and cannot be the responsibility of readers.
Registries data must accomplish their role of monitoring tools for the health system and the whole population, ensuring the correspondence between clinical recommendations and current clinical practice. Registries are surveillance services embedded within health systems, and data produced by the same systems feed them. Therefore, they are in a privileged position for evaluating the quality of cancer care, highlighting critical challenges and proposing advice for their improvements.
Future health policies need evidence on the clinical points of strength and weakness, based on reliable data.

The data input to registries is the footprints left by cancer patients in the databases of the health systems during their diagnostic and therapeutic experience. Data produced by registries must return to patients in the form of reliable evidence contributing to the development of the knowledge of cancer and improvement of the quality of cancer care.

Therefore, cancer registration may be imagined as a circular process that starts and ends with data, both of which represent the real experience of patients.

The primary aims of this research have represented the improvement of quality and usefulness of registry data using skin melanoma as a common thread. However, the leading concepts of the thesis apply to registration as a whole, that means that they may be extended to any tumour.
This thesis addresses some relevant aspects, but others have been partly or left aside. For example, among the former, the need to evaluate the quality of primary sources. Examples of not addressed topics may be the right of citizens and patients to be more informed about the role of environmental (e.g. UV), recreational (e.g. tanning beds) and professional (e.g. outdoor workers) exposure, as other relevant topics, e.g., rules and problems with coding, extent of the use of algorithms for automatic decisions, have been left in the background.

However, I hope that the general message that the maximum potential of registry data must be recognised, used and exploited has been clearly stated.

10. Recommendations

The topics discussed in the present thesis may be summarised in the form of some recommendations which could contribute to improving the role of registries in the short term:

- Traditional methods for quality of data evaluation are well established and reliable. However, a formal process of quality assessment with clear cut-offs for each indicator is still lacking. International registries associations are expected to agree on guidelines on such topics. Moreover, possible violations of Newcomb-Benford law could be included in the process of data quality assessment with the recommendation for registrars to devote particular attention to traditional indexes in case of anomalies. Also, the traditional sources of information (e.g., hospital admission forms, pathology reports) and any other digital data which feed the process of registry should be tested, first for abidance and then for violation of such law.

- The proposed index of internal heterogeneity (r/R) could be routinely added in International projects which include dozens or even hundreds of registries, as well as in local publications. This index would help readers to understand how far the summary index - in the example for incidence - represents sub-areas of each registry. Further applications of the same approach could also be investigated and applied for other measures (e.g., survival, prevalence) for which identical question (Internal homogeneity or heterogeneity?) may still be open.

  Each registry individually and networks of registries and scientific societies as well should strengthen the collaboration with clinicians, to join forces (and competences) to provide the most useful cancer-specific analysis. Therefore, each cancer registry has to make available data on those variables which currently address diagnosis and treatment.

- Lack of timeliness is one of the critical points of weakness of registries, and the weak interest in data may partly be due to this. Short-term projections could bridge the gap and provide, e.g., policymakers, with reliable estimates of the burden of cancer, necessary for health policy planning.

- The maximisation of the value and utility of the data could be ensured:
  - A more extensive involvement of stakeholders in planning the agenda of a registry is vital.
  - Cancer registries should establish active and continuous collaborations with patients. Cancer patients may be extremely supporting, lobbying the need for registries. Registrars should plan with them some patient-centred research (e.g., infertility, late effect of treatments, economic effects of cancer, time to cure). Communication of the risk may be one of the issues for which registrars could test and propose technical alternatives to mainstream approaches.
  - Healthy people have to be also part of the planning of a registry's activity. There are specific topics: e.g. environmental pollution, local specific exposures, for which a registry is expected to provide reliable population-based evidence.

- Collaboration with cancer patients, clinicians and policymakers will exploit the role of registries within the health system to monitor the appropriateness of provided care and focus on local weakness.

- The last recommendation is to remind the innovative paradigm proposed in the present work: the interrelationship between data production, data analysis and data dissemination/interpretation. These represent phases of a continuous process and any improvement in each step affects all the others.

11. Future research

The future research on cancer registration (and subsequently on cancer epidemiology) will be conditioned and inspired by the ongoing technological development, which will increasingly affect the availability of information and machines' capability of managing it.

- In this manner, it will be possible to exploit automation in almost all phases of cancer registration: collection of sources of information, data selection, linkages, identification of possible cases, variables' collection, coding, quality evaluation and production of epidemiological measures. Systems for data-mining (e.g., automated data extraction and classification, artificial-intelligence-assisted procedures for data recording) could be exploited to gather more information from digitalised sources. In each specific assessment cost/effectiveness should be compared between manual extraction and automated procedure, considering timeliness, accuracy, reproducibility and resources availability. The involvement of artificial intelligence is supposed to increase in the next future, but this will be a gradual process. This change would move human resources from data collection and coding to their quality evaluation, shortening the time for information harvesting and valorising professionalism. Automation may offer unprecedented opportunities for more value-added work, shortening the time of production and expanding the deliverables.

This new scenario will require a radical rethinking of well-established principles. On the other hand, it will offer an opportunity to extend the capability of registries to produce more information, for more stakeholders, in a shorter time.

Presumably, the concept of quality of data is one of the paradigms that registrars will have to reconsider.

Until now, the quality of data is evaluated at the end of the full registration process. Once quality is considered excellent enough, data analysis is performed, for any possible purpose. Presently we consider quality in absolute terms.

This type of approach contributes to the lack of timeliness of registry data. Indeed, data are unused at all until the full registration process is accomplished. The side effect is that even high-quality data on clinical events that had occurred several years before may not raise the interest of policymakers, patients or clinicians.

With more automation in the registration process, we could think of diverse types of data with a different timing and a diverse quality.

The latter could be called - paraphrasing Zygmunt Bauman (Bauman, 2000) - a liquid quality.
A high quality, appropriate for producing reliable information for a specific aim, but possibly different from the high quality necessary for another purpose.

The quality of data has a cost. It requests resources and time. Why should a public health system pay for the highest possible quality for data which may appear out-of-date for some of the possible purposes when published and are not valuable and useful for clinicians and patients?

Future research should adopt the quality for specific purposes. The general aim will be to employ data with the best necessary quality for reliably answering specific questions.
For example, a study involving an automatic extraction of data from hospital admissions produced an equal number of incident breast cancers as the traditional cancer registration (Ferretti, 2009). However, the cases were not the same. Therefore, the quality of the information produced with this simplified approach was perfect for estimating the burden, but not enough for identifying the people. At that time, the authors of the paper, including myself, gave more relevance to the negative result, reinforcing their belief on the conventional approach to registration (Ferretti, 2009).

Regrettably, this delayed the opportunity for registries to directly provide - with sufficient quality – information on the burden before the full cohort of patients had been carefully identified. However, research on decisional algorithms is going on and their capability to correctly identified patients improves thanks to the availability of a gold standard (registry) (Rasmussen, 2018; Wu, 2018; Goldsbury, 2017; Kemp, 2013). Therefore, their development has to be carefully monitored to utilize them when a registry is unavailable (Goldsbury, 2017; Kemp, 2013) but also to consider their implementation in the registry's process.

Registries have to respond to changing demands trying to use as much as possible the potentiality offered by automation. At the same moment, they have to enhance their professionalism on data, on their quality and interpretation. These precious competencies have to complement the expected speed from automation in a good collaboration.
Conversely, there is the risk that registries will be replaced by other systems capable of producing a flow of information at a remarkable pace, but without reliable expertise on the knowledge and interpretation of its content.

Registrars have both the interest and the responsibility of steering this technological change.

12. REFERENCES

Aase A, Bentham G. (1994). The Geography of Malignant Melanoma in the Nordic Countries: The Implications of Stratospheric Ozone Depletion. *Geografiska Annaler: Series B, Human Geography*. 76(2), 129-139.

Aase A, Bentham G. (1996). Gender, geography and socio-economic status in the diffusion of malignant melanoma risk. *Soc Sci Med*. 42(12), 1621-37.

Abu-Rustum NR, Iasonos A, Zhou Q, Oke E, Soslow RA, Alektiar KM, Chi DS, Barakat RR. (2008). Is there a therapeutic impact to regional lymphadenectomy in the surgical treatment of endometrial carcinoma? *Am J Obstet Gynecol*. 198(4),457.

Adamo M, Dickie L, Ruhl J. (January 2018). SEER Program Coding and Staging Manual 2018. National Cancer Institute, Bethesda, MD 20892.

Aguiar JP, Cardoso Borges F, Murteira R, Ramos C, Gouveia E, Passos MJ, Miranda A, da Costa FA. (2018). Using a cancer registry to capture signals of adverse events following immune and targeted therapy for melanoma. *Int J Clin Pharm*. 40(4):852-861.

AIOM, Italian medical oncologic association. (2010). Guidelines: melanoma, 2010. Available at http://www.aiom.it/area + pubblica/area + medica/ prodotti + scientifici/linee + guida/Archivio/1,333-333 [Accessed 15 April 2013].

Airtum working group. Italian cancer figures, report 2009. (2009). Cancer trend (1998–2005). *Epidemiol Prev*. 33(4-5 Suppl. 1), 54–55.

Aitken JF, Janda M, Elwod M, Youl PH, Ring IT, Lowe JB. (2006). Clinical outcomes from skin screening clinics within a community-based melanoma screening program. *J Am Acad Dermatol.* 54(1), 105–114.

Aitken JF, Youlden DR, Baade PD, Soyer HP, Green AC, Smithers BM. (2018). Generational shift in melanoma incidence and mortality in Queensland, Australia, 1995-2014. (2018). *Int J Cancer*. 142(8), 1528-1535.

Ajiki W, Tsukuma H, Oshima A. (1998). Index for evaluating completeness of registration in population-based cancer registries and estimation of registration rate at the Osaka Cancer registry between 1966 and 1992 using this index. *Nihon Koshu Eisei Zasshi.* 45(10), 1011–7.

Allemani C, Harewood R, Johnson CJ, Carreira H, Spika D, Bonaventure A, Ward K, Weir HK, Coleman MP. (2017). Population-based cancer survival in the United States: Data, quality control, and statistical methods. *Cancer*. 123(Suppl. 24), 4982-4993.

Amato L, Fusco D, Acampora A, Bontempi K, Rosa AC, Colais P, Cruciani F, D'Ovidio M, Mataloni F, Minozzi S, Mitrova Z, Pinnarelli L, Saulle R, Soldati S, Sorge C, Vecchi S, Ventura M, Davoli M. (2017). Volume and health outcomes: evidence from systematic reviews and from evaluation of Italian hospital data. *Epidemiol Prev*. 41(5-6 Suppl. 2), 1-128.

Amayiri N, Bouffet E. (2014). Pediatric neuro-oncology. In: Stefan DC, Rodriguez-Galidndo C, eds. Pediatric neuro-oncology in countries with limited resources: a practical manual. New York: Springer Science & Business manual. New York, pp. 365-76.

Amersi F, Morton DL. (2007). The role of sentinel lymph node biopsy in the management of melanoma. *Adv Surg*. 41, 241–256.

Andersen MR, Storm HH; Eurocourse Work Package 2 Group. (2015). Cancer registration, public health and the reform of the European data protection framework: Abandoning or improving European public health research? *Eur J Cancer*. 51(9), 1028-38.

Andreano A, Anghinoni E, Autelitano M, Bellini A, Bersani M, Bizzoco S, Cavalieri d'Oro L, Decarli A, Lucchi S, Mannino S, Panciroli E, Rebora P, Rognoni M, Sampietro G, Villa M, Zocchetti C, Zucchi A, Valsecchi MG, Russo AG; OSSERVA Working Group. (2016). Indicators based on registers and administrative data for breast cancer: routine evaluation of oncologic care pathway can be implemented. *J Eval Clin Pract*. 22(1), 62-70.

Andreano A, Ansarin M, Alterio D, Bruschini R, Valsecchi MG, Russo AG. (2018). Cancer of the head and neck: a set of indicators based on register and administrative data. *Acta Otorhinolaryngol Ital*. 38(1), 13-23.

Araghi M, Fidler MM, Arnold M, Jemal A, Bray F, Soerjomataram I. (2018). The Future Burden of Colorectal Cancer Among US Blacks and Whites. *J Natl Cancer Inst.* 110(7):791-793.

Armitage P, & Doll R. (1954) The age distribution of cancer and a multi-stage theory of carcinogenesis. *Br J Cancer*. 8(1), 1–12.

Arnold M, Holterhues C, Hollestein LM, Coebergh JW, Nijsten T, Pukkala E, Holleczek B, Tryggvadóttir L, Comber H, Bento MJ, Diba ChS, Micallef R, Primic-Žakelj M, Izarzugaza MI, Perucha J, Marcos-Gragera R, Galceran J, Ardanaz E, Schaffar R, Pring A, de Vries E. (2014). Trends in incidence and predictions of cutaneous melanoma across Europe up to 2015. *J Eur Acad Dermatol Venereol*. 28(9), 1170-8.

Balch CM, Buzaid AC, Atkins MB, Cascinelli N, Coit DG, Fleming ID, Houghton A Jr, Kirkwood JM, Mihm MF, Morton DL, Reintgen D, Ross MI, Sober A, Soong SJ, Thompson JA, Thompson JF, Gershenwald JE, McMasters KM. (2000). A new American Joint Committee on Cancer staging system for cutaneous melanoma. *Cancer*. 88 (6), 1484–1491.

Balch CM, Buzaid AC, Soong S-J, Atkins MB, Cascinelli N, Coit DG, Fleming ID, Gershenwald JE, Houghton A Jr, Kirkwood JM, McMasters KM, Mihm MF, Morton DL, Reintgen DS, Ross MI, Sober A, Thompson JA, Thompson JF. (2001). Final version of the American Joint Committee on cancer staging system for cutaneous melanoma. *J Clin Oncol.* 19 (16), 3635–3648.

Balch CM, Gershenwald JE, Soong S, Thompson JF, Atkins MB, Byrd DR, Buzaid AC, Cochran AJ, Coit DG, Ding S, Eggermont AM, Flaherty KT, Gimotty PA, Kirkwood JM, McMasters KM, Mihm MC Jr, Morton DL, Ross MI, Sober AJ, Sondak V. (2009). Final version of 2009 AJCC melanoma staging and classification. *JCO*. 3, 6199–6209.

Balch CM, Soong S-J, Atkins MB, Buzaid AC, Cascinelli N, Coit DG, Fleming ID, Gershenwald JE, Houghton A Jr, Kirkwood JM, McMasters KM, Mihm MF, Morton DL, Reintgen DS, Ross MI, Sober A, Thompson JA, Thompson JF. (2004). An evidence-based staging system for cutaneous melanoma. *CA Cancer J Clin*. 54(3), 131–149.

Barlin JN, Zhou Q, St Clair CM, Iasonos A, Soslow RA, Alektiar KM, Hensley ML, Leitao MM Jr, Barakat RR, Abu-Rustum NR. (2013). Classification and regression tree (CART) analysis of endometrial carcinoma: Seeing the forest for the trees. *Gynecol Oncol*. 130(3), 452-6.

Bauman Z. (2000). Liquid modernity. Polity Press, Cambridge, UK.

Baumert J, Schmidt M, Giehl KA, Volkenandt M, Plewig G, Wendtner C, Schmid-Wedntner MH. (2009). Time trends in tumour thickness vary in subgroups: analysis of 6475 patients by age, tumour site and melanoma type. Melanoma Res. 19, 24–30.

Benford F. (1938). The law of anomalous numbers. *Proc Am Philos Soc*. 78, 551–72.

Beral V, Peto R. (2010). UK cancer survival statistics are misleading and make survival look worse than it is. *BMJ.* 341, c4112.

Berg JW. (1996). Morphologic classification of human cancer. In: Schottenfeld D & Fraumeni JF Jr. Cancer Epidemiology and Prevention, 2nd edition, Chapter 3 of Section 1: Basic Concepts. Oxford, New York, Oxford University Press, 1996, pp. 28-44.

Berger O, Grønberg BH, Loge JH, Kaasa S, Sand K. (2018). Cancer patients' knowledge about their disease_and_treatment_before, during and after_treatment: a prospective, longitudinal study. *BMC Cancer*. 18(1), 381.

Bergvall M, Skullman S, Kodeda K, Larsson PA. (2019) Better survival for patients with colon cancer operated on by specialised colorectal surgeons - A nationwide population-based study in Sweden 2007-2010. *Colorectal Dis.* doi: 10.1111/codi.14760.

Berrino F, De Angelis R, Sant M, Rosso S, Bielska-Lasota M, Coebergh JW, Santaquilani M (2007). Survival for eight major cancers and all cancers combined for European adults diagnosed in 1995-99: results of the EUROCARE-4 study. *Lancet Oncol*. 8(9), 773-83.

Betti R, Martino P, Vergani R, Gualandri L, Crosti C. (2009). Nodular melanomas: analysis of the casistic and relationship with thick melanomas and diagnostic delay. *J Dermatol.* 35(10), 643–650.

Bhardwaj A, Bahl C, Sharma S, Singh N, Behera D. (2018). Interactive potential of genetic polymorphism in Xenobiotic metabolising and DNA repair genes for predicting lung cancer predisposition and overall survival in North Indians. *Mutat Res Genet Toxicol Environ Mutagen*. 826, 15-24.

Bieri U, Moch H, Dehler S, Korol D, Rohrmann S. (2015). Changes in autopsy rates among cancer patients and their impact on cancer statistics from a public health point of view: a longitudinal study from 1980 to 2010 with data from Cancer registry Zurich. *Virchows Arch*. 466(6), 637-43.

Bilimoria KY, Balch CM, Bentrem DJ, Talamonti MS, Ko CY, Lange JR, Winchester DP, Wayne JD. (2008). Complete lymph node dissection for sentinel node-positive melanoma: assessment of practice patterns in the United States. *Ann Surg Oncol.* 15(6), 1566–1576.

Bilimoria KY, Raval MV, Bentrem DJ, Wayne JD, Balch CM, Ko CY. (2009). National assessment of melanoma care using formally developed quality indicators. *J Clin Oncol*. 27, 5445–5551.

Blazer DG 3rd, Sondak VK, Sabel MS. (2007). Surgical therapy of cutaneous melanoma. *Semin Oncol*. 34, 270–280.

Bonaventure A, Harewood R, Stiller CA, Gatta G, Clavel J, Stefan DC, Carreira H, Spika D, Marcos-Gragera R, Peris-Bonet R, Piñeros M, Sant M, Kuehni CE, Murphy MFG, Coleman MP, Allemani C; CONCORD Working Group. (2017). Worldwide comparison of survival from childhood leukaemia for 1995-2009, by subtype, age, and sex (CONCORD-2): a population-based study of individual data for 89828 children from 198 registries in 53 countries. *Lancet Haematol.* 4(5), e202-e217.

Borgia JA, Basu S, Faber LP, Kim AW, Coon JS, Kaiser-Walters KA, Fhied C, Thomas S, Rouhi O, Warren WH, Bonomi P, Liptay MJ. (2009). Establishment of a multi-analyte serum biomarker panel to identify lymph node metastases in non-small cell lung cancer. *J Thorac Oncol*. 4, 338-47.

Borgognoni L, Urso C, Vaggelli L, Brandani P, Gerlini G, Reali UM. (2004). Sentinel node biopsy procedures with an analysis of recurrence patterns and prognosis in melanoma patients: technical advantages using computer- assisted gamma probe with adjustable collimation. *Melanoma Res*. 14, 311–319.

Boyle P, Parkin DM. (1991). Statistical methods for registries. In: Jensen OM, Parkin DM, MacLennan R, Muir CS, Skeet, RG, editors. Cancer registration: principles and methods, Lyon, France: IARC Scientific Publication (No. 95), pp. 126–157.

Bray F, Colombet M, Mery L, Piñeros M, Znaor A, Zanetti R and Ferlay J, editors. (2017). Cancer Incidence in Five Continents, Vol. XI (electronic version). Lyon: International Agency for Research on Cancer. Available from: http://ci5.iarc.fr, accessed [12/7/2019]

Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. (2018). Global cancer statistics_2018: GLOBOCAN_estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 68(6):394-424.

Bray F, Kohler B, Ferlay J. (2014). Chapter 5: Data comparability and quality. In: D. Forman, F. Bray, D.H. Brewster, C. Gombe Mbalawa, B. Kohler, M. Piñeros, E. Steliarova-Foucher, R. Swaminathan and J. Ferlay, (eds.) Cancer Incidence in 5 Continents volume X. IARCc Scientific Publication n.164, Lyon, France, pp. 89-97.

Bray F, Parkin MD. (2009). Evaluation of data quality in the cancer registry: Principles and methods. Part I: Comparability, validity and timeliness. *Eur J Cancer*. 45, 747-755.

Bread WA. (1908). Bills of Mortality, reprinted from Transactions of the Medico-Legal Society,1907–1908] London Baillière, Tindal and Cox, 8, Henrietta Street Covent Gardned, 1908. Available at https://en.wikisource.org/wiki/Bills_of_Mortality [Accessed 30 November 2018].

Breiman L. (2001). Random Forests. Available at http://oz.berkeley.edu/~breiman/randomforest2001.pdf [Accessed 18 October 2018].

Breiman L. Friedman JH, Olshen RA, Stone CJ. (1984). Classification and regression trees. New York: Chapman & Hall (Wadsworth, Inc).

Brenner H, Schmidtmann I. (1996). Determinants of homonym and synonym rates of record linkage in disease registration. *Methods Inf Med*. 35, 19-24.

Brenner H. (1995). Limitations of the death certificate only index as a measure of incompleteness of cancer registration. Br J Cancer. 72, 506-10.

Bruno C, Tumino R, Fazzo L, Cascone G, Cernigliaro A, De Santis M, Giurdanella MC, Nicita C, Rollo PC, Scondotto S, Spata E, Zona A, Comba P. (2014). Incidence of pleural mesothelioma in a community exposed to fibres with fluoro-edenitic composition in Biancavilla (Sicily, Italy). *Ann Ist Super Sanità*. 50(2), 111-118.

Buckman R. (1996). Talking to patients about cancer. BMJ. 313, 699.

Bullard J, Coleman MP, Robinson D, Lutz JM, Bell J, Peto J. (2000). Completeness of cancer registration: a new method for routine use. *Br J Cancer*. 82(5), 1111-6.

Burton EC, McDonald KM, Goldman L, Page P. (2003). Changes in rates of autopsy-detected. *JAMA*. 289(21), 2849–2856.

Caldarella A, Amunni G, Angiolini C, Crocetti E, Di Costanzo F, Di Leo A, Giusti F, Pegna AL, Mantellini P, Luzzatto L, Paci E. (2012). Feasibility of evaluating quality cancer care using registry data and electronic health records: a population-based study. *Int J Qual Health Care.* 24(4), 411–418.

Campbell SM, Roland MO, Buetow SA. (2000). Defining quality of care. *Soc Sci Med.* 51(11), 1611–25.

Capocaccia R, De Angelis R. (1997). Estimating the completeness of prevalence based on cancer registry data. *Stat Med*. 16(4), 425-40.

Capocaccia R, Gatta G, Roazzi P, Carrani E, Santaquilani M, De Angelis R, Tavilla A; EUROCARE Working Group. (2003). The EUROCARE-3 database: methodology of data

collection, standardisation, quality control and statistical analysis. *Ann Oncol*. 14 (Suppl. 5), v14-27.

Carli P, De Giorgi V, Betti R, Vergani R, Catricala C, Mariani G, Simonacci M, Bettacchi A, Bottoni U, Lo Scocco G, Mulas P, Giannotti B. (2003). Relationship between cause of referral and diagnostic outcome in pigmented lesion clinics aimed to melanoma screening: a multicenter survey of the Italian Multidisciplinary group on Melanoma (GIPMe). *Melanoma Res*. 13(2), 207–211.

Carli P, De Giorgi V, Chiarugi A, Nardini P, Weinstock MA, Crocetti E, Stante M, Giannotti B. (2004). Addition of dermoscopy to conventional naked-eye examination in melanoma screening: a randomized study. *J Am Acad Dermatol*. 50(5), 683–689.

Carli P, De Giorgi V, Giannotti B. (2001). Dermoscopy and early diagnosis of melanoma. The light and the dark. *Arch Dermatol*. 137, 1641–1644.

Carli P, De Giorgi V, Nardini P, Mannone F, Palli D, Giannotti B. (2002). Melanoma detection rate and concordance between self-skin examination and clinical evaluation in patients attending a pigmented lesion clinic in Italy. *Br J Dermatol*. 146(2), 261–266.

Carli P, Palli D (2003). Re: melanocytic nevi, solar keratoses, and divergent pathways to cutaneous melanoma. *J Natl Cancer Inst*. 95, 1801.

Carsin AE, Drummond FJ, Black A, van Leeuwen PJ, Sharp L, Murray LJ, Connolly D, Egevad L, Boniol M, Autier P, Comber H, Gavin A. (2010). Impact of PSA testing and prostatic biopsy on cancer incidence and mortality: comparative study between the Republic of Ireland and Northern Ireland. *Cancer Causes Control*. 21(9), 1523-31.

Carstensen B. (2007). Age-period-cohort models for the Lexis diagram. *Stat Med*. 26, 3018-45.

Cartwright LA, Dumenci L,Siminoff LA, Matsuyama RK. (2014). Cancer Patients' Understanding of Prognostic Information. *J Cancer Educ.* 29(2), 311–317.

Case LD, Kimmick G, Paskett ED, Lohman K, Tucker R. (2002). Interpreting measures of treatment effect in cancer clinical trials. *Oncologist*. 7(3), 181-187.

Castanon A, Landy R, Pesola F, Windridge P, Sasieni P. (2018). Prediction of cervical cancer incidence in England, UK, up to 2040, under four scenarios: a modelling study. *Lancet Public Health*. 3, e34-e43.

Caulfield S, Menezes G, Marignol L, Poole C. (2018). Nomograms are key decision-making tools in prostate cancer radiation therapy. *Urol Oncol*. 36(6), 283-292.

Chakera AH, Hesse B, Burak Z, Ballinger JR, Britten A, Caraco` C, Cochran AJ, Cook MG, Drzewiecki KT, Essner R, Even-Sapir E, Eggermont AM, Stopar TG, Ingvar C, Mihm MC Jr, McCarthy SW, Mozzillo N, Nieweg OE, Scolyer RA, Starz H, Thompson JF, Trifirò G, Viale G, Vidal-Sicart S, Uren R, Waddington W, Chiti A, Spatz A, Testori A; European Association of Nuclear Medicine-European Organisation for Research. (2009). EANM-EORTC general recommendations for sentinel node diagnostics in melanoma. *Eur J Nucl Med Mol Imaging*. 36(10), 1713–1742.

Chang Y.J., Chen L.J., Chung K.P., & Lai M.S. (2012). Risk groups defined by recursive partitioning analysis of patients with colorectal adenocarcinoma treated with colorectal resection. *BMC Medical Research Methodology.* 12(2), 2.

Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, Larkin J, Dummer R, Garbe C, Testori A, Maio M, Hogg D, Lorigan P, Lebbe C, Jouary T, Schadendorf D, Ribas A, O'Day SJ, Sosman JA, Kirkwood JM, Eggermont AM, Dreno B, Nolop K, Li J, Nelson B, Hou J, Lee RJ, Flaherty KT, McArthur GA; BRIM-3 Study Group. (2011). Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N Engl J Me*d. 364(26), 2507–2516.

Chellini E, Crocetti E, Carli P, Martini A, Giovannetti L. (2007). The melanoma epidemic debate: some evidence for a real phenomenon from Tuscany, Italy. *Melanoma Res.* 17, 129–130.

Chun FK, Karakiewicz PI, Briganti A, Walz J, Kattan MW, Huland H, Graefen M. (2007). A critical appraisal of logistic regression-based nomograms, artificial neural networks, classification and regression-tree models, look-up tables and risk-group stratification models for prostate cancer. *BJU Int*. 99(4), 794-800.

Claeson M, Andersson EM, Wallin M, Wastensson G, Wennberg AM, Paoli J, Gonzalez H. (2012). Incidence of cutaneous melanoma in Western Sweden, 1970-2007. *Melanoma Res*. 22(5), 392-8.

Clayton D, Schifflers E. (1987a) Models for temporal variation in cancer rates. I: age-period and age-cohort models. *Stat Med*. 6, 449–467.

Clayton D, Schifflers E. (1987b). Models for temporal variation in cancer rates. II: age-period-cohort models. *Stat Med*. 6, 469–481.

Clegg LX, Feuer EJ, Midthune DN, Fay MP, Hankey BF. (2002). Impact of reporting delay and reporting error on cancer incidence rates and trends. *J Natl Cancer Inst*. 94(20), 1537-45.

Clemmenson, J. (1995). Statistical studies in the aetiology of malignant neoplasms. *Acta Pathol Microbiol Scand*. 1, Suppl. 17.

Clèries R, Ameijide A, Marcos-Gragera R, Pareja L, Carulla M, Vilardell ML, Esteban L, Buxó M, Espinàs JA, Puigdefàbregas A, Ribes J, Izquierdo A, Galceran J, Borrás JM. (2018). Predicting the cancer burden in Catalonia between 2015 and 2025: the challenge of cancermanagement in the elderly. *Clin Transl Oncol*. 20(5):647-657.

Coebergh JW, van den Hurk C, Louwman M, Comber H, Rosso S, Zanetti R, Sacchetto L, Storm H, van Veen EB, Siesling S, van den Eijnden-van Raaij J. (2015). EUROCOURSE recipe for cancer surveillance by visible population-based cancer RegisTrees in Europe: From roots to fruits. *Eur J Cancer*. 51, 1050-63.

Coit DG, Andtbacka R, Bichakjian CK, Dilawari RA, Dimaio D, Guild V, Halpern AC, Hodi FS, Kashani-Sabet M, Lange JR, Lind A, Martin L, Martini MC, Pruitt SK, Ross MI, Sener SF, Swetter SM, Tanabe KK, Thompson JA, Trisal V, Urist MM, Weber J, Wong MK; NCCN Melanoma Panel. (2009). Melanoma. *J Natl Compr Canc Netw*. 7(3), 250–275.

Coleman MP, Quaresma M, Berrino F, Lutz JM, De Angelis R, Capocaccia R, Baili P, Rachet B, Gatta G, T, Micheli A, Sant M, Weir HK, Elwood JM, Tsukuma H, Koifman S, E Silva GA, Francisci S, Santaquilani M, Verdecchia A, Storm HH, Young JL; CONCORD Working Group. (2008). Cancer survival in five continents: a worldwide population-based study (CONCORD). *Lancet Oncol*. 9, 730-56.

Colonna M, Boussari O, Cowppli-Bony A, Delafosse P, Romain G, Grosclaude P, Jooste V; French Network of Cancer Registries (FRANCIM). (2018). Time trends and short term projections of cancer prevalence in France. *Cancer Epidemiol*. 56:97-105.

Colonna M, Sauleau EA. (2013). How to interpret and choose a Bayesian spatial model and a Poisson regression model in the context of describing small area cancer risks variations. *Rev Epidemiol Sante Publique*. 61, 559–567.

Contiero P, Tagliabue, G, Gatta, G, Crocetti, E, Scaburri, A, Grell Geographic Analysis Working Group. (2018). Geographic variation in cancer incidence in Latin-language countries. Oral presentation at Acta XLIII Réunion du groupe pour l'épidémiologie et

l'enregistrement du cancer dans les pays de langue latine (GRELL) Grell, Ascension Meeting, Trento, Italy, May 16th – 18th, 2018.

Cook MG, Di Palma S. (2008). Pathology of sentinel lymph nodes for melanoma. *J Clin Pathol.* 61, 897–902.

Cook MG, Green MA, Anderson B, Eggermont AM, Ruiter DJ, Spatz A, Kissin MW, Powell BW; EORTC Melanoma Group. (2003). The development of optimal pathological assessment of sentinel lymph nodes for melanoma. *J Pathol*. 200(3),314–319.

Corazziari I, Quinn M, Capocaccia R. (2004). Standard cancer patient population for age standardising survival ratios. *Eur J Cancer*. 40(15), :2307-16.

Cormier JN, Xing Y, Ding M, Lee JE, Mansfield PF, Gershenwald JE, Ross MD, Du XL. (2005). Population-based assessment of surgical treatment trends for patients with melanoma in the era of sentinel lymph node biopsy. *J Clin Oncol*. 23(25), 6054–6062.

Costa FA, Ramos C, Murteira R, Almodovar T, Passos-Coelho JL, Carvalho MI, Costa L, Brito MJ, Ramos S, Ferreira M, Miranda AC. (2019). The cancer registry as an ally in monitoring treatment effectiveness. *Pulmonology*. 25(1):3-8.

Cox DR. (1972). Regression models and life tables. J R Stat Soc. 34, 187–220.

Crichton NJ, Hinde JP & Marchini J. (1997). Models for diagnosing chest pain: is CART helpful? *Statistics in Medicine*. 16, 717–727.

Criscione VD, Weinstock MA. (2010). Melanoma thickness trends in the United Stases, 1988–2006. *J Invest Dermatol*. 130(3), 793-7 [Epub 2009].

Crocetti E, Baldacchini F, Bucchi L, Giuliani O, Mancini S, Ravaioli A, Falcini F. (2017a). Applicazione di un indice di macro variabilità geografica all'incidenza per tumori: USA, 2014. XLI Convegno dell'Associazione Italiana di Epidemiologia, Mantova, 25-27 Ottobre.

Crocetti E, Battisti L, Betta A, Palma PD, Paci E, Piffer S, Pojer A, Polla E, Zappa M. (2007). The cytological screening turned out effective also for adenocarcinoma: a population-based case-control study in Trento, Italy. *Eur J Cancer Prev*. 16(6), 564-7.

Crocetti E, Bossard N, Uhry Z, Roche L, Rossi S, Capocaccia R, Faivre J; GRELL EUROCARE-5 Working Group. (2017b). Trends in net survival from 15 cancers in six European Latin countries: the SUDCAN population-based study material. *Eur J Cancer Prev.* 6, S3-S8.

Crocetti E, Buzzoni C. (2017c). Do big numbers assure high-quality of data? *Lancet Haematol.* 4, e309.

Crocetti E, Caldarella A, Chiarugi A, Nardini P, Zappa M. (2010). The thickness of melanomas has decreased in central Italy, but only for thin melanomas, while thick melanomas are as thick as in the past. *Melanoma Res*. 20(5), 422–426.

Crocetti E, Capocaccia R, Casella C, Guzzinati S, Ferretti S, Rosso S, Sacchettini C, Spitale A, Stracci F, Tumino R.; Network of the Italian Cancer Registries (AIRT). (2004). Population-based incidence and mortality cancer trends (1986–1997) from the network of Italian cancer registries. *Eur J Cancer Prev*. 13(4), 287–295.

Crocetti E, Carli P. (2003). Unexpected reduction of mortality rates from melanoma in males living in central Italy. *Eur J Cancer*. 39(6), 818-821.

Crocetti E, Ciatto S, Buzzoni C, Zappa M. (2010). Prostate cancer incidence rates have started to decrease in central Italy. *J Med Screen*. 17(1), 50-1

Crocetti E, Gori S, Falcini F. (2018). Metastatic breast cancers: Estimates for Italy. *Tumori.* 104(2):116-120

Crocetti E, Mangone L, Lo Scocco G, Carlo P. (2006). Prognostic variables and prognostic groups for malignant melanoma. The information from Cox and Classificiation and regression tree analysis: an Italian population-based study. *Melanoma Res*. 16(5), 429-433.

Crocetti E, Miccinesi G, Paci E, Zappa M. (2001). An application of the two-source capture-recapture method to estimate the completeness of the Tuscany Cancer registry, Italy. *Eur J Cancer Prev*. 10(5), 417-23.

Crocetti E, Ravaioli A, Amadori D, Mancini S, Vattiato R, Giuliani O, Baldacchini F, Falcini F. (2019). Female breast cancers (T1-2, N0, M0, HR+, HER2-) with an intermediate genetic-based recurrence risk: a real-world estimate in Italy. *Tumori*. May 15:300891619849291.

Crocetti E, Sacchettini C, Caldarella A, Paci E. (2005). [Automatic coding of pathologic cancer variables by the search of strings of text in the pathology reports. The experience of the Tuscany Cancer registry]. *Epidemiol Prev*. 29(1), 57-60.

Dal Maso L, Guzzinati S, Buzzoni C, Capocaccia R, Serraino D, Caldarella A, Dei Tos AP, Falcini F, Autelitano M, Masanotti G, Ferretti S, Tisano F, Tirelli U, Crocetti E, De Angelis R; AIRTUM Working group. (2014). Long-term survival, prevalence, and cure of cancer: a population-based estimation for 818 902 Italian patients and 26 cancer types. *Ann Oncol.* 25(11), 2251-60.

Dal Maso L, Lise M, Zambon P, Falcini F, Crocetti E, Serraino D, Cirilli C, Zanetti R, Vercelli M, Ferretti S, Stracci F, De Lisi V, Busco S, Tagliabue G, Budroni M, Tumino R, Giacomin A, Franceschi S; AIRTUM Working Group. (2011). Incidence of thyroid cancer in Italy, 1991-2005: time trends and age-period-cohort effects. *Ann Oncol*. 22(4), 957-63.

Daniels J, Caetano SJ, Huyer D, Stephen A, Fernandes J, Lytwyn A, Hoppe FM. (2017). Benford's Law for Quality Assurance of Manner of Death Counts in Small and Large Databases. *J Forensic Sci*. 62(5), 1326-1331.

de Vries E, Bray F, Coebergh JWW, Parkin DM (2003). Changing epidemiology of malignant cutaneous melanoma in Europe 1953–1997: rising trends in incidence and mortality but recent stabilizations in Western Europe and decreases in Scandinavia. *Int J Cancer*. 107(1), 119–126.

de Vries E, Bray F, Eggermont AM, Coebergh JW; European Network of Cancer Registries. (2004). Monitoring stage-specific trends in melanoma incidence across Europe reveals the need for more complete information on diagnostic characteristics. *Eur J Cancer Prev*. 13(5), 387-95.

de Vries E, Coebergh JM. (2004). Cutaneous malignant melanoma incidence in Europe. *Eur J Cancer*. 40(16), 2355–2366.

Dewdney SB, Lachance J. (2016). Electronic Records, Registries, and the Development of "Big Data": Crowd-Sourcing Quality toward Knowledge. *Front Oncol*. 6, 268.

Diaconis P, Freedman DP. (1979). On rounding percentages. *J Am Stat Assoc*. 74(366), 359-364.

Diessner J, Wischnewsky M, Stüber T, Stein R, Krockenberger M, Häusler S, Janni W, Kreienberg R, Blettner M, Schwentner L, Wöckel A, Bartmann C. (2016). Evaluation of clinical parameters influencing the development of bone metastasis in breast cancer. *BMC Cancer*. 12(16), 307.

Doll R, Muir C, Waterhouse J. (1970). Cancer incidence in 5 Continents, volume II. UICC.

Doll R, Payne, P, Waterhouse JAH, eds. (1966). Cancer Incidence in Five Continents, Vol. I Union Internationale Contre le Cancer, Geneva.

Doll R. (1972). Cancer in Five Continents. Proc Roy Soc Med. 65, 49-55.

Donnelly C, Cairnduff V, Chen JJ, Kearney T, Fitzpatrick D, Fox C, Gavin A. (2017). The completeness and timeliness of cancer registration and the implications for measuring cancer burden. *Cancer Epidemiol*. 49, 101-107.

Donnelly C, Quaife S, Forbes L, Boylan J, Tishelman C, Gavin A. (2017). Do perceived barriers to clinical presentation affect anticipated time to presenting with cancer symptoms: an ICBP study. *Eur J Public Health*. 27(5), 808-813.

Dronkers EAC, Hoesseini A, de Boer MF, Offerman MPJ. (2018). Communication of prognosis in head and neck cancer patients; a descriptive qualitative analysis. Oral Oncol. 84, 76-81.

Durtschi C, Hillison W, Pacini C. (2004). The effective use of Benford's law to assist in detecting fraud in accounting data. *J Forensic Accounting*. VV, 17–34.

Dyba T, Hakulinen T, Paivarinta L. (1997). A simple non-linear model in incidence prediction. *Stat Med*.16(20), 2297–2309.

Earnest A, Evans SM, Sampurno F, Millar J. (2019). Forecasting annual incidence and mortality rate for prostate cancer in Australia until 2022 using autoregressive integrated moving average (ARIMA) models. *BMJ Open*. 19;9(8):e031331.

Eddy DJ. (2003). Surgical treatment of melanoma. Br J Dermatol. 149, :2–12.

Ederer F, Axtell L, Cutler S. (1961). The relative survival rate: A statistical methodology. *National Cancer Institute Monograph*. 6, 101–121.

Ederer F, Heise H. (1959). Instructions to IBM 650 programmers in processing survival computations. methodological note no. 10, end results evaluation section. Tech. rep., National Cancer Institute, Bethesda, MD.

Egger ME, Callender GG, McMasters KM, Ross MI, Martin RC 2nd, Edwards MJ, Urist MM, Noyes RD, Sussman JJ, Reintgen DS, Stromberg AJ, Scoggins CR. (2013). Diversity of stage III melanoma in the era of sentinel lymph node biopsy. *Ann Surg Oncol*. 20(3), 956-63.

Eggermont AM. (2010). Advances in systemic treatment of melanoma. *Ann Oncol*. 21, vii339–vii344.

Einwachter-Thompson J, MacKie RM. (2008). An evidence base for reconsidering current follow-up guidelines for patients with cutaneous melanoma less than 0.5 mm thick at diagnosis. *Br J Dermatol*. 159(2), 337–341.

Elwood JM, Jopson J. (1997). Melanoma and sun exposure: an overview of published studies. *Int J Cancer*. 73(2), 198–203.

Engholm G, Ferlay J, Christensen N, Kejs AMT, Hertzum-Larsen R, Johannesen TB, Khan S, Leinonen MK, Ólafsdóttir E, Petersen T, Schmidt LKH, Trykker H, Storm HH. (2016).

NORDCAN: cancer incidence, mortality, pre- valence and survival in the nordic countries, version 7.3. Association of the Nordic Cancer Registries. Danish Cancer Society. Available at http://www.ancr.nu. [Accessed 19 December 2016].

Engqvist Boman L, Sandelin K, Wengström Y, Silén C. (2017). Patients' learning and understanding during their breast cancer trajectory. *Patient Educ Couns*. 100(5), 795-804.

Erdmann F, Lortet-Tieulent J, Schüz J, Zeeb H, Greinert R, Breitbart EW, Bray F. (2013). International trends in the incidence of malignant melanoma 1953-2008--are recent generations at higher or lower risk? *Int J Cancer*. 132(2), 385-400.

Erickson JL, Velasco JM, Hieken TJ. (2008). Compliance with melanoma treatment guidelines in a community teaching hospital: time trends and other variables. *Ann Surg Oncol*. 15(4), :1211–1217.

Essig S, Michel G, Dupont C, Kiss A, Bergstraesser E, Tinner EM, Kuehni CE; Swiss Pediatric Oncology Group (SPOG). (2019). Communicating "cure" to pediatric oncology patients: A mixed-methods study. *Pediatr Blood Cancer*. 66(6), e27661.

Estève J, Benhamou E, Raymon L (eds). (1994). Statistical Methods in Cancer Research Volume IV Descriptive Epidemiology. IARC Scientific Publications No.128 International Agency for Research on Cancer, Lyon.

EU. Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2016.119.01.0001.01.ENG, [Accessed 3/12/2017]).

Ferlay J, Bray F, Steliarova-Foucher E, Forman D (2014). Cancer incidence in five continents, CI5plus IARC CancerBase No 9. Lyon, France: International Agency for Research on Cancer.

Ferlay J, Heanue M, Boyle P, editors. (2007). Cancer incidence in five continents. Lyon: IARC Scientific Publications. Vol. IX, pp. 319.

Ferlay J. (2014). Classificantion and coding. In: D. Forman, F. Bray, D.H. Brewster, C. Gombe Mbalawa, B. Kohler, M. Piñeros, E. Steliarova-Foucher, R. Swaminathan and J. Ferlay (eds) Cancer Incidence in 5 Continents volume X. IARC Scientific Publication n.164, Lyon, France, pp. 68-78.

Fernández MI, Williams SB, Willis DL, Slack RS, Dickstein RJ, Parikh S, Chiong E, Siefker-Radtke AO, Guo CC, Czerniak BA, McConkey DJ, Shah JB, Pisters LL, Grossman HB, Dinney CP, Kamat AM. (2017). Clinical risk stratification in patients with surgically resectable micropapillary bladder cancer. *BJU Int.* 119(5), 684-691.

Ferretti S, Guzzinati S, Zambon P, Manneschi G, Crocetti E, Falcini F, Giorgetti S, Cirilli C, Pirani M, Mangone L, Di Felice E, Del Lisi V, Sgargi P, Buzzoni C, Russo A, Paci E. (2009). [Cancer incidence estimation by hospital discharge flow as compared with cancer registries data]. *Epidemiol Prev*. 33(4-5), 147-53.

Feuer EJ, Rabin BA, Zou Z, Wang Z, Xiong X, Ellis JL, Steiner JF, Cynkin L, Nekhlyudov L, Bayliss E, Hankey BF. (2014). The Surveillance, Epidemiology, and End Results Cancer Survival Calculator SEER*CSC: validation in a managed care setting. *J Natl Cancer Inst Monogr*. (49), 265-74.

Fisher RA. (1921). Probable Error of a Coefficient of Correlation Deduced from a Small Sample. *Metron*. 1, 3-32.

Forman D, Bray F, Brewster DH, Gombe Mbalawa C, Kohler B, Piñeros M, Steliarova-Foucher E, Swaminathan R and Ferlay J, eds. (2013). Cancer Incidence in Five Continents, Vol. X (Electronic Version). Lyon: International Agency for Research on Cancer.

Forman D, Brewster DH, Kohler B, Piñeros M. (2014). Chapter 1: Introduction. In Cancer Incidence in Five Continents. In: Edited by D. Forman, F. Bray, D.H. Brewster, C. Gombe Mbalawa, B. Kohler, M. Piñeros, E. Steliarova-Foucher, R. Swaminathan and J. Ferlay, editors. Vol. X, IARC Scientific Publication No. 164, Lyon, France.

Francisci S, Capocaccia R, Grande E, Santaquilani M, Simonetti A, Allemani C, Gatta G, Sant M, Zigon G, Bray F, Janssen-Heijnen M; EUROCARE Working Group. (2009). The cure of cancer: a European perspective. *Eur J Cancer*. 45(6), 1067-79.

Frangos JE, Duncan LM, Piris A, Nazarian RM, Mihm MC Jr, Hoang MP, Gleason B, Flotte TJ, Byers HR, Barnhill RL, Kimball AB. (2012). Increased diagnosis of thin superficial spreading melanomas: a 20- year study. *J Am Acad Dermatol*. 67, 387–394.

Freeman A, Marshall M, Treadwell J, Spiegelhalter D, Rafi I. (2017). Communicating risk in primary care: what the Academy of Medical Sciences' report means in practice. *Br J Gen Pract*. 67(661), 346-347.

Friedman JH, Meulman JJ. (2003). Multiplice additive regression trees with application in epidemiology. *Stat Med*. 22(9), 1365–1381.

Garbe C, Buttner P, Bertz J, Burg G, d'Hoedt B, Drepper H, Guggenmoos-Holzmann I, Lechner W, Lippold A, Orfanos CE. (1995). Primary cutaneous melanoma. Identification of prognostic groups and estimation of individual prognosis for 5093 patients. *Cancer*. 75(10), 2484–2491.

Garbe C, Eigentler TK. (2007). Diagnosis and treatment of cutaneous melanoma: state of the art 2006. *Melanoma Res*. 17(2), 117–127.

Garbe C, Ellwanger U, Tronnier M, Brocker EB, Orfanos CE. (2002). The new American Joint Committee on cancer staging system for cutaneous melanoma. A critical analysis based on data of the German central malignant melanoma registry. *Cancer*. 94(8), 2305–2307.

Garbe C, Leiter U. (2009). Melanoma epidemiology and trends. *Clin Dermatol*. 27, 3–9.

Garbe C, Peris K, Hauschild A, Saiag P, Middleton M, Spatz A, Grob JJ, Malvehy J, Newton-Bishop J, Stratigos A, Pehamberger H, Eggermont AM; European Dermatology Forum; European Association of Dermato-Oncology; European Organization of Research and Treatment of Cancer. (2012). Diagnosis and treatment of melanoma. European consensus-based interdisciplinary guideline – update 2012. *Eur J Cancer.* 48(15), 2375–2390.

Garbe C, Peris K, Hauschild A, Saiag P, Middleton M, Spatz A, Grob JJ, Malvehy J, Newton-Bishop J, Stratigos A, Pehamberger H, Eggermont A. (2010). Diagnosis and treatment of melanoma: European consensus-based interdisciplinary guideline. *Eur J Cancer*. 46(2), 270–283.

Gatta G, Capocaccia R, Sant M, Bell CM, Coebergh JW, Damhuis RA, Faivre J, Martinez-Garcia C, Pawlega J, Ponz de Leon M, Pottier D, Raverdy N, Williams EM, Berrino F. (2000). Understanding variations in survival for colorectal cancer in Europe: a EUROCARE high resolution study. *Gut*. 47(4), 533-8.

Gatta G, Ciccolallo L, Capocaccia R, Coleman MP, Hakulinen T, Møller H, Berrino F; EUROCARE Working Group. (2003). Differences in colorectal cancer survival between European and US populations: the importance of sub-site and morphology. *Eur J Cancer*. 39(15), 2214-22.

Gatta G, Sant M, Coebergh JW, Hakulinen T. (1996). Substantial variation in therapy for colorectal cancer across Europe: EUROCARE analysis of cancer registry data for 1987. *Eur J Cancer*. 32A, 831-5.

Gatta G, Trama A, Capocaccia R; RARECARENet Working Group. (2017). Epidemiology of rare cancers and inequalities in oncologic outcomes. Eur J Surg Oncol. 45(1), 3-11.

Geller AC, Elwood M, Swetter SM, Brooks DR, Aitken J, Youl PH, Demierre MF, Baade PD. (2009). Factors related to the presentation of thin and thick melanoma from a population-based cancer registry in Queensland Australia. *Cancer.* 115(6), 1318–1327.

Golbeck J. (2015). Benford's law applies to online social networks. (2015). *PLoS One*. 10(8), e0135169.

Goldsbury D, Weber M, Yap S, Banks E, O'Connell DL, Canfell K. (2017). Identifying incident colorectal and lung cancer cases in health service utilisation databases in Australia: a validation study. *BMC Med Inform Decis Mak*. 17(1), 23.

Gondos A, Holleczek B, Janssen-Heijnen M, Brewster DH, Bray F, Rosso S, Hakulinen T, Brenner H; EUNICE Survival Working Group. (2009). Model-based projections for deriving up-to-date cancer survival estimates: an international evaluation. *Int J Cancer.* 125(11):2666-72.

Gordon EJ, Daugherty CK. (2003). 'Hitting you over the head': Oncologists' disclosure of prognosis to advanced cancer patients. *Bioethics*. 17(2), 142-68.

Graunt J. (1662). Natural and Political Observations mentioned in a following Index and made upon the Bills of Mortality.

Greene FL, Page DL, Fleming ID, Fritz AG, Balch CM, Haller DG, Morrow M. (2002). AJCC cancer staging manual. 6th ed. Chicago, IL: Springer.

Grilli R, Chiesa V. (2018). Overuse in cancer care: do European studies provide information useful to support policies? *Health Res Policy Syst*. 16(1), 12.

Grönberg H, Adolfsson J, Aly M, Nordström T, Wiklund P, Brandberg Y, Thompson J, Wiklund F, Lindberg J, Clements M, Egevad L, Eklund M. (2015). Prostate cancer screening in men aged 50–69 years (STHLM3): a prospective population-based diagnostic study. *Lancet Oncol*. 16(6), 1667–1676.

Gumbel EJ, Keeney RD. (1950). The extremal quotient. Ann Math Stat. 21(4), 523-538.

Guo JC, Wu Y, Chen Y, Pan F, Wu ZY, Zhang JS, Wu JY, Xu XE, Zhao JM, Li EM, Zhao Y, Xu LY. (2018). Protein-coding genes combined with long noncoding RNA as a novel transcriptome molecular staging model to predict the survival of patients with esophageal squamous cell carcinoma. *Cancer Commun (Lond*). 38(1), 4.

Gupta A. (2012). Pesticide use in south and south-east Asia: environmental public health and legal concerns. *Am J Environ Sci*. 8(2), 152.

Guzzinati S, Virdone S, De Angelis R, Panato C, Buzzoni C, Capocaccia R, Francisci S, Gigli A, Zorzi M, Tagliabue G, Serraino D, Falcini F, Casella C, Russo AG, Stracci F, Caruso B, Michiara M, Caiazzo AL, Castaing M, Ferretti S, Mangone L, Rudisi G, Sensi F, Mazzoleni G, Pannozzo F, Tumino R, Fusco M, Ricci P, Gola G, Giacomin A, Tisano F, Candela G, Fanetti AC, Pala F, Sardo AS, Rugge M, Botta L, Maso LD. (2018). Characteristics of people living in Italy after a cancer diagnosis in 2010 and projections to 2020. *BMC Cancer*. 18(1):169.

Hadley LG, Rouma BS, Saad-Eldin Y. (2012). Challenge of pediatric oncology in Africa. *Semin Pediatr Surg*. 21(2), 136-41.

Haenszel W. (1975). The United States network of cancer registries. In: Grundmann, E. & Pedersen, E., eds., Cancer registry, Berlin, Springer.

Hahn EJ, Rayens MK, Wiggins AT, Gan W, Brown HM, Mullett TW. (2018). Lung cancer incidence and the strength of municipal smoke-free ordinances. *Cancer*. 124(2), 374-380.

Hakulinen T. (1982). Cancer survival corrected for heterogeneity in patient withdrawal. *Biometrics*. 38(4), 933–942.

Hansen BT, Campbell S, Nygård M. (2018). Long-term incidence trends of HPV-related cancers, and cases preventable by HPV vaccination: a registry-based study in Norway. *BMJ Open*. 8(2), e019005.

He C, Plattner R, Rangnekar V, Zhou B, Liu C, Stewart RL, Huang B, Wang C, Tucker TC. (2019). Potential protein markers for breast cancer recurrence: a retrospective cohort study. *Cancer Causes Control*. 30(1), 41-51.

Hein J, Zobrist R, Konrad C, Schuepfer G. (2012). Scientific fraud in 20 falsified anesthesia papers: detection using financial auditing methods. *Anaesthesist*. 61(6), 543–9.

Hernandez AF, Menendez P. (2016). Linking pesticide exposure with pediatric leukemia: potential underlying mechanisms. *Int J Mol Sci*. 17(4), 461.

Hershman DL, Wright JD. (2012). Comparative effectiveness research in oncology methodology: observational data. *J Clin Oncol*. 30(34), 4215-22.

Hèry C, Tryggvadóttir L, Sigurdsson T, Ólafsdóttir E, Sigurgeirsson B, Jonasson JG, Olafsson JH, Boniol M, Byrnes GB, Doré JF, Autier P. (2010). A melanoma epidemic in Iceland: Possible influence of sunbed use. *Am J Epidemiol*. 172(7), 762-7.

Hess K.R., Abbruzzese M.C., Lenzi R., Raber M.N. & Abbruzzese J.L. Classification and regression tree analysis of 1000 consecutive patients with unknown primary carcinoma. (1999). *Clinical Cancer Research*. 5(11), 3403–3410.

Hiatt RA, Tai CG, Blayney DW, Deapen D, Hogarth M, Kizer KW, Lipscomb J, Malin J, Phillips SK, Santa J, Schrag D. (2015). Leveraging State Cancer Registries to Measure and Improve the Quality of Cancer Care: A Potential Strategy for California and Beyond. *J Natl Cancer Inst*. 107(5), pii: djv047.

Hill PH. (1995). A statistical derivation of the significant-digit law. *Stat Sci*. 10(4), 354–63.

Hodi FS, O'Day SJ, Mcdermott DF, Weber RW, Sosman JA, Haanen JB, Gonzalez R, Robert C, Schadendorf D, Hassel JC, Akerley W, van den Eertwegh AJ, Lutzky J, Lorigan P, Vaubel JM, Linette GP, Hogg D, Ottensmeier CH, Lebbé C, Peschel C, Quirt I, Clark JI, Wolchok JD, Weber JS, Tian J, Yellin MJ, Nichol GM, Hoos A, Urba WJ. (2010). Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med*. 363(8), 711–723.

Hoffman FL. (1930). Discussion to the paper of F. C. Wood. Am. J. Public Health. 20, 19-20.

Holford TR. (2006). Approaches to fitting age-period-cohort models with unequal intervals. *Stat Med*. 25(6), 977-93.

Huang L, Midthune D, Krapcho M, Zou Z, Horner MJ, Feuer EJ. (2013). Adjusting for reporting delay in cancer incidence when combining different sets of cancer registries. *Biom J*. 55(5), 755-70.

Hüllemann S, Schüpfer G, Mauch J. (2017). Application of Benford's law: a valuable tool for detecting scientific papers with fabricated data? A case study using proven falsified articles against a comparison group. *Anaesthesist*. 66(10), 795-802.

Hutchison, L.L, Roffers, SD& Fritz AG. (eds). (1997). 6. Cancer registry Management, Principles and Practice, Kendal Hunt, Iowa.

ITT Melanoma WG. (2007). Clinical recommendations for skin melanoma. Florence (Italy), Tuscany Tumor Institute – ITT: Scientific Press.

Jain P, Kantarjian HM, Ghorab A, Sasaki K, Jabbour EJ, Nogueras Gonzalez G, Kanagal-Shamanna R, Issa GC, Garcia-Manero G, Kc D, Dellasala S, Pierce S, Konopleva M, Wierda

WG, Verstovsek S, Daver NG, Kadia TM, Borthakur G, O'Brien S, Estrov Z, Ravandi F, Cortes JE. (2017). Prognostic factors and survival outcomes in patients with chronic myeloid leukemia in blast phase in the tyrosine kinase inhibitor era: Cohort study of 477 patients. *Cance*r. 123(22), 4391-4402.

Janssen-Heijnen ML, van Steenbergen LN, Voogd AC, Tjan-Heijnen VC, Nijhuis PH, Poortmans PM, Coebergh JW, van Spronsen DJ. (2014). Small but significant excess mortality compared with the general population for long-term survivors of breast cancer in the Netherlands. *Ann Oncol*. 25: 64-68.

Jensen OM, Storm HH. (1991). Reporting of results. In: Cancer registration: principles and methods. O.M. Jensen, D.M. Parkin, R. MacLennan, C.S. Muir and R.G. Skeet, (eds). WHO, IARC, Lyon, France, pp. 108-125.

Jensen OM, Whelan S. (1991). Planning a cancer registry. In: O.M. Jensen, D.M. Parkin, R. MacLennan, C.S. Muir and R.G. Skeet, eds. Cancer Registration: Principles and Methods. IARC Scientific Publications No. 95 Lyon, pp. 22-28.

Jochems A, Schouwenburg MG, Leeneman B, Franken MG, van den Eertwegh AJ, Haanen JB, Gelderblom H, Uyl-de Groot CA, Aarts MJ, van den Berkmortel FW, Blokx WA, Cardous-Ubbink MC, Groenewegen G, de Groot JW, Hospers GA, Kapiteijn E, Koornstra RH, Kruit WH, Louwman MW, Piersma D, van Rijn RS, Ten Tije AJ, Vreugdenhil G, Wouters MW, van der Hoeven JJ. (2017). Dutch Melanoma Treatment Registry: Quality assurance in the care of patients with metastatic melanoma in the Netherlands. *Eur J Cancer*. 72, 156-165.

Johnson CH, Peace S, Adamo P, Fritz A, Percy-Laurry A, Edwards BK. (2007). The 2007 Multiple Primary and Histology Coding Rules. National Cancer Institute, Surveillance, Epidemiology and End Results Program. Bethesda, MD.

Jouhet V, Defossez G, Burgun A, le Beux P, Levillain P, Ingrand P, Claveau V. (2012). Automated Classification of Freetext Pathology Reports for Registration of Incident Cases of Cancer. *Methods Inf Med*. 51(3), 242-51.

Judge G, Schechter JG. (2009). Detecting problems in survey data using Benford's law. *J Human Resour*. 44, 1–24.

Jung KW, Won YJ, Kong HJ, Lee ES. (2019). Prediction of Cancer Incidence and Mortality in Korea, 2019. Cancer Res Treat. 51(2):431-437.

Kaplan EL, Meier P. (1958). Non-parametric estimation from incomplete observations. *J Am Stat Assoc*. 53, 457–481.

Karthik D, Stelzer G, Gershanov S, Baranes D, Salmon-Divon M. (2016). Elucidating tissue specific genes using the Benford distribution. BMC Genomics. 17, 595.

Katayama K, Narimatsu H. (2016). Prediction of Female Breast Cancer Incidence among the Aging Society in Kanagawa, Japan*. PLoS One.* 11(8), e0159913.

Kearney TM, Donnelly C, Kelly JM, O'Callaghan EP, Fox CR, Gavin AT. (2015). Validation of the completeness and accuracy of the Northern Ireland Cancer registry. *Cancer Epidemiol.* 39(3), 401-4.

Keiding, G. (1973). Annotation zur Krebsepidemiologie. Hamburg. Ärzteblatt. 8, 27.

Kelly M, O'Brien KM, Lucey M, Clough-Gorr K, Hannigan A. (2018). Indicators for early assessment of palliative care in lung cancer patients: a population study using linked health data. *BMC Palliat Care*. 17(1), 37.

Kemp A, Preen DB, Saunders C, Holman CD, Bulsara M, Rogers K, Roughead EE. (2013). Ascertaining invasive breast cancer cases; the validity of administrative and self-reported data sources in Australia. *BMC Med Res Methodol*. 13, 17.

Kienle S. (2015). What Benford can tell us about cover pools. An empirical analysis. *Int Businness Econ Res J*. 14, 829–34.

Kim HJ, Fay MP, Feuer EJ, Midthune DN. (2000). Permutation tests for joinpoint regression with applications to cancer rates. *Stat Med*. 19(3), 335-51 (correction: 2001;20:655).

Kim HJ, Son BH, Ahn SH. (2014). Daily collection of self-reporting sleep disturbance data via a smartphone app in breast cancer patients receiving chemotherapy: a feasibility study. *J Med Internet Res*. 16, e135.

Kim JY, Hildebrandt MA, Pu X, Ye Y, Correa AM, Vaporciyan AA, Wu X, Roth JA. (2012). Variations in the vascular endothelial growth factor pathway predict pulmonary complications*. Ann Thorac Surg*. 94(4), 1079-84.

Kobayashi H, Okuma T, Oka H, Okajima K, Ishibashi Y, Zhang L, Hirai T, Ohki T, Tsuda Y, Ikegami M, Sawada R, Shinoda Y, Akiyama T, Kawano H, Goto T, Tanaka S. (2019). Body composition as a predictor of toxicity after treatment with eribulin for advanced soft tissue sarcoma. *Int J Clin Oncol*. 24(4), 437-444.

Køster B, Meyer MK, Andersson TM, Engholm G, Dalum P. (2018). Sunbed use 2007-2015 and skin cancer projections of campaign results 2007-2040 in the Danish population: repeated cross-sectional surveys. *BMJ Open*. 8(8), v e022094.

Køster B, Thorgaard C, Philip A, Clemmensen IH. (2010). Prevalence of sunburn and sun-related behaviour in the Danish population: a cross-sectional study. *Scand J Public Health*. 38(5),548-52.

Kraemer H. (1992). Evaluating medical tests: objective and quantitative guidelines. Newbury Park, CA: Sage Publications.

Lamborn K.R., Chang S.M. &, Prados M.D. (2004). Prognostic factors for survival of patients with glioblastoma: recursive partitioning analysis. *Neurological Oncology*. 6(3), 227–235.

Langendijk JA, Slotman BJ, van der Waal I, Doornaert P, Berkof J, Leemans CR. (2005). Risk-group definition by recursive partitioning analysis of patients with squamous cell head and neck carcinoma treated with surgery and postoperative radiotherapy. *Cancer*. 104(7), 1408-1417.

Lasithiotakis K, Kruger-Krasagakis S, Ioannidou D, Ioannidou D, Pediaditis I, Tosca A. (2004). Epidemiological differences for cutaneous melanoma in a relatively dark-skinned Caucasian population with chronic sun exposure. *Eur J Cancer*. 40(16), 2502–2507.

Lavie A, Desouches C, Casanova D, Bardot J, Grob JJ, Legrée´ R, Magalon G. (2007). Surgical management of cutaneous malignant melanoma. Review. *Ann Chir Plast Esthet.* 52(1), 1–13.

Lee J, de Carvalho M. (2019). Technological improvements or climate change? Bayesian modeling of time-varying conformance to Benford's Law. *PLoS One*. 14(4), e0213300.

Leemis LM, Schmeiser BW, Evans DL. (2000). Survival distributions satisfying Benford's law. *Am Stat*. 54(4), 236–41.

Lehnich AT, Rusner C, Chodick G, Katz R, Sella T, Stang A. (2019). Actual frequency of imaging during follow-up of testicular cancer in Israel-a comparison with the guidelines. *Eur Radiol*. 29(7), 3918-3926.

Leinonen MK, Miettinen J, Heikkinen S, Pitkäniemi J, Malila N. (2017). Quality measures of the population-based Finnish Cancer registry indicate sound data quality for solid malignant tumours. *Eur J Cancer*. 77, 31-39.

Leiter U, Buettner PG, Eigentler TK, Garbe C. Prognostic factors of thin cutaneous melanoma: an analysis of the central malignant melanoma registry of German dermatological society. (2004). *J Clin Oncol*. 22(18), 3651–3653.

Lemon SC, Roy J, Clark MA, Friedmann PD, Rakowski W. (2003). Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. *Ann Behav Med*. 26(3), 172–181.

Lens M. (2008). Current clinical overview of cutaneous melanoma. *Br J Nurs*. 17(5), 300–305.

Lens MB, Dawes M, Goodacre T, Newton-Bishop JA. (2002). Elective lymph node dissection in patients with melanoma: systematic review and meta-analysis of randomized controlled trials. *Arch Surg.* 137(4), 458–461.

Lens MB, Dawes M. (2004). Global perspectives of contemporary epidemiological trends of cutaneous malignant melanoma. Br J Dermatol. 150(2), 179–185.

Lens MB, Nathlan P, Bataille V. (2007). Excision margins for primary cutaneous melanoma: updated pooled analysis of randomized controlled trials. *Arch Surg*. 142(9), 885–891.

Lewis DR, Chen HS, Cockburn MG, Wu XC, Stroup AM, Midthune DN, Zou Z, Krapcho MF, Miller DG, Feuer EJ. (2018). Early estimates of cancer incidence for 2015: Expanding to include estimates for white and black races. *Cancer*. 124(10):2192-2204.

Lewis DR, Chen HS, Midthune DN, Cronin KA, Krapcho MF, Feuer EJ. (2015). Cancer. Early estimates of SEER cancer incidence for 2012: Approaches, opportunities, and cautions for obtaining preliminary estimates of cancer incidence. *Cancer*. 121(12), 2053-62.

Lewis RJ. (2000). An introduction to classification and regression tree (CART) analysis. Presented at the 2000 Annual Meeting of the Society for Academic Emergency Medicine in San Francisco California (www.saem.org/ download/lewis1.pdf).

Lilford LJ, Brown CA, Nicholl J. (2007). Use of process measures to monitor the quality of clinical practice. *BMJ.* 335(7621), 648–650.

Linos E, Swetter SM, Cockburn MG, Colditz GA, Clarke CA. (2009). Increasing burden of melanoma in the United States. J Invest Dermatol. 129(7), 1666–1674.

Lipsker D, Engel F, Cribel B, Velten M, Hedelin G. (2007). Trends in melanoma epidemiology suggest three types of melanoma. *Br J Dermatol*. 157(2), 338–343.

Lipsker DM, Hedelin G, Heid E, Grosshans EM, Cribier BJ. (1999). Striking increase of thin melanomas contrasts with stable incidence of thick melanoma. *Arch Dermatol*. 135(12), 1451–1456.

Livi L, Paiar F, Meldolesi E, Bianchi S, Cardona G, Cataliotti L, Crocetti E, Distante V, Simontacchi G, Scoccianti S, Talamonti C, Pallotta S, Biti GP. (2005). The management of elderly patients with T1-T2 breast cancer treated with or without radiotherapy. *Eur J Surg Oncol.* 31(5):473-8.

Lu L, Katsaros D, Canuto EM, Biglia N, Risch HA, Yu H. (2016). LIN-28B/let-7a/IGF-II axis molecular subtypes are associated with epithelial ovarian cancerprognosis. *Gynecol Oncol*. 141(1), 121-7.

Luo G, Zhang Y, Guo P, Wang L, Huang Y, Li K. (2017). Global patterns and trends in stomach cancer incidence: Age, period and birth cohort analysis. *Int J Cancer*. 141(7), 1333-1344.

Mackenzie LJ, Carey ML, Suzuki E, Sanson-Fisher RW, Asada H, Ogura M, D'Este C, Yoshimura M, Toi M. (2018). Agreement between patients' and radiation oncologists' cancer diagnosis and prognosis perceptions: A cross sectional study in Japan. *PLoS One*. 13(6), e0198437.

MacLennan R. (1991). Items of patient information which may be collected by registries. In: O.M. Jensen, D.M. Parkin, R. MacLennan, C.S. Muir and R.G. Skeet, eds. Cancer Registration: Principles and Methods. IARC Scientific Publications No. 95 Lyon, pp. 29-42.

Magrath I, _Steliarova-Foucher E, Epelman S, Ribeiro RC, Harif M, Li CK, Kebudi R, Macfarlane SD, Howard SC. (2013). Paeditric cancer in low-income and middle-income countries. *Lancet Oncol*. 14(3), e104–16.

Mainz J, Hansen AM, Palshof T, Bartels P. (2009). National quality measurement using clinical indicators: the Danish National Indicator Project. J Surg Oncol. 99(8), 500–504.

Mainz J. (2003). Defining and classifying clinical indicators for improvement. *Int J Qual Health Care*. 15(6), 523–530.

Makkouk A, Sundaram V, Chester C, Chang S, Colevas AD, Sunwoo JB, Maecker H, Desai M, Kohrt HE. (2017). Characterizing CD137 upregulation on NK cells in patients receiving monoclonal antibody therapy. *Ann Oncol*. 28(2), 415-420.

Malin JL, Schneider EC, Epstein AM, Adams J, Emanuel EJ, Kahn KL. (2006). Results of the National Initiative for Cancer Care Quality: how can we improve the quality of cancer care in the United States? *J Clin Oncol.* 24(4), :626–34.

Mantel N. (1966). Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemother Rep*. 1966; 50(3), :163–170.

Mariotto AB, Zou Z, Zhang F, Howlader N, Kurian AW, Etzioni R. (2018). Can We Use Survival Data from Cancer Registries to Learn about Disease Recurrence? The Case of Breast Cancer. *Cancer Epidemiol Biomarkers Prev*. 27(11), 1332-1341.

Marsden JR, Newton-Bishop JA, Burrows L, Cook M, Corrie PG, Cox NH, Gore ME, Lorigan P, Mackie R, Nathan P, Peach H, Powell B, Walker C; British Association of Dermatologists (BAD) Clinical Standards Unit. (2010). Revised UK guidelines for management of cutaneous melanoma 2010. Br J Dermatol. 163(9), 238–256.

Marshall RJ. (2001). The use of classification and regression trees in clinical epidemiology. *J Clin Epidemiol*. 54(6), 603–609.

Martos C, Crocetti E (Coordinator), Visser O, Rous B and the Cancer Data Quality Checks Working Group. (2014). One common procedure for data quality checks for European cancer registries. Luxembourg: Publications Office of the European Union.

Martos C., Giusti F, Crocetti E, Randi G, Neamtiu L, Dyba T, Dimitrova N, N. Carvalho R, Bettio M.. (2018). Understanding geographical variations in the incidence of bladder cancer in Europe. oral presentation at Grell, Ascension Meeting, Trento, Italy, May 16th–18th.

Mazzucco W, Cusimano R, Zarcone M, Mazzola S, Vitale F. (2017). Funnel plots and choropleth maps in cancer risk communication: a comparison of tools for disseminating population-based incidence data to stakeholders. *BMJ Open*. 7(3), e011502.

McArdle J. (2012). Exploratory data mining using CART in the behavioral sciences. In: Cooper H, Camic P, Long D, Panter A, Rindskopf D, Sher K, editors. APA Handbook of Research Methods in Psychology. Washington, DC: American Psychological Association; pp. 405–421.

McMasters KM, Noyes RD, Reintgen DS, Goydos JS, Beitsch PD, Davidson BS, Sussman JJ, Gershenwald JE, Ross MI; Sunbelt Melanoma Trial. (2004). Lessons learned from the Sunbelt Melanoma Trial. *J Surg Oncol*. 86(4), 212–223.

Menck, H. & Smart, C. (editors). (1994). Central Cancer Registries, Design, Management, and Use. Hargood Academic Publishers, Switzerland.

Miccinesi G, Crocetti E, Morino P, Fallai M, Piazza M, Cavallini V, Paci E. (2003). Palliative home care reduces time spent in hospital wards: a population-based study in the Tuscany Region, Italy. *Cancer Causes Control*. 14(10), 971-7.

Min YH, Lee JW, Shin YW, Jo MW, Sohn G, Lee JH, Lee G, Jung KH, Sung J, Ko BS, Yu JH, Kim HJ, Son BH, Ahn SH. (2014). Daily collection of self-reporting sleep disturbance data via a smartphone app in breast cancer patients receiving chemotherapy: a feasibility study. *J Med Internet Res*. 16(5), e135.

Minicozzi P, Innos K, Sánchez MJ, Trama A, Walsh PM, Marcos-Gragera R, Dimitrova N, Botta L, Visser O, Rossi S, Tavilla A, Sant M; EUROCARE-5 Working Group. (2017). Quality analysis of population-based information on cancer stage at diagnosis across Europe, with presentation of stage-specific cancer survival estimates: A EUROCARE-5 study. *Eur J Cancer*. 84, 335-353.

Minicozzi P, Walsh PM, Sánchez MJ, Trama A, Innos K, Marcos-Gragera R, Dimitrova N, Botta L, Johannesen TB, Rossi S, Sant M; EUROCARE-5 Working Group. (2018). Is low survival for cancer in Eastern Europe due principally to late stage at diagnosis? *Eur J Cancer*. 93, 127-137.

Mocellin S, Pasquali S, Nitti D. (2011). The impact of surgery on survival of patients with cutaneous melanoma: revisiting the role of primary tumor excision margins. *Ann Surg*. 2353(2), 238–243.

Moller B, Fekjaer H, Hakulinen T, Sigvaldason H, Storm HH, Talback M, Haldorsen T. (2003). Prediction of cancer incidence in the Nordic countries: empirical comparison of different approaches. *Stat Med*. 22(17), 2751–2766.

Morrison AS. (1985). Screening in Chronic Disease. Monographs in epidemiology and biostatistics. Oxford University Press. NY, USA.

Morton DL, Thompson JF, Cochran AJ, Mozzillo N, Elashoff R, Essner R, Nieweg OE, Roses DF, Hoekstra HJ, Karakousis CP, Reintgen DS, Coventry BJ, Glass EC, Wang HJ; MSLT Group. (2006). Sentinel-node biopsy or nodal observation in melanoma. *N Engl J Med*. 355(13), 1307–1317.

Muir CS, Demaret E, Boyle P. (1985). The cancer registry in cancer control: an overview. In: Parkin, D. M., Wagner, G. & Muir, C. S., eds, The Role of the Registry in Cancer Control

(IARC Scientific Publications No. 66), Lyon, International Agency for Research on Cancer, pp. 13-26.

Mullins BT, Basak R, Broughman JR, Chen RC. (2019). Patient-reported sexual quality of life after different types of radical prostatectomy and radiotherapy: Analysis of a population-based prospective cohort. *Cancer*. doi: 10.1002/cncr.32288.

Murray CS, Stockton DL, Doherty VR. (2005). Thick melanoma: the challenge persists. *Br J Dermatol*. 152(1), 104–109.

Naldi L, Lorenzo Imberti G, Parazzini F, Parazzini F, Gallus S, La Vecchia C. (2000). Pigmentary traits, modalities of sun reaction, history of sunburns, and melanocytic nevi as risk factors for cutaneous malignant melanoma in the Italian population: results of a collaborative case–control study. *Cancer*. 88(12) :2703–2710.

Newcomb S. (1881). Note on the frequency of use of the different digits in natural numbers. *Am J Math*. 4:39–40.

Newcombe HB. (1995), When "privacy" threatens public health. *Can J Public Health.* 86(3), 188-92.

Nguyen SM, Deppen S, Nguyen GH, Pham DX, Bui TD, Tran TV. (2019). Projecting Cancer Incidence for 2025 in the 2 Largest Populated Cities in Vietnam. *Cancer Control*. 26(1):1073274819865274.

Nicolai N, Biasoni D, Catanzaro MA, Colecchia M, Trama A; RARECAREnet Working Group. (2018). Testicular germ-cell tumours and penile squamous cell carcinoma: Appropriate management makes the difference. *Eur J Surg Oncol*. 45(1), 60-66.

Ogilvie GS, Naus M, Money DM, Dobson SR, Miller D, Krajden M, van Niekerk DJ, Coldman AJ. (2015). Reduction in cervical intraepithelial neoplasia in young women in British Columbia after introduction of the HPV vaccine: An ecological analysis. *Int J Cancer.* 137(8), 1931-7.

Oviedo M, Muñoz MP, Carmona G, Borrás E, Batalla J, Soldevila N, Domínguez A. (2012). The impact of immigration and vaccination in reducing the incidence of hepatitis B in Catalonia (Spain). *BMC Public Health*. 12(1), 614.

Paci E, Crocetti E, Benvenuti A, Buzzoni C, Caldarella A, Giovannetti L. (2007). Tuscany cance registry. In: , Edwards B, Shin HR, Storm H, Ferlay J, Heanue M, Boyle P, editors. Cancer Incidence in Five Continents, Vol. IX IARC Scientific Publications No. 160. Lyon: IARC.

Paci E, Crocetti E, Miccinesi G, Benvenuti A, Intrieri T, Sacchettini C, Giovannetti L. Tuscany Cancer Registry. In: Curado MP, Edwards B, Shin HR, Storm H, Italian medical oncologic association (AIOM). Guidelines: melanoma, 2010. Available at http://www.aiom.it/area + pubblica/area + medica/ prodotti + scientifici/linee + guida/Archivio/1,333-333 [Accessed 15 April 2013].

Pacini F, Castagna MG, Brilli L, Pentheroudakis G; ESMO Guidelines Working Group. (2012). Thyroid cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 23(Suppl 7), vii110-9.

Pak K, Uno H, Kim DH, Tian L, Kane RC, Takeuchi M, Fu H, Claggett B, Wei LJ. (2017). Interpretability of Cancer Clinical Trial Results Using Restricted Mean Survival Time as an Alternative to the Hazard Ratio. *JAMA Oncol*. 1(3), 1692-1696.

Palli D, Rosselli del Turco M, Buiatti E, Ciatto S, Crocetti E, Paci E. (1989). Time interval since last test in a breast cancer screening programme: a case-control study in Italy. *J Epidemiol Community Health*. 43(3), 241-8.

Parkin DM, Chen VW, Ferlay J, Galceran J, Storm HH, Whelan SL. (1994). Comparability and quality control in cancer registration. IARC Technical Report No. 19. IARC, Lyon.

Parkin DM, Hakulinen T. (1991). Analysis of survival. In: O.M. Jensen, D.M. Parkin, R. MacLennan, C.S. Muir and R.G. Skeet. Cancer Registration: Principles and Methods. IARC Scientific Publications No. 95 Lyon, 1991, pp. 159-176.

Parkin DM. (2006). The evolution of the population-based cancer registry. *Nat Rev*. 6(8), 603-12.

Parkin MD, Bray F. (2009). Evaluation of data quality in the cancer registry: Principles and methods. Part II: Completeness. *Eur J Cancer*. 45(5), 756-764.

Pilleron S, Sarfati D, Janssen-Heijnen M, Vignat J, Ferlay J, Bray F, Soerjomataram I. (2019). Global cancer incidence in older adults, 2012 and 2035: A population-based study. *Int J Cancer*. 144(1):49-58.

Pinilla J, López-Valcárcel BG, González-Martel C, Peiro S. (2018). Pinocchio testing in the forensic analysis of waiting lists: using public waiting list data from Finland and Spain for testing Newcomb-Benford's Law. *BMJ Open*. 8, e022079.

Pohar Perme M, Estève J, Rachet B. (2016). Analysing population-based cancer survival - settling the controversies. *BMC Cancer*. 16(1), 933.

Pohar Perme M, Stare J, Esteve J. (2012). On estimation in relative survival. *Biometrics.* 68(1), 113–120.

Pollach G, Brunkhorst F, Mipando M, Namboya F, Mndolo S, Luiz T. (2016). The "first digit law" - A hypothesis on its possible impact on medicine and development aid. *Med Hypotheses*. 97, 102-106.

Powell J. (1991). Data sources and reporting. In: O.M. Jensen, D.M. Parkin, R. MacLennan, C.S. Muir and R.G. Skeet. Cancer Registration: Principles and Methods. IARC Scientific Publications No. 95 Lyon, pp. 29-42.

Protopopoff N, Van Bortel W, Speybroeck N, Van Geertruyden JP, Baza D, D'Alessandro U, & Coosemans M. (2009). Ranking malaria risk factors to guide malaria control efforts in African highlands. *PLoS ONE*. 4, e8022.

Puliti D, Miccinesi G, Collina N, De Lisi V, Federico M, Ferretti S, Finarelli AC, Foca F, Mangone L, Naldoni C, Petrella M, Ponti A, Segnan N, Sigona A, Zarcone M, Zorzi M, Zappa M, Paci E; IMPACT Working Group. (2008). Effectiveness of service screening: a case-control study to assess breast cancer mortality reduction. *Br J Cancer*. 99(3), 423-7.

Rashid I. (2013). Gli effetti distorcenti della standardizzazione diretta nell'era dello standard Europa 2013: il caso dei tumori polmonari XXII. Oral presentation at Riunione scientifica AIRTUM, Venice, Italy, 11-13 April.

Rasmussen LA, Jensen H, Virgilsen LF, Jensen JB, Vedsted P. (2018). A validated algorithm to identify recurrence of bladder cancer: a register-based study in Denmark. *Clin Epidemiol*. 10, 1755-1763.

Rauch B, Göttsche M, Brähler G, Engel S. (2011). Fact and fiction in the EU- governmental economic data. *German Econ Rev*. 2, :243–55.

Reintgen DS, Jakub JW, Pendas S, Swor G, Giuliano R, Shivers S. (2004). The staging of malignant melanoma and the Florida Melanoma Trial. *Ann Surg Oncol.* 11(3 Suppl), 186S–191S.

Richardson A, Flechter L, Sneyd M, Cox B, Reeder AI. (2008). The incidence and thickness of cutaneous malignant melanoma in New Zealand 1994–2004. *N Z Med J*. 121(1279), 18–26.

Rigoni-Stem D. (1842). Fatti statistici relativi alle malattie cancerose che servirono di base alle poche cose dette dal dott. *Gior Servire Progr Path Therap*. 2, 507–517.

Riker AI, Kirksey L, Thompson L, Morris A, Cruse CW. (2006). Current surgical management of melanoma. *Expert Rev Anticancer Ther.* 6(11),1569–1583.

Robles SC, Marrett LD, Clarke EA, Rish HA. (1988). An application of capture–recapture methods to the estimation of completeness of cancer registration. *J Clin Epidemiol*. 41(5), 495–501.

Rokach L. & Maimon O. (2007). Data Mining with Decision Trees: Theory and Applications. World Scientific, River Edge, NJ.

Romain G, Boussari O, Bossard N, Remontet L, Bouvier AM, Mounier M, Iwaz J, Colonna M, Jooste V; French Network of Cancer Registries (FRANCIM). (2019). Time-to-cure and cure proportion in solid cancers in France. A population based study. *Cancer Epidemiol*. 60, 93-101.

Roscher I, Falk RS, Vos L, Clausen OPF, Helsing P, Gjersvik P, Robsahm TE. (2018). Validating 4 Staging Systems for Cutaneous Squamous Cell Carcinoma Using Population-Based Data: A Nested Case-Control Study. *JAMA Dermatol*. 154(4), 428-434.

Rosenberg PS, Anderson WF. (2011). Age-period-cohort models in cancer surveillance research: ready for prime time? *Cancer Epidemiol Biomarkers Prev*. 20(7), 1263-8.

Rossi S, Baili P, Capocaccia R, Caldora M, Carrani E, Minicozzi P, Pierannunzio D, Santaquilani M, Trama A, Allemani C, Belot A, Buzzoni C, Lorez M, De Angelis R; EUROCARE-5 Working Group. (2015). The EUROCARE-5 study on cancer survival in Europe 1999-2007: Database, quality checks and statistical analysis methods. *Eur J Cancer.* 51(15), 2104-2119.

Ryzhov A, Bray F, Ferlay J, Fedorenko Z, Goulak L, Gorokh Y, Soumkina O, Znaor A. (2018). Evaluation of data quality at the National Cancer registry of Ukraine. *Cancer Epidemiol*. 53, :156-165.

Sacchetto L, Zanetti R, Comber H, Bouchardy C, Brewster DH, Broganelli P, Chirlaque MD, Coza D, Galceran J, Gavin A, Hackl M, Katalinic A, Larønningen S, Louwman MWJ, Morgan E, Robsahm TE, Sanchez MJ, Tryggvadóttir L, Tumino R, Van Eycken E, Vernon S, Zadnik V, Rosso S. (2018). Trends in incidence of thick, thin and *in situ* melanoma in Europe. *Eur J Cancer*. 92, 108-118.

Sacerdote C, Bordon R, Pitarella S, Mano MP, Baldi I, Casella D, Di Cuonzo D, Frigerio A, Milanesio L, Merletti F, Pagano E, Ricceri F, Rosso S, Segnan N, Tomatis M, Ciccone G, Vineis P, Ponti A. (2013). Compliance with clinical practice guidelines for breast cancer

treatment: a population-based study of quality-of-care indicators in Italy. *BMC Health Serv Res*. 13, 28.

Sankila R, Black R, Coebergh JW, Démaret E, Forman D, Gatta G, Parkin MD. (2003). Evaluation of clinical care by cancer registries. IARC Scientific publication n. 37, IARC, Lyon, France.

Sant M, Allemani C, Capocaccia R, Hakulinen T, Aareleid T, Coebergh JW, Coleman MP, Grosclaude P, Martinez C, Bell J, Youngson J, Berrino F; EUROCARE Working Group. (2003) Stage at diagnosis is a key explanation of differences in breast cancer survival across Europe. *Int J Cancer*. 106(3), 416-22.

Sant M, Capocaccia R, Verdecchia A, Gatta G, Micheli A, Mariotto A, Hakulinen T, Berrino F. (1995). Comparisons of colon-cancer survival among European countries: The Eurocare Study. *Int J Cancer*. 63(1), 43-8.

Sargen MR, Kanetsky PA, Newton-Bishop J, Hayward NK, Mann GJ, Gruis NA, Tucker MA, Goldstein AM, Bianchi-Scarra G, Puig S, Elder DE. (2015). Histologic features of melanoma associated with CDKN2A genotype. J Am Acad Dermatol. 72(3), 496-507.e7.

Saumell Y, Sanchez L, González S, Ortiz R, Medina E, Galán Y, Lage A. (2017). Overall Survival of Patients with Locally Advanced or Metastatic Esophageal Squamous Cell Carcinoma Treated with Nimotuzumab in the Real World. *Adv Ther*. 34(12), 2638-2647.

Schain WS. (1980). Patients' rights in decision making: the case for personalism versus paternalism in health care. *Cancer*. 46(4 Suppl), 1035-1041.

Schenker N, Gentleman JF. (2001). On judging the significance of differences by examining the overlap between confidence intervals. Am Stat. 55(3), 182–186.

Schiffner R, Schiffner-Rohe J, Landthaler M, Stolz W. (2003). Long-term dermoscopic follow-up of melanocyitc naevi: clinical outcome and patient compliance. *Br J Dermatol*. 149(1):79–86.

Schneider EC, Epstein AM, Malin JL, Kahn KL, Emanuel EJ. (2004). Developing a system to assess the quality of cancer care: ASCO's national initiative on cancer care quality. *J Clin Oncol*. 22(15), 2985–91.

Schouten LJ, Straatman H, Kiemeney LA, Gimbrere CH, Verbeek AL. (1994). The capture–recapture method for estimation of cancer registry completeness: a useful tool? *Int J Epidemiol*. 23(6), 1111–6.

Schröder FH. (1995). Screening, early detection, and treatment of prostate cancer: a European view. *Urology*. 46(3 Suppl A), 62-70.

Segi, M. (1960). Cancer Mortality for Selected Sites in 24 Countries. (1950–57), Sendai, Tohoku University School of Public Health.

Shanmugaratnam K. (1991). Introduction. In O.M. Jensen, D.M. Parkin, R. MacLennan, C.S. Muir and R.G. Skeet. Cancer Registration: Principles and Methods. IARC Scientific Publications No. 95 Lyon, pp. 1-2.

Shariat SF, Capitanio U, Jeldres C, Karakiewicz PI. (2009). Can nomograms be superior to other prediction tools? *BJU Int*. 103(4), 492-5.

Shin JA, El-Jawahri A, Parkes A, Schleicher SM, Knight HP, Temel JS. (2016). Quality of Life, Mood, and Prognostic Understanding in Patients with Metastatic Breast Cancer. J Palliat Med. 19(8), 863-9.

Shryock HS, Siegel JS, Stockwell EG. (1976). The methods and materials of demography. New York: Academic Press.

Siesling S, Louwman WJ, Kwast A, van den Hurk C, O'Callaghan M, Rosso S, Zanetti R, Storm HH. (1988). Completeness of cancer registration in Denmark1943-1966 and efficacy of linkage procedures. *Int J Epidemiol*. 17(1), 44-9.

Siesling S, Louwman WJ, Kwast A, van den Hurk C, O'Callaghan M, Rosso S, Zanetti R, Storm H, Comber H, Steliarova-Foucher E, Coebergh JW. (2015). Uses of cancer registries for public health and clinical research in Europe: Results of the European Network of Cancer Registries survey among 161 population-based cancer registries during 2010-2012. *Eur J Cancer*. 51(9), 1039-49.

Sladden MJ, Balch C, Barzilai DA, Berg D, Freiman A, Handiside T, Hollis S, Lens MB, Thompson JF. (2009). Surgical excision margins for primary cutaneous melanoma. *Cochrane Database Syst Rev*. 4, CD004835.

Slijepĉević S, Blašković B. (2014). Statistical detection of fraud in the reporting of Croatian public companies. *Financ Theory Prac*. 38, :81–96.

Smittenaar CR, Petersen KA, Stewart K, Moitt N. (2016). Cancer incidence and mortality projections in the UK until 2035. *Br J Cancer*. 115(9), 1147–1155.

Sondak VK, Zager JS, Messina JL. (2007). Sentinel lymph node biopsy as the standard of care for cutaneous melanoma: pro. *Clin Adv Hematol Oncol*. 5(6), 483–486.

Souliotis K, Peppou L-E, Tzavara C, Agapidaki E, Varvaras D, Buonomo O, Debiais D, Hasurdjiev S, Sarkozy F. (2018). Cancer patients' organisation participation in heath policy decision-making: a snapshot/cluster analysis of the EU-28 countries. *BMJ Open*. 8(8), e018896.

Spiegelhalter DJ, Riesch H. (2011). Don't know, can't know: embracing deeper uncertainties when analysing risks. *Philos Trans A Math Phys Eng Sci*. 369(1956), 4730-50.

Spillane AJ, Saw RP, Tucker M, Byth K, Thompson JF. (2008). Defining lower limb lymphedema after inguinal or ilio-inguinal dissection in patients with melanoma using classification and regression tree analysis. *Ann Surg*. 248(2), 286-93.

Spitale A, Ortelli L, Mazzucchelli L, Bordoni A. (2017). Quality indicators of colorectal cancer care in southern Switzerland: results from a population-based study. *Swiss Med Wkly.* 147, w14530.

Spurgeon SE, Hsieh YC, Rivadinera A, Beer TM, Mori M, Garzotto M. (2006). Classification and regression tree analysis for the prediction of aggressive prostate cancer on biopsy. *J Urol*. 175 (3 Pt 1), 918-22.

Start RD, Firth JA, Macgillivray F, Cross SS, Road BH. (1995). Have declining clinical necropsy rates reduced the contribution of necropsy to medical research? *J Clin Pathol*. 48(5), 402–404

Staudt M, Lasithiotakis K, Leiter U, Meier F, Eigentler T, Bamberg M, Tatagiba M, Brossart P, Garbe C. (2010). Determinants of survival in patients with brain metastases from cutaneous melanoma. *Br J Cancer*. 102(8), 1213-8.

Steliarova-Foucher E, Colombet M, Ries LAG, Moreno F, Dolya A, Bray F, Hesseling P, Shin HY, Stiller CA; IICC-3 contributors. (2017). International incidence of childhood cancer, 2001-10: a population-based registry study. *Lancet Oncol*. 18(6), 2017.

Steliarova-Foucher E, Stiller C, Lacour B, Kaatsch P. (2005). International Classification of Childhood Cancer, Third Edition. *Cancer*. 103(7), 1457- 1467.

Stordeur S, Vrijens F, Devriese S, Beirens K, Van Eycken E, Vlayen J. (2012). Developing and measuring a set of process and outcome indicators for breast cancer. *Breast*. 21(3), 253-60.

Storm H, Comber H, Steliarova-Foucher E, Coebergh JW. (2015). Uses of cancer registries for public health and clinical research in Europe: Results of the European Network of

Cancer Registries survey among 161 population-based cancer registries during 2010-2012. *Eur J Cancer*. 51(9), 1039-49.

Storm HH. (1988). Completeness of cancer registration in Denmark 1943-1966 and efficacy of linkage procedures. *Int J Epidemiol*. 17(1), 44-9.

Su X., Azuero A., Cho J., Kvale E., Meneses K.M. & McNees M.P. (2011). An introduction to tree-structured modeling with application to quality of life data. *Nursing Research*. 60(4), 247–255.

Sun Z, Chen T, Thorgeirsson SS, Zhan Q, Chen J, Park JH, Lu P, Hsia CC, Wang N, Xu L, Lu L, Huang F, Zhu Y, Lu J, Ni Z, Zhang Q, Wu Y, Liu G, Wu Z, Qu C, Gail MH. (2013). Dramatic reduction of liver cancer incidence in young adults: 28-year follow-up of etiological interventions in an endemic area of China. *Carcinogenesis*. 34(8), 1800-5.

Suppa M, Gandini S, Njimi H, Bulliard JL, Correia O, Duarte AF, Peris K, Stratigos AJ, Nagore E, Longo MI, Bylaite-Bucinskiene M, Karls R, Helppikangas H; Euromelanoma Working Group, Del Marmol V. (2019). Prevalence and determinants of sunbed use in thirty European countries: data from the Euromelanoma skin cancer prevention campaign. *J Eur Acad Dermatol Venereol*. 33 (Suppl. 2), 13-27.

Tejera-Vaquerizo A, Mendiola-Fernandez M, Fernandez-Orland A, Herrera-Ceballos E. (2008). Thick melanoma: the problem continues. *J Eur Acad Dermatol Venerol*. 22(5), 575–579.

Teppo L, Thomas DB, editors. (2002). Cancer incidence in five continents. Vol VIII. Lyon: IARC Scientific Publications no. 155, pp. 362–363.

Testori A, De Salvo GL, Montesco MC, Trifiro` G, Mocellin S, Landi G, Macripò G, Carcoforo P, Ricotti G, Giudice G, Picciotto F, Donner D, Di Filippo F, Soteldo J, Casara D, Schiavon M, Vecchiato A, Pasquali S, Baldini F, Mazzarol G, Rossi CR; Italian Melanoma Intergroup. (2009). Clinical considerations on sentinel node biopsy in melanoma from an Italian multicentric study on 1313 patents (SOLISM_IMI). *Ann Surg Oncol.* 16(7), 2018–2027.

Thompson JF, Soong SJ, Balch CM, Gershenwald JE, Ding S, Coit DG, Flaherty KT, Gimotty PA, Johnson T, Johnson MM, Leong SP, Ross MI, Byrd DR, Cascinelli N, Cochran AJ, Eggermont AM, McMasters KM, Mihm MC Jr, Morton DL, Sondak VK. (2011). Prognostic significance of mitotic rate in localized primary cutaneous melanoma: an analysis of

patients in the multi-institutional American Joint Committee on Cancer melanoma staging database. *J Clin Oncol*. 29(16), 2199–220.

Tödter KH. (2009). Bedford's laws as an indicator of fraud in economics. *German Econ Rev*. 10(8), 339–51.

Trinquart L, Jacot J, Conner SC, Porcher R. (2016). Comparison of Treatment Effects Measured by the Hazard Ratio and by the Ratio of Restricted Mean Survival Times in Oncology Randomized Controlled Trials. *J Clin Oncol*. 34(15), 1813-9.

Tu Y-K, Smith GD, Gilthorpe MS. (2011). A New Approach to Age-Period-Cohort Analysis Using Partial Least Squares Regression: The Trend in Blood Pressure in the Glasgow Alumni Cohort. *Plos ONE*. 6(4), e19401.

Ursin G, Sen S, Mottu JM, Nygård M. (2017). Protecting Privacy in Large Datasets-First We Assess the Risk; Then We Fuzzy the Data. *Cancer Epidemiol Biomarkers Prev*. 2017 Jul 28. doi:26(8), 1219-1224.

Vaccarella S, Lortet-Tieulent J, Plummer M, Franceschi S, Bray F. (2013). Worldwide trends in cervical cancer incidence: impact of screening against changes in disease risk factors. *Eur J Cancer*. 49(15), 3262-73.

Venkatasubramaniam A, Wolfson J, Mitchell N, Barnes T, Jaka M, and French S. (2017). Decision trees in epidemiological research. Emerging Themes in Epidemiology. 14, 11.

Verdasca J, Costa FA, Ramos C, Murteira R, Miranda A. (2019). The South Region Cancer Registry: an evaluation of its exhaustiveness in a cohort of lung cancer patients. *Thorac Cancer*. 10(2):330-334.

Vicentini M, Zorzi M, Bovo E, Mancuso P, Zappa M, Manneschi G, Mangone L, Giorgi Rossi P; Colorectal Cancer Screening IMPATTO study working group. (2019). Impact of screening programme using the faecal immunochemical test on stage of colorectal cancer: Results from the IMPATTO study. *Int J Cancer*. 145(1), 110-121.

Vlayen J, Vrijens F, Devriese S, Beirens K, Van Eycken E, Stordeur S. (2012). Quality indicators for testicular cancer: a population-based study. *Eur J Cancer*. 48(8), 1133-40.

Vrijens F, De Gendt C, Verleye L, Robays J, Schillemans V, Camberlin C, Stordeur S, Dubois C, Van Eycken E, Wauters I, Van Meerbeeck JP. (2018). Quality of care and variability in lung cancer management across Belgian hospitals: a population-based study using routinely available data. *Int J Qual Health Care*. 30(4), 306-312.

Wagner G. (1985). Cancer registration: Historical aspects. In: Parkin, D. M., Wagner, G. & Muir, C. S., eds, The Role of the registry in Cancer Control (IARC Scientific Publications No. 66), Lyon, International Agency for Research on Cancer, pp. 3-12.

Wagner G. (1991). History of cancer registration. In: O.M. Jensen, D.M. Parkin, R. MacLennan, C.S. Muir and R.G. Skeet (eds) IARC Scientific Publication No. 95, chapter 2.

Wagner JD, Gordon MS, Chuang TY, Coleman JJ. (2000). Current therapy of cutaneous melanoma. *Plast Reconstr Surg*. 105(5), 1774–1799.

Waldmann A, Nolte S, Weinstock MA, Breitbart EW, Eisemann N, Geller AC, Greinert R, Volkmer B, Katalinic A. (2012). Skin cancer screening participation and impact on melanoma incidence in Germany – an observational study on incidence trends in regions with and without population-based screening. *Br J Cancer*. 106(5), 970–974.

Wallace WA. (2002). Assessing the quality of data used for benchmarking and decision making. *J Gov Financ Manag*. 51(3), 16–22.

Walsh KM, Chokkalingam AP, Hsu LI, Metayer C, de Smith AJ, Jacobs DI, Dahl GV, Loh ML, Smirnov IV, Bartley K, Ma X, Wiencke JK, Barcellos LF, Wiemels JL, Buffler PA. (2013). Associations between genome-wide Native American ancestry, known risk alleles and B-cell ALL risk in Hispanic children. *Leukemia*. 27(12), 2416–19.

Ward MH, Colt JS, Metayer C, Gunier RB, Lubin J, Crouse V, Nishioka MG, Reynolds P, Buffler PA. (2009). Residential exposure to polychlorinated biphenyls and organochlorine pesticides and risk of childhood leukemia. *Environ Health Perspect*. 117(6), 1007–13.

Ward ZJ, Yeh JM, Bhakta N, Frazier AL, Atun R. (2019). Estimating the total incidence of global childhood cancer: a simulation-based analysis.*Lancet Oncol*. 20(4):483-493.

Warycha MA, Christos PJ, Mazumdar M, Darvishian F, Shapiro RL, Berman RS, Pavlick AC, Kopf AW, Polsky D, Osman I. (2008). Changes in the presentation of nodular and superficial spreading melanomas over 35 years. Cancer. 113(12), 3341–3348.

Welch HG, Black WC. (2010). Overdiagnoisis in cancer. *J Natl Cancer Inst*. 102(9); 115-16.

Welch HG, Woloshin S, Schwartz LM. (2005). Skin biopsy rates and incidence of melanoma: population based ecological study. *BMJ*. 2005; 331/7515), :481.

Whelan SL. (2010). The International Association of Cancer Registries: a history. Asian Pacific Journal of Cancer Prevention vol II, 2010: IACR Supplement.

Whiteman DC, Watt P, Purdie DM, Hughes MC, Hayward NK, Green AC. (2003). Melanocytic nevi, solar keratoses, and divergent pathways to cutaneous melanoma. *J Natl Cancer Inst*. 95(11), 806–812.

Wilkinson E. (2009). Questions remain over validity of EUROCARE data. *Lancet*. 374(9694), 964-5.

Winn AN, Shah GL, Cohen JT, Lin PJ, Parsons SK. (2015). The real world effectiveness of hematopoietic transplant among elderly individuals with multiple myeloma. *J Natl Cancer Inst*. 107(8), pii: djv139.

Wood FC. (1930). Need for cancer morbidity statistics. *Am J Publ Health*. 20(1), 11-20.

Woods LM, Coleman MP, Lawrence G, Rashbass J, Berrino F, Rachet B. (2011). Evidence against the proposition that "UK cancer survival statistics are misleading": simulation study with National Cancer Registry data. *BMJ*. 342, 3399.

Wright F, Spithoff K, Easson A, Murray C, Toye J, McCready D, Petrella T; Melanoma Disease Site Group of Cancer Care Ontario's Program in Evidence-based Care. (2011). Primary excision margins and sentinel lymph node biopsy in clinically node-negative melanoma of the trunk or extremities. *Clin Oncol*. 23(9), 572–578.

Wu JW, Azoulay L, Huang A, Paterson M, Wu F, Secrest MH, Filion KB. (2018). Identification of incident pancreatic cancer in Ontario administrative health data: A validation study. *Pharmacoepidemiol Drug Saf*. Oct 4. doi: 10.1002/pds.4641.

Xie L, Semenciw R, Mery L. (2015). Cancer incidence in Canada: trends and projections (1983-2032). *Health Promot Chronic Dis Prev Can*. 35(Suppl. 1), 2-186.

Xu R, Adak S. (2002). Survival analysis with time-varying regression effects using a tree-based approach. *Biometrics*. 58(2), 305–315.

Xu Y, Kong S, Cheung WY, Bouchard-Fortier A, Dort JC, Quan H, Buie EM, McKinnon G, Quan ML. (2019). Development and validation of case-finding algorithms for recurrence of breast cancer using routinely collected administrative data. *BMC Cancer.* 19(1), 210.

Yabroff KR, Mariotto AB, Feuer E, Brown ML. (2008). Projections of the costs associated with colorectal cancer care in the United States, 2000–2020. *Health Econ*. 17(8), 947-59.

Young JK. (1991). The hospital-based cancer registry. In: Cancer registration: principles and methods. O.M. Jensen, D.M. Parkin, R. MacLennan, C.S. Muir and R.G. Skeet, (eds.). WHO, IARC, Lyon, France, pp. 177-184.

Yu XQ, Luo Q, Hughes S, Wade S, Caruana M, Canfell K, O'Connell DL. (2019). Statistical projection methods for lung cancer incidence and mortality: a systematic review. *BMJ Open.* 9(8): e028497.

Yuan H, Ai QY, Kwong DL, Fong DY, King AD, Vardhanabhuti V, Lee VH, Khong PL. (2017). Cervical nodal volume for prognostication and risk stratification of patients with nasopharyngeal carcinoma, and implications on the TNM-staging system. *Sci Rep.* 7(1), 10387.

Zanetti R, Sacchetto L, Coebergh JW, Rosso S. (2018). To accelerate cancer prevention in Europe: Challenges for cancer registries. *Eur J Cancer*. 104, 151-159.

Zanetti R, Schmidtmann I, Sacchetto L, Binder-Foucard F, Bordoni A, Coza D, Ferretti S, Galceran J, Gavin A, Larranaga N, Robinson D, Tryggvadottir L, Van Eycken E, Zadnik V, Coebergh JWW, Rosso S. (2015). Completeness and timeliness: Cancer registries could/should improve their performance. *Eur J Cancer*. 51(9), 1091–1098.

Ziegler LE, Craigs CL, West RM, Carder P, Hurlow A, Millares-Martin P, Hall G, Bennett MI. (2018). Is palliative care support associated with better quality end-of-life care indicators for patients with advanced cancer? A retrospective cohort study. *BMJ Open.* 8(1), e018284.

Zijlstra M, Timmermans M, Fransen H, van der Aa M, Reyners A, Raijmakers N, van de Poll-Franse L. (2019). Treatment patterns and associated factors in patients with advanced epithelial ovarian cancer: a population-based study. *Int J Gynecol Cancer*. 29(6), 1032-1037.

Znaor A, Laversanne M, Bray F. (2017). Less overdiagnosis of kidney cancer? An age-period-cohort analysis of incidence trends in 16 populations worldwide. *Int J Cancer*. 2017;141(5):925-932.

Appendix 1: List of related abstracts presented in Congresses

Oral presentation, 21st May 2009

Diverging figures for thin and thick skin melanoma.

GRELL 2009. XXXIV Group for cancer registration and epidemiology in Latin language countries. 20-22 May 2009, Lugano, Switzerland.

Emanuele Crocetti, Carlotta Buzzoni, Antonella Corbinelli, Francesco Giusti, Teresa Intrieri, Gianfranco Manneschi, Liuba Nemcova, Claudio Sacchettini, Alessandras Chiarugi, Paolo Nardini, Eugenio Paci

Tuscany Cancer Registry, Clinical and descriptive epidemiology unit and *Unit for early diagnosis of melanoma, ISPO, Florence, Italy

Objective
Breslow's thickness is the most relevant prognostic factor for skin melanoma. Most of thick melanomas are of nodular type. Early diagnosis aims at detecting melanoma when it is thin and it can be successfully removed with very good prognosis. This study focuses on the evaluation of melanoma incidence trends according to thickness and morphology.
Methods
We analysed data from the Tuscany Cancer Registry for four subsequent 5-year periods from 1985 to 2004. The following morphology were analysed (ICD-0 3 morphology code): superficial spreading (8743), on lentigo (8742), nodular (8721), other types, not otherwise specified (8720). Thickness was categorized as follows: thin (0.01-0.99 mm), intermediate (1.00-1.99 mm), thick (2+mm).
Results
The incidence of skin melanoma has markedly increased over time. The number of newly diagnosed skin melanoma has particularly increased (+57%) from 1985-89 (n. 442) to 2000-04 (n. 1020). The proportion of thin melanoma has grown from 18.6% to 45.5%, while the proportion of thick ones has decreased from 28.1 to 20.5. Also, intermediate melanoma has decreased from 17.9 to 12.9, by the same amount of the cases who were lacking information on Breslow's thickness (from 35.5 to 21.1%). The main cause of the growth of incidence lies in the growth of superficial spreading melanoma. In the meantime, the median thickness went from 1.7 mm to 0.8 mm. The thinning of the median thickness for SSM went from 1.20 mm to 0.68 mm. Among thick melanoma (>=2 mm) from 1985-1989 to 2000-2004 there were no statistically significant changes, neither of the mean thickness (from 4.2 mm to 4.9 mm) nor of that median (from 3.3 to 3.7 mm).In the most recent period (2000-2004), nodular melanomas represent 6.2% of all the cases but 24.5% of the thick ones. During the analysed period there was no statistically significant change in the median thickness for nodular melanomas.
Conclusions
The overall decrease of the mean thickness at diagnosis for skin melanomas observed in central Italy from mid-80ths to mid 2000ths was due to the diagnosis of a number of increasingly thinning melanomas. On the other hand, thick melanomas – most of them of nodular type - did not show any evident sign of modification regardless of the

time of diagnosis. Nodular melanomas did show stable incidence over time and their median thickness did not change. This type of deadly melanoma, more than others, is the target for future prevention actions.

Oral presentation, 5th May 2016

A possible contribution to the quality evaluation of cancer registry data may come from the Benford's mathematical law

Grell - Group for cancer registration and epidemiology in Latin language countries. 4 - 6 May 2016, Albi, France
https://docs.wixstatic.com/ugd/6510a0_e79879a68cd44b4093c6ca71c48dacac.pdf

Emanuele Crocetti, Giorgia Randi, Tadek Dyba, Raquel Carvalho, Francesco Giusti, Carmen Martos, Roisin Rooney, Manola Bettio

European Commission, DG Joint Research Centre (JRC), Institute for Health and Consumer Protection, Public Health Policy Support Unit

Objectives
According to the Benford's law, the distribution of the occurrences of the first digit (FSD) in many large collections of numbers is not uniform. We evaluated, in a sample of GRELL cancer registries, whether incidence rates followed Bedford's law, as a possible contribution in their quality check process.
Materials and methods
Data from six European population-based cancer registries from GRELL countries (Belgium, France, Spain, Italy, Portugal and Switzerland) were retrieved from the Cancer Incidence in 5 Continents-X website. Crude incidence rates were computed for the main cancer groups. The distribution of FSD of incidence rates was computed for the six registries altogether, and separately by registry and sex. The adherence of the observed FSD frequency distributions to the Benford's law was evaluated both graphically and through several statistical tests.
Results
All the FSD distributions of incidence rates were positively skewed, as expected for those following Benford's law. The ratios between the frequency of the first and the second digit ranged around Benford's expected value (1.7), as that between the first and the ninth (6.6). The coefficient of correlation was overall high but not always as expected, ranging from 0.82 to 0.99. Also the distance measures from the observed and the expected FSD distribution, although generally small, showed slightly higher values for some registries.
Discussion and conclusions
The quality evaluation of cancer data is complex and implies assessing several different dimensions. The FSD distribution of incidence rates, in this GRELL setting, appeared to adhere to the Benford's law. The analyzed data had already been checked and approved for publication in Cancer Incidence in 5 Continents-X. However, some minor deviations from the Benford's distribution may still suggest possible revision or integration of the applied criteria for data validation.

Poster, 7-10 June 2016

Can We Apply Benford's Law To Check Quality Of Cancer Incidence Data?

IARC 50 years Global cancer. Occurrence, causes and avenues to prevention. 7-10 June 2016, Lyon, France.
http://www.iarc-conference2016.com/index.php?onglet=21&idUser=&emailUser=&acces=&recherche=crocetti

Emanuele Crocetti, Giorgia Randi, Tadeusz Dyba, Raquel Carvalho, Francesco Giusti, Carmen Martos, Roisin Rooney, Lydia Voti, Manola Bettio

European Commission, DG Joint Research Centre, Institute for Health and Consumer Protection, Public Health Policy Support Unit, Ispra (VA), Italy

Purpose
Benford's law states that the distribution of the occurrence of the first significant digit (FSD) of a number, in many large collections of numbers, is not uniform. The aim of this study was to evaluate whether population-based cancer incidence rates follow Benford's law and if this can be used in their data quality checking process.
Methods
Detailed databases from six population-based cancer registries (from Africa, North and South America, Asia, Europa and Oceania) were retrieved from the Cancer Incidence in 5 Continents-X website. These datasets consisted of 244 combinations of topography and morphological groups, 18 age groups and two sexes. The distribution of FSD was evaluated for the whole dataset, plus for some subgroups as cancer registries, cancer types and sexes. Several statistics, including Pearson's coefficient of correlation, distance measures and specific tests, were applied to check for consistency between calculated FSD frequency distribution and the theoretical Benford's one.
Results
The distribution of FSD, calcutated for each combination, consistently showed mean values greater than the medians and were positively skewed. For the whole dataset (22,180 observations), and for single cancer registries (from 1,546 to 6,296 observations), the coefficient of correlation was high, ranging from 0.918 to 0.997. Also the distance measures were very low. Very similar results were obtained for major cancer sites, and sexes. The need for statistical tests, not influenced by sample size, was confirmed.
Conclusions
The data analyzed in this study had already been checked and approved for publication in Cancer Incidence in 5 Continents-X. Therefore, their quality was expected to be good. This study demonstrated that cancer incidence rates follow Benford's law. This suggests using the adherence to Benford's law of the FSD distribution of incidence rates as a quick tool in their quality evaluation, in order to identify possible deviations for further investigations.

Oral presentation, 6th October 2016

Cancer incidence rates and Benford's law: a useful liaison

ENCR Scientific Meeting and General Assembly. Joining forces for better cancer registration in Europe 5-7 October 2016, Baveno, Italy
http://publications.jrc.ec.europa.eu/repository/bitstream/JRC102769/encr%202016%20conference%20book%20%28print%29%20%28secured%29_2.pdf

Emanuele Crocetti, Giorgia Randi, Raquel Carvalho, Tadeusz Dyba, Francesco Giusti, Carmen Martos, Roisin Rooney, Lydia Voti, Manola Bettio

European Commission, DG Joint Research Centre, Ispra, Italy

Background and Introduction
In 1938 Benford described an odd distribution of the first significant digit (FSD) in many numerical collections: the probability to have 1 as FSD is 30.1 % then it slightly and consistently lowers up to 9 which is least frequent FSD. This pattern is already used to identify possible violations in numerical data (e.g. in accounting). We evaluated whether population-based cancer incidence rates follow Benford's law (BL), to detect possible violations during data quality assessment of cancer registry (CR) data.
Materials and Methods
We randomly sampled from CI5C-X web site the detailed databases of two population-based CRs
for each of the following regions: Africa, north
and south America, Asia, Europe and Oceania. The distribution of the FSD of crude incidence rates was evaluated for each separate registry, and for all of them together in a single dataset. The observed FSD distribution was plotted against the Benford theo- retical one, and the following statistics were computed: Person's r, distances' measures, and Chi2/n to check if the data are well-modelled by BL. A summary index was also computed (the lowest the index, the best the fitting).
Results
The distributions of FSD of crude incidence rates (overall on 40493 observations) showed a mean greater than the median and a positive skewness, typical of Benford-like distributions. In fact, it fitted almost perfectly Benford distribution (r = 0.997; m=0.01; d*=0.02; Chi2/n=0.05). Individual selected CRs (from 779 to 5 376 observations) had generally very good fitting; however, one registry had all the four statistics in the worst duo-decile (p=0.00005).
Conclusions
Crude cancer incidence rates adhere to BL. This suggests using BL as a quick, easy and objective screening tool for assessing CR data quality. The CR with the worst adherence to BL had a warning also in CI5C-X. We propose to use the BL as a screening tool in cancer data quality evaluation, identifying anomalies worthy of further inspection.

Poster, 5-7 October 2016

Cancer registries should additionally provide information on cancer incidence geographical variability

ENCR Scientific Meeting and General Assembly. Joining forces for better cancer registration in Europe 5-7 October 2016, Baveno, Italy.
http://publications.jrc.ec.europa.eu/repository/bitstream/JRC102769/encr%202016%20conference%20book%20%28print%29%20%28secured%29_2.pdf

Emanuele Crocetti, Giorgia Randi, Raquel Carvalho, Tadeusz Dyba, Francesco Giusti, Carmen Martos, Roisin Rooney, Lydia Voti, Manola Bettio

European Commission, DG Joint Research Centre, Ispra, Italy

Background and Introduction
The frequency of cancer in a defined population and in a certain period is reported by cancer reg- istries (CR) as a rate, age-standardised to allow for reliable comparisons. This is the standard statistic computed by all CRs independent of the size of the population at risk. The measure of variability provided – standard error – refers to the precision of the estimator and does not reflect the heterogeneity in cancer incidence within the area. Some national CRs are publishing incidence rates at a lower geo- graphical level providing thus more insight into intra-CR incidence variability.
Materials and Methods
We retrieved from Nordcan (http://www-dep.iarc. fr/NORDCAN) the European age-standardised incidence rates (ASR) for Denmark and Finland in 2013, at national level as well as for five regions in each country, for all sites excluding non-melanoma skin cancer, for men. Differences between and within countries were evaluated.
Results
The national ASR for Denmark was 479.4, and for Finland 422.2 per 100000 py. The standard errors, 3.68 and 3.46, were negligible due to a big number of incident cases considered, 17 582 and 15 517 respec- tively. ASRs for the Danish regions ranked from 458.7 in Zealand to 495.9 in Southern Denmark (range: 37.2). As regards Finland, regional ASRs varied between 362.1 in Turku to 490.4 in Tampere, with a range of 128.3. Significant differences were observed both between and within countries.
Conclusions
Although both Danish and Finnish ASRs are cor- rect, the analysed example demonstrates that the national ASR may reflect more (Denmark) or less accurately (Finland) the incidence of cancer in the different regions of a country. When heterogeneity is present regional rates are more informative than the national ones. The unavailability of a unique population-unit for sub-areas makes comparisons difficult. However, CR should start to deal with the need to provide information on internal cancer incidence variability as well as just incidence.

Poster, 19-21 October 2016

No hints on geographical heterogeneity of cancer incidence in cancer registry rates.

The 38th Annual IARC Conference Marrakech (Morocco) 19 - 21 October 2016
http://www.iacr.com.fr/images/AnnualMeetings_Abstracts/20161012_IARC_WEB-Abstracts-Marrakesh_V8.pdf

Francesco Giusti, Emanuele Crocetti, Giorgia Randi, Raquel Negrao Carvalho, Tadeusz Dyba, Carmen Martos, Roisin Rooney, Lydia Voti, Manola Bettio

Joint Research Centre / European Commission, Ispra (VA), Italy

Background
As a standard practice worldwide, cancer registries (CRs) express the frequency of cancer in a defined population and in a certain period as a rate, independently of the size of the population at risk. However, this single measure may not describe the variability of incidence within a country. In fact, the only measure of variability provided – the standard error - refers to the precision of the estimator.
Methods
We retrieved from the CDC website
(https://nccd.cdc.gov/uscs/) age-adjusted (US 2000) incidence rates (ASR) and 95 confidence intervals (CI) for all cancer sites combined, for the period 2008-12, in the whole United States (US). We compared them across sub-geographical areas (4), regions, states (50) and a few cities, using the overlap between CI of ASR.
Results
The overall US ASR, for men and women together, was 462.0 cases per 100.000 person/year (95 CI 461.6-462.3). The national rate was lower (95 CI did not overlap) than the ASRs in Northeast and Midwest and higher (95 CI did not overlap) than in South and West regions. 26 states had incidence rates higher, 18 lower and six states within the variability of the national average rate. Additionally, within states, the overall ASR did not reflect accurately single cities (e.g. California and sub-areas)
Conclusions
In this exploratory analysis, the US ASRs are used to demonstrate how a national (or supranational) ASR may not reflect incidence of subareas. Although we did not statistically quantify this variability major differences appeared between and within regions, states and cities. When heterogeneity is present among regions, regional rates are more informative than the single national one. CRs should then start providing information on internal cancer incidence variability as well as on incidence level.

Poster, 19-21 October 2016

Cancer incidence rates are numbers which must abide by mathematical laws.

The 38th Annual IARC Conference Marrakech (Morocco) 19 - 21 October 2016
http://www.iacr.com.fr/images/AnnualMeetings_Abstracts/20161012_IARC_WEB-Abstracts-Marrakesh_V8.pdf

Giorgia Randi, Emanuele Crocetti, Raquel Negrao Carvalho, Tadeusz Dyba, Francesco Giusti, Carmen Martos, Roisin Rooney, Lydia Voti, Manola Bettio

Joint Research Centre / European Commission, Ispra (VA), Italy

Background
Some numerical series abide by Benford's law (BL). BL describes the distribution of the first significant digit (FSD) of these numbers, which is unexpectedly skewed towards small figures. Violations of BL are already considered evidences in trials for frauds in accounting. We evaluated whether population-based cancer incidence rates follow BL, to use possible violations during the quality assessment of cancer registry (CR) data.
Methods
We randomly sampled from CI5C-X website the detailed databases of two population-based CRs for each of the following regions: Africa, north and south America, Asia, Europe and Oceania. The distribution of the FSD of crude incidence rates was evaluated for each registry separately, as well as for all of them together in a single dataset. The observed FSD distribution was plotted against the Benford theoretical one, and the following statistics were computed: Person's r, distances' measures, and Chi2/n to check if the data were fitting Benford's distribution. A summary index was also computed (the lower the index the best the fitting).
Results
The distributions of FSD of crude incidence rates (overall on 40493 observations) showed a mean greater than the median and a positive skewness, typical of Benford-like distributions. In fact, FSD of rates fitted almost perfectly BL (r=0.997; m=0.01; d*=0.02; Chi2/n = 0.05). Single CRs (having from 779 up to 5376 observations) had generally very good fitting; however, one registry had all the four statistics in the worst decile (p=0.00005).
Conclusions
Crude cancer incidence rates adhere to BL. BL is very simple, quick and easy to be understood and computed. Moreover, it does not rely on subjective opinion or personal professional expertise of any reviewer. Therefore, we propose to add the use of BL, as an objective screening tool in cancer data quality assessment, to identify anomalies worthy of further inspection.

Oral presentation, 25th May 2017

Incidence rates and intra-area variability.

Grell - Group for cancer epidemiology and registration in Latin language countries. Ascension Reunion, 24-26 May 2017, Brussels

Emanuele Crocetti, Francesco Giusti, Giorgia Randi, Tadeusz Dyba, Carmen Martos, Manola Bettio

European Commission, Directorate General Joint Research Centre, Directorate F – Health, Consumers and Reference Materials, Health in Society Unit Via E. Fermi, 2749. 21027 Ispra (VA), Italy

Objectives.
Usually the new cases of cancer diagnoses within a specific area are reported in the form of incidence rate. The standard error (SE) of the rate expresses the precision of the estimator. Incidence rates in themselves do not provide any information on the variability in a geographical comparison among different areas. The objective of this study is the evaluation of this issue, and related suggestion for a measure reflecting correctly this variation.
Methods.
We retrieved from Nordcan (www-dep.iarc.fr/NORDCAN) the age standardised incidence rates (ASR, European population) and 95% confidence intervals (CI) for all cancer sites but non-melanoma skin cancers, in 2014, among men. We compared, looking at 95% CI overlaps, the overall ASR for Nordic Countries with each country-level ASRs, and the latter with the Regional ASRs. We tested as a summary measure to report on variation the ratio (r) between ASRs for sub-areas and the overall ASR (R) of the area, r/R.
Results.
The overall ASR for Nordic Countries is 453.1 cases for 100.000 inhabitants (SE 1.6); 95% CI of this rate are above those of Faroe Islands (251.0; 28.8), Finland (404.5; 3.4), Iceland (387.2; 14.8) and Sweden (428.5; 2.6) and below those of Denmark (504.4; 3.7) and Norway (509.9; 4.1). This heterogeneity becomes evident when looking at the r/R value of 46%, meaning that the range between ASRs of single Countries is 46% of the overall ASR. The same pattern applies for the comparison within Countries, where the r/R is small in Iceland (r/R 10%), Denmark and Norway (14%), and wider in Finland (32%) and Sweden (38%).
Conclusions.
This exploratory analysis confirms that the overall ASR for a well-defined area may not reflect correctly the variation occurring among the different sub-areas. The adoption of the proposed r/R ratio in addition to the traditional ASTs would help to underline such heterogeneity.

Appendix 2: Sub-analysis originally not included in the presented papers ('data not shown').

Addendum to paper n. 1

Sub-analysis by sex.

Comparison of expected (Benford's law) distribution of first significant digits with observed ones by sex.

Sex: men (n. 70,722 observations)

Table 16: Incidence rates for men: comparison of the observed and expected percentage distribution of each first significant digit.

First digit	Observed	Expected
1	31.4	30.1
2	17.54	16.61
3	12.52	12.49
4	9.6	9.6
5	7.7	7.9
6	6.3	7.7
7	5.8	5.8
8	4.7	5.1
9	4.7	4.6
Mean	3.4	3.4

Pearson correlation coefficients (r) = 0.999

Maximum distance in absolute terms between expected and observed frequencies for each of the nine digits (m) = 0.01267

χ^2/n: $\chi 2$ divided by the sample size = 117.2

the normalised Euclidean distance between the two distributions divided by the maximum possible distance (d*) = 0.0143

Z statistic (z) = 0.82

Sex: women (n. 75,868 observations).

Table 17: Incidence rates for women: comparison of the observed and expected percentage distribution of each first significant digit.

First digit	Observed	Expected
1	31.6	30.1
2	17.7	16.61
3	12.1	12.49
4	9.2	9.6
5	8.1	7.9
6	6.0	7.7
7	5.8	5.8
8	4.8	5.1
9	4.8	4.6
Mean	3.4	3.4

Pearson correlation coefficients (r) = 0.999
Maximum distance in absolute terms between expected and observed frequencies for each of the nine digits (m) = 0.01461
χ^2/n: $\chi 2$ divided by the sample size = 117.2
the normalised Euclidean distance between the two distributions divided by the maximum possible distance (d*) = 0.0174
Z statistic (z) = 1.22

Sub-analysis by registries.
Comparison of the expected (Benford's law) mean (3.44) and median (3) of first significant digits distribution with the observed values by registries.
The expected (Benford's law) values are: 3.44 for the mean and 3 for the median.

Table 18: Distribution of first significant digits by registries: observed mean and median.

Registry	Mean	Median
1	3.5	3
2	3.5	3
3	3.5	3
4	3.4	3
5	2.8	2
6	3.7	3
7	3.5	3
8	3.7	3
9	3.5	3
10	3.6	3
11	3.2	2
12	3.4	2
13	3.1	2
14	3.4	3
15	3.6	3
16	3.1	2
17	3.3	3
18	3.3	2
19	3.4	2
20	3.1	2
21	3.4	3
22	3.2	2
23	2.9	2
24	3.3	2
25	2.9	2
26	3.5	2
27	3.6	3
28	4.2	3
29	3.5	3

30	3.6	3
31	3.4	2
32	3.6	3
33	3.2	2
34	3.7	3
35	3.2	2
36	3.6	3
37	3.0	2
38	3.2	2
39	3.5	3
40	3.5	3
41	3.3	2
42	3.2	2
43	3.2	2

Addendum to paper n. 2

Table 19: Country layers, number of incident cases of 'all sites excluding non-melanoma skin cancer', in men, in 2012, resident population, European age-standardised incidence rates (ASR), standard error (SE) of the ASR, lower (LCI) and upper (UCI) 95% confidence intervals and R/R (Range/Rate : Range is the absolute difference in ASRs between the greatest and lowest ASR for areas in the lower layer; Rate is the ASR) Data from Nordcan (http://www-dep.iarc.fr/NORDCAN).

Area	Numbers	ASR(E)	SE	LCI	LCS	range	r/R
Nordic countries	76,603	452.6	1.67	449.3	455.9	152.9	0.34
Denmark	18,819	522.6	3.87	515.0	530.2	52.5	0.10
North Jutland	2,107	523.3	11.68	500.4	546.2		
Central Jutland	3,887	483.4	7.85	468.0	498.8		
Southern	4,200	521	8.12	505.1	536.9		
The capital	5,300	535.9	7.44	521.3	550.5		
Zealand	3,294	521	9.46	502.5	539.5		
Faroe Islands	112	373.2	35.62	303.4	443.0		
Finland	14,865	415.1	3.47	408.3	421.9	61.9	0.15
Helsinki	4,638	431.1	6.4	418.6	443.6		
Kuopio	2,227	363.7	7.95	348.1	379.3		
Oulu	1,935	386.2	8.96	368.6	403.8		
Tampere	3,277	425.6	7.6	410.7	440.5		
Turku	2,778	445.1	8.68	428.1	462.1		
Greenland	94	380.9	41.62	299.3	462.5		
Iceland	755	440.1	16.24	408.3	471.9	1.5	0.00
Reykjavik-Reykjanes	504	441.6	19.9	402.6	480.6		
Outside the capital	251	437.1	28.14	381.9	492.3		
Norway	15,634	526.1	4.26	517.8	534.4	64.8	0.12
Central	2,233	517.2	11.13	495.4	539.0		
Northern	1,470	479.9	12.69	455.0	504.8		
South-Eastern	8,743	530.6	5.74	519.3	541.9		
Western	3,188	544.7	9.74	525.6	563.8		
Sweden	26,530	404.8	2.56	399.8	409.8	64.7	0.16
Northern	2,528	376	7.8	360.7	391.3		
Stockholm-Gotland	5,284	420.1	5.84	408.7	431.5		
Southern	5,217	436.3	6.25	424.1	448.6		
South-Eastern	3,109	422.5	7.98	406.9	438.1		
Uppsala-Örebro	5,515	371.6	5.22	361.4	381.8		
Western	4,877	402.9	5.95	391.2	414.6		

Addendum to paper n. 3

Table 20: Time trends of median thickness for skin melanoma thick >=3 and >=4 mm.

Thickness	1985-1989	1990-1994	1995-1999	2000-2004
>=3mm	4.0	4.5	4.5	4.68
>=4 mm	5.1	6.0	5.35	5.8

Appendix N. 3: List of the six published papers discussed in the thesis.

frontiers in Public Health

ORIGINAL RESEARCH
published: 13 October 2016
doi: 10.3389/fpubh.2016.00225

Using the Benford's Law as a First Step to Assess the Quality of the Cancer Registry Data

Emanuele Crocetti and Giorgia Randi*

Health in Society Unit, Directorate F Health, Consumers and Reference Materials, Joint Research Centre (JRC), European Commission, Ispra, Italy

Background: Benford's law states that the distribution of the first digit different from 0 [first significant digit (FSD)] in many collections of numbers is not uniform. The aim of this study is to evaluate whether population-based cancer incidence rates follow Benford's law, and if this can be used in their data quality check process.

Methods: We sampled 43 population-based cancer registry populations (CRPs) from the Cancer Incidence in 5 Continents-volume X (CI5-X). The distribution of cancer incidence rate FSD was evaluated overall, by sex, and by CRP. Several statistics, including Pearson's coefficient of correlation and distance measures, were applied to check the adherence to the Benford's law.

Results: In the whole dataset (146,590 incidence rates) and for each sex (70,722 male and 75,868 female incidence rates), the FSD distributions were Benford-like. The coefficient of correlation between observed and expected FSD distributions was extremely high (0.999), and the distance measures low. Considering single CRP (from 933 to 7,222 incidence rates), the results were in agreement with the Benford's law, and only a few CRPs showed possible discrepancies from it.

Conclusion: This study demonstrated for the first time that cancer incidence rates follow Benford's law. This characteristic can be used as a new, simple, and objective tool in data quality evaluation. The analyzed data had been already checked for publication in CI5-X. Therefore, their quality was expected to be good. In fact, only for a few CRPs several statistics were consistent with possible violations.

Keywords: cancer registry, incidence, data quality, Benford, methodology

OPEN ACCESS

Edited by:
Ming Wu,
Jiangsu Provincial Center for Disease Control and Prevention, China

Reviewed by:
Stefano Guzzinati,
Istituto Oncologico Veneto, Italy
Xiaojin Yu,
Southeast University, China

****Correspondence:***
Emanuele Crocetti
emanuele.crocetti@ec.europa.eu

Specialty section:
This article was submitted to Epidemiology, a section of the journal Frontiers in Public Health

Received: *11 August 2016*
Accepted: *28 September 2016*
Published: *13 October 2016*

Citation:
Crocetti E and Randi G (2016) Using the Benford's Law as a First Step to Assess the Quality of the Cancer Registry Data. Front. Public Health 4:225. doi: 10.3389/fpubh.2016.00225

INTRODUCTION

The Benford's law (1), originally identified by Newcomb (2), states that in many numerical series the distribution of the first significant digits (FSDs) (the first non-zero digit on the left side of a number) is not uniform. In fact, for numbers which adhere to this law, the probability of 1 to be the FSD is 30.1%, and this probability steadily decreases for the following digits up to 9, which is the least common leading digit (4.6% of the cases). A distribution abides by the Benford's law if the frequency $[F(x)]$ of the FSD, $x \in \{1, \ldots, 9\}$, follows the logarithmic relation, $F(x) = \log_{10}\left(1 + \frac{1}{x}\right)$ (1). The law of "anomalous numbers" applies also to the frequency of digits in other positions (1).

Not all the numbers abide by the Benford's law, but for those which do, violations raise concerns. For example, in accounting and auditing, also at a Governmental level, the Benford's law has been widely used to detect possible frauds (3–5).

Population-based cancer registries produce a great amount of numbers: the cancer incidence rates. The evaluation of their quality is rather complex, involving different aspects, and it is mainly based on the knowledge of the clinical, diagnostic, and therapeutic pathways of patients and on the process of data collection and registration (6, 7).

The most renowned publication on cancer incidence is Cancer Incidence in 5 Continents (CI5) (8). The cancer registries submitting their data to CI5 have to pass a formal quality evaluation before being accepted. The data quality assessment implies checking the internal coherence, consistency, completeness, and comparability with the final decision taken by a group of experts in the field.

The aim of this study is to evaluate if cancer incidence rates adhere to the Benford's law to use this mathematical characteristic as a further and objective tool for their quality evaluation.

MATERIALS AND METHODS

In the website of the CI5 volume X (CI5-X) (9), the data of the 290 population-based cancer registries included in the publication are available, detailed by all the 424 cancer registry populations (CRPs), as each cancer registry can provide information not only for the whole population but also for different racial and/or ethnic subgroups within the same population.

The CI5-X data include aggregated information for 244 combinations of cancer site and morphological group, specified for 19 age groups (5-year age groups from 0–4 to 85+, plus unknown age) and for the two sexes.

We drew a pseudorandom sample of 10% of the available CRPs, stratified by continent (considering South and North America separately), setting a random number seed to make the sampling reproducible.

Overall, 43 CRPs (from 40 cancer registries) were sampled and included in the analysis: 1 from Africa (Malawi, Blantyre), 3 from Central and South America (Argentina, Tierra del Fuego; Brazil, San Paolo and Ecuador, Quito), 18 from USA (Virginia, Asian and Pacific Islanders; Nebraska, Black; Ohio; Vermont; Montana; Michigan; Georgia; Indiana, White; Missouri, White; NPCR-National program of cancer registries – including 42 States; Colorado, Asian and Pacific Islanders; Arkansas, Black; Alabama, White; Arkansas, White; California, Asian and Pacific Islanders; Connecticut, Black; Virginia, Black; and California), 7 from Asia (India, Karunagappally; Singapore, Malay; Turkey, Edirne; Israel, Jews; Japan, Hiroshima Prefecture; Japan, Fukui Prefecture; and Israel), 11 from Europe (France, Isère; Germany, North Rhine – Westphalia; France, Hérault; UK, England; Estonia; Switzerland, St Gall-Appenzell; Bulgaria; Malta; Ukraine; Spain, Navarra; and Italy, Sondrio), and finally 2 from Oceania (New Zealand; Other and USA, and Hawaii).

The cancer data corresponding to the age group 19 (age unknown) were excluded from the analysis.

After the exclusion of those combinations of cancer morphology and site with no cases, 146,590 combinations were included in the analysis.

Crude incidence rates were computed for each sex, age group, and topography and morphology combination dividing the number of cases by the corresponding population, and expressed per 100,000 inhabitants. The FSD distribution for crude incidence rates was then calculated for all the CRPs together, by sex, and by CRP. Moreover, a sensitivity analysis has been performed randomly excluding half of the most important cancer sites (prostate, lung, breast, and colon–rectum).

For checking the adherence of observed FSD distributions to the Benford's one, we used different methods.

Since the Benford's distribution has mean greater than median and is positively skewed (10), these figures have been evaluated for cancer incidence rates.

Theoretical and observed distributions were plotted for a graphical comparison.

According to the literature, we did not use those tests (e.g., χ^2 and the Kuiper's statistic) that are extremely sensitive in rejecting the null hypothesis (being a distribution Benford-like) for large samples (4, 11–13). To test the goodness of fit, we used the following tests:

- r: the Pearson correlation. This is commonly used to measure how closely a distribution follows the Benford's law (11, 12). The most the coefficient "r" is close to +1 the highest the correlation between Benford's law and the observed FSD distribution is.
- χ^2/n: the χ^2 divided by the sample size (4, 14).
- m: the maximum distance in absolute terms between expected and observed frequencies for each of the nine digits (1–9). The statistics may vary between 0 (no differences between the two distributions) to $+\infty$ (maximum difference) and the corresponding formula is $m = \max_{i=1,2,\ldots,9}\{|b_i - e_i|\}$ (12), where b_i is the frequency expected by Benford and e_i is the observed frequency for each digit i.
- d^*: the normalized Euclidean distance between the two distributions divided by the maximum possible distance, which would occur when the FSD was 9 for all the numbers. The corresponding formula is:

$$d^* = \sqrt{\sum_{i=1}^{9}(b_i - e_i)^2} \Big/ \sqrt{\sum_{i=1}^{8}(b_i)^2 + (1 - e_9)^2}$$

  where b_i is the frequency expected by Benford and e_i is the observed frequency for each digit i. The statistic may vary between 0 (no differences) to 1 (maximum difference) (12).
- Z statistic: the average of the Z values for each comparison between the nine observed and theoretical digits distributions (5):

$$Z = \frac{1}{9}\sum_{i=1}^{9}\sqrt{n}\left[\frac{|b_i - e_i| - 1/(2n)}{\sqrt{b_i\,(1 - b_i)}}\right]$$

where $i = 1, \ldots, 9$ is a fixed digit, b_i is the frequency expected by Benford, and e_i is the observed frequency for each digit i. The

cut-off value for statistical significance, with alpha = 0.05 and one side tail, is 1.64.

For providing an inter-CRP comparison, the mean, the median, and the 10th or the 90th (the one including the most extreme values) percentile of each statistic were computed.

A summation index has been computed for rating the CRPs according to the statistics' results. Each CRP received one point for each statistic in the 10th or 90th percentile (whichever represents the worst values). The summation index could vary from 0 (no statistic beyond the threshold) up to 5 (all statistics beyond the threshold). The probability for each statistic to be in the most extreme decile was 0.1 (approximately 4/43) assuming independence between statistics, considering that the summation index follows a binomial distribution (pr = 0.1, $n = 5$) the random probability for a CRP to have the summation index equal to 0 is 0.59, to 1 is 0.33, to 2 is 0.07, to 3 is 0.008, to 4 is 0.0005, and to 5 is 0.00001.

The analysis has been performed with Stata v. 12, using specific commands for extracting the sample ("sample" and "seed") and for computing observed and Benford FSD distributions ("digdis").

RESULTS

When considering all the cancer incidence rates together (146,590 observations), the distribution of the FSDs appeared to be positively skewed (0.84), with the mean (3.38) greater than the median (3.0). These values were close to those of the theoretical Benford's distribution (skewness 0.8, mean 3.44, and median 3.0), as were the ratios between 1st vs. 9th (observed 6.6 vs. Benford 6.6), and between 1st vs. 2nd (1.8 vs. 1.7) FSD.

These results let suppose that the FSD distribution of cancer incidence rates might adhere to the Benford's pattern. In fact, when the observed FSD distribution was graphically compared to the theoretical one, as shown in **Figure 1**, they were almost overlapping.

The Pearson's correlation coefficient, r, showed an almost perfect direct correlation between the observed FSD distribution and the expected one (0.999); moreover, all the measures of the distance between the distributions were very low ($m = 0.014$ and $d^* = 0.015$), and the average Z was below the significance level. Finally, the χ^2 test, weighted on the number of observations (χ^2/n), was also very low (0.002).

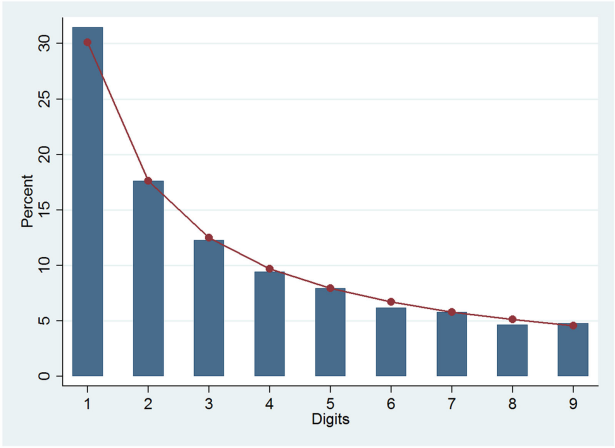


FIGURE 1 | Benford (line) and observed (columns) distributions of first digits for all crude cancer incidence rates.

The analysis has been repeated by sex and confirmed the same results (data not shown). Also, after the exclusion of half of the rates for the major cancer sites were excluded the overall results confirmed the adherence of the FSD distribution to Benford's law ($r = 0.999$, $m = 0.014$, $d^* = 0.016$, $\chi^2/n = 0.002$).

When single CRPs were evaluated, each FSD distribution was positively skewed, and the mean was greater than the median (data not shown).

In **Figure 2**, the FSD distribution of all cancer incidence rates and the Benford distribution were compared for each of the 43 analyzed CRPs. The shapes of all distributions generally resembled the Benford's one, with a decreasing percentage of FSD from 1 to the 9. However, a few possible differences were shown.

The Pearson correlation coefficients were very high for the majority of the CRPs (median = 0.97); however, some values were relatively low (0.85 representing the 10th percentile). Also, the other measures of distance were generally low (median: $m = 0.05$, $d^* = 0.07$), but still the corresponding 90th percentiles reached rather higher values (0.10 and 0.12, for m and d^*, respectively). For the ratio between the χ^2 and the number of rates, the 90th percentile was almost 3-time the median (90th percentile = 0.14 and 50th percentile = 0.05), and, finally, for the average Z, the value of the 90th percentile corresponded to the value of statistical significance (1.64).

Although the majority of the CRPs reported statistics showing an agreement with the Benford's law, for a few of them, the values seemed to indicate possible discrepancies.

For 35 CRPs, the summation index was 0, for 4 CRPs was 1, and for 2 CRPs was 2. Only one CRP (Argentina, Tierra del Fuego) reported a summation index of 3 ($r = 0.839$; $d^* = 0.125$; $\chi/n = 0.146$) and another one (USA, Virginia, Black) had all the five statistics in the worst classes ($r = 0.82$; $m = 0.13$; $d^* = 0.147$; $\chi/n = 0.182$; $Z = 1.88$). The probability for the two latter results to happen by chance is very low. Therefore, for such CRPs, a possible violation of the Benford's law should be considered.

DISCUSSION

In the present study, a considerable and heterogeneous sample of CRPs, included in CI5-X, was analyzed to evaluate, for the first time to our knowledge, if the FSD distribution of cancer incidence rates abided by the Benford's law.

The results showed a substantial adherence of FSD distribution of cancer incidence rates to the Benford's law.

This was not surprising. In fact, FSD distribution of cancer incidence rates had *a priori* some characteristics for being Benford prone. Indeed, they are the second generation distribution, being the result of the division of the number of cases diagnosed in a time span by the corresponding resident population, they comprise a large range of numbers covering several orders of magnitude (from units to thousands per 100,000 people, according to different ages and cancer types), and they are not influenced by human thought (15, 16).

FIGURE 2 | Theoretical (line) and observed distributions (columns) of first digits for all the analyzed incidence rates, by registries and populations (reg). 1: Malawi, Blantyre; 2: Argentina, Tierra del Fuego; 3: Brazil, San Paolo; 4: Ecuador, Quito; 5: USA, Virginia, Asian and Pacific Islanders; 6: USA, Nebraska, Black; 7: USA, Ohio; 8: USA, Vermont; 9: USA, Montana; 10: USA, Michigan; 11: USA, Georgia; 12: USA, Indiana, White; 13: USA, Missouri, White; 14: USA, NPCR-National program of cancer registries (including 42 States); 15: USA, Colorado, Asian and Pacific Islanders; 16: USA, Arkansas, Black; 17: USA, Alabama, White; 18: USA, Arkansas, White; 19: USA, California, Asian and Pacific Islanders; 20: USA, Connecticut, Black; 21: USA, Virginia, Black; 22: USA, California; 23: India, Karunagappally; 24: Singapore, Malay; 25: Turkey, Edirne; 26: Israel, Jews; 27: Japan, Hiroshima Prefecture; 28: Japan, Fukui Prefecture; 29: Israel; 30: France, Isère; 31: Germany, North Rhine – Westphalia; 32: France, Hérault; 33: UK, England; 34: Estonia; 35: Switzerland, St Gall-Appenzell; 36: Bulgaria; 37: Malta; 38: Ukraine; 39: Spain, Navarra; 40: Italy, Sondrio; 41: Germany, Brandemburg; 42: New Zealand: Other; 43: USA, Hawaii.

We verified that cancer incidence rates respect the quantitative measures suggested by Wallace (10) to assess whether a distribution may be expected to obey the Benford's law. In fact, the mean of their observed FSD is greater than the median, and their distribution has a positive skewness.

In the present study using graphical visualization, correlation coefficient, and some distance statistics, we observed that FSD distribution of cancer incidence rates abide by the Benford's law when analyzed overall, by sexes, excluding half of the rates for the major cancer sites (female breast, colon–rectum, and lung and prostate cancers) and generally by CRP.

We have analyzed data which had been already examined for their quality and proved as good for publication in CI5-X (8). Therefore, no major problems in data quality were expected. In fact, our results showed that for almost all the CRPs the FSD distribution substantially adhere to the Benford's law. When the 43 CRPs were analyzed individually, the plot of their FSD distribution seemed to be in agreement with the Benford's law. It must be mentioned that, due to sampling, two CRPs were subgroups of the same registry (USA, Arkansas Black and White; USA Virginia, Asian and Pacific Islanders and Black) and two others included a subgroup and the whole population of the same registry (Israel and Israel, Jews and USA California and USA California, Asian and Pacific Islanders). No large difference within those cancer registries has been shown. Therefore, quality of cancer registry data and related activity (in terms of data availability, data collection, etc.) seemed not related to racial/ethnic subgroups at least in the analyzed registries.

The cancer registry data quality evaluation is not a perfect process, and some residual heterogeneity could exist also in CRPs included in CI5-X. In fact, in the introduction of CI5-X, it is stated that in the registry specific pages for some CRPs "an asterisk

preceding the registry title indicates that special considerations (which may include underregistration) must be taken into account in interpreting the published rates or indicators of quality. . ." (8). Overall, asterisks were reported for 114/424 CRPs in CI5-X (26.9%), and in 11/43 (25.6%) in our sample. One of the two CRPs which had three or more statistics with the worst values for Benford's compliance had the asterisks (50%), in comparison with the others in the sample (10/31, 24.3%).

We evidenced that, although the majority of CRPs seemed to adhere to the Benford's law, at least two of them showed possible violation. Random fluctuations could have driven the observed results (14), even if with a very low probability, but the coherence across the different applied statistics made, for these CRPs, the inconsistency with the Benford's distribution more probable.

According to our experience, based on the analyzed dataset that has been already checked for data quality and accepted for publication (CI5-X), cancer registries showing the poorest results had r value below 0.9 and m, χ^2/n, and d^* values higher than 0.10; presumably in a wild situation, greater values are expected.

The adherence to Benford's law has been widely used not only to detect fraudulent data in business and administration (3) but also to test data irregularities in scientific research (17). Frauds in cancer incidence data are not expected. However, non-adherence to the law may be a clue for further evaluation. The distance from the expected distribution may be the consequence of selections or incompleteness of the data collection, of rounding of small rates (18), of errors in data recoding or in data transfer.

The meaning of Benford's violation is a red flag showing an unusual behavior requesting further data examination (14). Once the suspect for a violation is raised, a CRP, which owns more data than those we analyzed, should try to find out clues for the possible problem. Our suggestion is to look for the Benford patter for incident cases based on different (combinations of) sources of information (pathology reports, hospitalization, death certificate, etc.) to detect any source-specific pattern. Moreover, it should be evaluated the stability over time of the data flow, for each information source and cancer site.

CONCLUSION

Checking for adherence to the Benford's law is not suggested in place of the traditional cancer registry data quality process, but it could be used as a simple and objective tool in the first steps to identify those cancer registries to be evaluated with great attention.

AUTHOR CONTRIBUTIONS

EC conceived the idea of the study, planned and designed it, and drafted the first draft. GR made substantial contribution to the statistical analysis and revised critically the paper. Both authors edited and approved the final version of the manuscript. Both authors are accountable for all aspects of the work.

REFERENCES

1. Bedford F. The law of anomalous numbers. *Proc Am Philos Soc* (1938) 78:551–72.
2. Newcomb S. Note on the frequency of use of the different digits in natural numbers. *Am J Math* (1881) 4:39–40. doi:10.2307/2369148
3. Tödter KH. Bedford's laws as an indicator of fraud in economics. *German Econ Rev* (2009) 10:339–51. doi:10.1111/j.1468-0475.2009.00475.x
4. Rauch B, Göttsche M, Brähler G, Engel S. Fact and fiction in the EU-governmental economic data. *German Econ Rev* (2011) 12:243–55. doi:10.1111/j.1468-0475.2011.00542.x
5. Slijepčević S, Blašković B. Statistical detection of fraud in the reporting of Croatian public companies. *Financ Theory Prac* (2014) 38:81–96. doi:10.3326/fintp.38.1.4
6. Bray F, Parkin MD. Evaluation of data quality in the cancer registry: principles and methods. Part I: comparability, validity and timeliness. *Eur J Cancer* (2009) 45:747–55. doi:10.1016/j.ejca.2008.11.032
7. Parkin MD, Bray F. Evaluation of data quality in the cancer registry: principles and methods. Part II: completeness. *Eur J Cancer* (2009) 45:756–64. doi:10.1016/j.ejca.2008.11.033
8. Forman D, Brewster DH, Kohler B, Piñeros M. Chapter 1: introduction. In: Forman D, Bray F, Brewster DH, Gombe Mbalawa C, Kohler B, Piñeros M, Steliarova-Foucher E, Swaminathan R, Ferlay J, editors. *Cancer Incidence in Five Continents, Vol. X*. Lyon: IARC (2013). 24 p. [IARC Scientific Publication No. 164].
9. Forman D, Bray F, Brewster DH, Gombe Mbalawa C, Kohler B, Piñeros M, et al., editors. *Cancer Incidence in Five Continents, Vol. X (Electronic Version)*. Lyon: International Agency for Research on Cancer (2013).
10. Wallace WA. Assessing the quality of data used for benchmarking and decision making. *J Gov Financ Manag* (2002) 51(3):16–22.
11. Golbeck J. Benford's law applies to online social networks. *PLoS One* (2015) 10(8):e0135169. doi:10.1371/journal.pone.0135169
12. Judge G, Schechter JG. Detecting problems in survey data using Benford's law. *J Human Resour* (2009) 44:1–24. doi:10.3368/jhr.44.1.1
13. Kienle S. What Benford can tell us about cover pools. An empirical analysis. *Int Businness Econ Res J* (2015) 14:829–34. doi:10.19030/iber.v14i6.9489
14. Leemis LM, Schmeiser BW, Evans DL. Survival distributions satisfying Benford's law. *Am Stat* (2000) 54:236–41. doi:10.2307/2685773
15. Hill PH. A statistical derivation of the significant-digit law. *Stat Sci* (1995) 10:354–63.
16. Durtschi C, Hillison W, Pacini C. The effective use of Benford's law to assist in detecting fraud in accounting data. *J Forensic Accounting* (2004) V:17–34.
17. Hein J, Zobrist R, Konrad C, Schuepfer G. Scientific fraud in 20 falsified anesthesia papers: detection using financial auditing methods. *Anaesthesist* (2012) 61:543–9. doi:10.1007/s00101-012-2029-x
18. Freedman DP. On rounding percentages. *J Am Stat Assoc* (1979) 74:359–64. doi:10.1080/01621459.1979.10482518

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

OPEN

Variability of cancer risk within an area: time to complement the incidence rate

Emanuele Crocetti, Francesco Giusti, Carmen Martos, Giorgia Randi, Tadeusz Dyba and Manola Bettio

The aim of this study was to show that age-adjusted cancer incidence rates for an area may not be representative of the incidence in subareas. We propose a simple measure to show the amount of geographical variability. European age-standardized incidence rates (ASRs) for 'all sites excluding nonmelanoma skin cancer', for men, in 2014, for Nordic countries as a whole, for each country (Denmark, Faroe Islands, Finland, Greenland, Iceland, Sweden and Norway) and for their regions, were retrieved from the Nordcan with corresponding standard errors SEs. We compared the ASR for Nordic countries versus single country and single country versus specific regions. The overlapping of 95% confidence intervals was used for ASRs comparisons. As a measure of variability, we computed the range between the highest and the lowest ASR within an area and the ratio between this range and the ASR of the overall area, $r/R = (\text{range}/\text{ASR}) \times 100$. The 95% confidence interval of the ASR for Nordic countries as a whole did not overlap those of the majority of the single countries; in fact, the r/R – which provides a clue for the amount of underlying geographical variability – was rather large (57.1%). Within countries, the variability was negligible in Iceland ($r/R = 9.6\%$), whereas the highest value was found in Sweden (37.1%). The ASR does not provide any information on underlying geographical variability. Therefore, its interpretation could be misleading. When data for subareas are available, the r/R, which is simple to compute and to understand, should be added to the ASR for providing more truthful information. *European Journal of Cancer Prevention* **26:442–446 Copyright © 2017 The Author(s). Published by Wolters Kluwer Health, Inc.**

European Journal of Cancer Prevention 2017, **26**:442–446

Keywords: age-adjusted, comparability, epidemiology, geographical variability, incidence, population-based cancer registries, rates

European Commission, DG Joint Research Centre, Ispra (VA), Italy

Correspondence to Emanuele Crocetti, MD, European Commission, DG Joint Research Centre, Via E. Fermi, 2749 – TP127, 21027 Ispra (VA), Italy
Tel: +39 0332 789926; fax: +39 0332 783858;
e-mail: emanuelecrocetti@yahoo.com

Received 16 March 2017 **Revised** 19 May 2017 **Accepted** 13 June 2017

Introduction

As a standard practice worldwide, population-based cancer registries (CRs) express the occurrence of cancer in a defined population in a certain period as the ratio between the newly diagnosed cancers and the at-risk resident population. This ratio is called the crude incidence rate (Boyle and Parkin, 1991). Cancer incidence increases with the ageing of the population. Therefore, incidence rates are strongly dependent on the age structure of the underneath population. Consequently, rates are computed using a standard age structure as a reference (age-standardized rate, ASR) to enable reliable comparisons across time and countries (Boyle and Parkin, 1991). Crude rates and ASRs are the standard indicators reported by all CRs independent of the size of the population at risk. These statistics are usually complemented by a measure of precision, the standard error (SE) of the rate and/or the 95% confidence intervals (CIs).

An incidence rate expresses the summary probability of developing cancer in the area covered by the CR. It provides no clues on the homogeneity or the heterogeneity of incidence rates across subareas.

To gain more insight into this topic and to explore the possible variability in ASRs among subareas of CRs, we analysed the data of the Association of the Nordic Cancer Registries, which makes data available in the Nordcan project (Engholm *et al.*, 2016).

Methods

We retrieved from the Nordcan the European ASRs for 'all sites excluding nonmelanoma skin cancer', for men, in 2014.

ASRs are available for three geographical layers as presented in Table 1: (a) Nordic countries as a whole; (b) single country: Denmark, Faroe Islands, Finland, Greenland, Iceland, Sweden, Norway; and (c) regions: five in Denmark: North Jutland, Central Jutland, Southern Denmark, The Capital and Zealand region; five in Finland: Helsinki, Kuopio, Oulu, Tampere and Turku region; two in Iceland: Reykjavik-Reykjanes and Outside The Capital;

DOI: 10.1097/CEJ.0000000000000389

Table 1 Data from Nordcan (*http://www-dep.iarc.fr/NORDCAN*): country layers, country/region name, number of incident cases of 'all sites excluding nonmelanoma skin cancer', in men, in 2014, resident population, European age-standardized incidence rates, SE of the ASR, lower and upper 95% confidence intervals and *r/R*

Layer	Area	*N*	Resident population	ASR (European)	SE	LCI	UCI	*r/R* (%)
1	Nordic countries	79 441	1 307 5123	453.1	1.6	449.9	456.3	57.1
2	Denmark	19 031	2 799 895	504.4	3.7	497.1	511.7	14.4
3	North Jutland	2004	292 697	474.8	10.9	453.4	496.2	
3	Central Jutland	4165	639 192	496.5	7.8	481.2	511.8	
3	Southern	4230	600 667	493.0	7.8	477.7	508.3	
3	The Capital	5245	860 818	507.1	7.1	493.2	521.0	
3	Zealand	3387	406 521	547.5	9.7	528.4	566.6	
2	Faroe Islands	78	25 039	251.0	28.8	194.6	307.4	
2	Finland	15 142	2 686 119	404.5	3.4	397.9	411.1	31.9
3	Helsinki	4849	922 582	427.2	6.2	415.0	439.4	
3	Kuopio	2194	403 920	345.5	7.7	330.5	360.5	
3	Oulu	1970	372 534	382.2	8.8	364.9	399.5	
3	Tampere	3811	545 749	474.6	7.9	459.1	490.1	
3	Turku	2303	441 350	348.5	7.5	333.7	363.3	
2	Greenland	99	29 742	384.1	40.6	304.5	463.7	
2	Iceland	694	164 257	387.2	14.8	358.1	416.3	9.6
3	Reykjavik	481	115 443	401.2	18.5	365.0	437.4	
3	Outside	213	48 818	364.1	25.4	314.3	413.9	
2	Norway	15 865	2 581 421	509.9	4.1	501.9	517.9	13.9
3	Central	2362	357 476	526.9	11.0	505.3	548.5	
3	Northern	1489	242 918	468.8	12.4	444.6	493.0	
3	South-Eastern	8685	1 433 445	502.9	5.5	492.2	513.6	
3	Western	3329	547 582	539.8	9.4	521.3	558.3	
2	Sweden	28 709	4 843 303	428.5	2.6	423.4	433.6	37.8
3	Northern	2521	444 391	364.5	7.6	349.6	379.4	
3	Stockholm-Gotland	6735	1 111 680	526.3	6.5	513.6	539.0	
3	Southern	5375	872 866	438.1	6.2	425.9	450.3	
3	South-Eastern	3130	510 943	415.8	7.7	400.6	431.0	
3	Uppsala-Örebro	5852	1 002 193	385.8	5.3	375.5	396.1	
3	Western	5096	901 258	407.1	5.9	395.6	418.6	

The smaller the *r/R* the lower the variability across subareas ASRs.
range/Rate% (*r/R*): range is the absolute difference in ASRs between the greatest and the lowest ASR of subareas in the lower layer; rate is the ASR.
ASR, age-standardized incidence rate; LCI, lower 95% confidence interval; UCI, upper 95% confidence interval.

six in Sweden: Northern, Stockholm–Gotland, Southern, South-Eastern, Uppsala-Örebro and Western region; and four in Norway: Central, Northern, South-Eastern and Western region.

The overall male resident population in Nordic countries in 2014 was 13 075 123. Residents in single countries and regions are shown in Table 1.

ASRs express the number of new cases diagnosed among 100 000 men in 2014 according to the observed age-specific rates and the age-groups of the European standard population.

We also retrieved from Nordcan the SE of the ASRs and we computed the 95% CIs according to the method of the binomial approximation (Boyle and Parkin, 1991) (Table 1).

We evaluated whether two rates were different inspecting the overlap between specific 95% CI (Schenker and Gentleman, 2001). The precision of the age-specific rates that concur in the calculation of ASR increases when the number of cases in this group increases. This applies to each age group and thus to ASR as the whole entity. The overall numbers observed yearly in the analysed series (Table 1) were, with the exception of Faroe Islands and Greenland, in the order of several hundreds or even thousands. SEs of the ASR are greater when the numbers on which they are based are small.

We compared the ASR at each geographical level with the level underneath: Nordic countries versus Denmark, Faroe Islands, Finland, Greenland, Iceland, Sweden, Norway and single country versus specific regions.

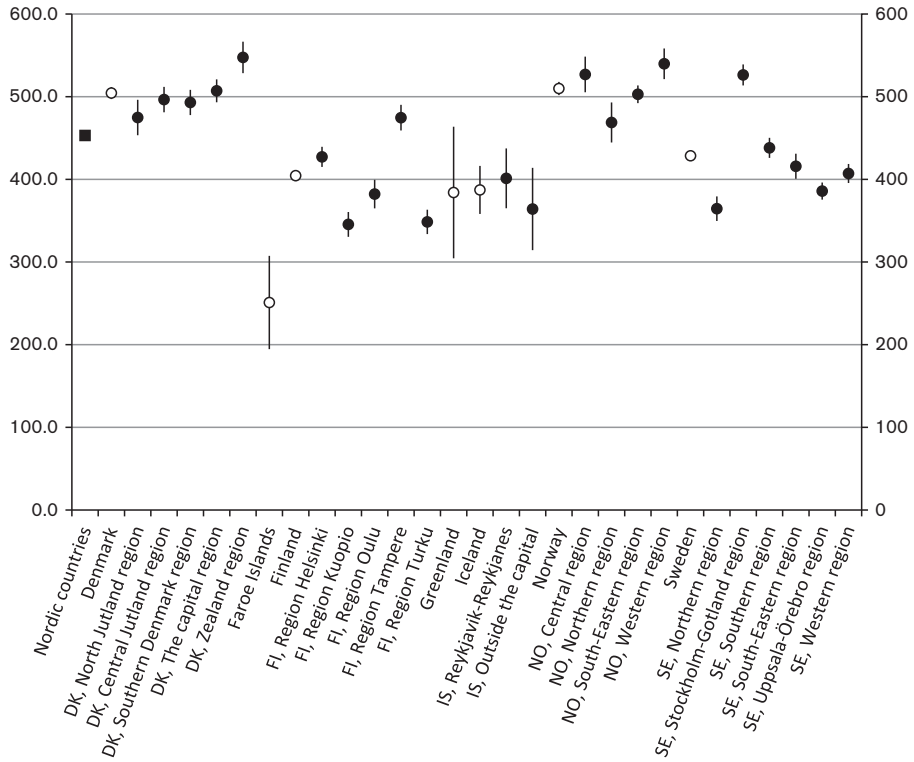
Moreover, we computed the absolute difference (range) between the highest and the lowest ASR within a nested layer (the range between countries for Nordic countries and between regions for a specific country). Then we calculated the percent ratio between this range and the ASR of the level above interpreted as a summary value of the subareas (for Nordic countries or a single country, respectively), $r/R = (\text{range}/\text{ASR}) \times 100$.

The *r/R* provides a measure of the variability across the available ASRs of the nested level for which the ASR represents the summary measure. The smaller the *r/R* (minimum 0%), the lower the variability across subarea ASRs.

Results

In Fig. 1, the ASRs for Nordic countries as a whole, for single countries and for country-specific regions, are shown with the corresponding 95% CI. The ASRs for countries appear to be scattered in the picture. In fact,

Fig. 1



European age-adjusted incidence rates for 'all sites excluding nonmelanoma skin cancer', for men, in 2014, for Nordic countries, single countries and regions. From Nordcan (*http://www-dep.iarc.fr/NORDCAN*).

the ASR for Nordic countries (453.1) is compatible with the Greenland's one only (384.1) because of the wide range of variability of the latter, because of the small number of cases on which it is based and the resulting imprecision in its computation (wide 95% CI). In contrast, Denmark (504.4) and Norway (509.9) showed greater ASRs that the Nordic countries one and Faroe Islands (251.0), Finland (404.5), Iceland (387.2) and Sweden (428.5) have lower values than the supranational summary ASR (Fig. 1).

In Table 1, for each area (Nordic countries, country and region), the number of cancer cases for 'all sites excluding nonmelanoma skin cancer', for men, in 2014 and the resident population are reported together with the ASR, the SE and the 95% CI.

The inverse relationship between number of observed cases and the SE is evident. In fact, SE is only 1.6 (cases per 100 000 men in 2014) for Nordic countries (on the basis of 79 441 analysed cases), whereas it is 40.6 for Greenland (99 cases).

In Table 1, the *r*/*R* is also reported for geographical level 1 (Nordic countries vs. countries) and 2 (single countries vs. regions).

When the ASR of Nordic countries is evaluated together with the *r*/*R*, the value of $r/R = 57.1\%$ provides a clear hint of a huge intercountries variability in ASRs, clearly shown in Fig. 1. In fact, this *r*/*R* means that the range between the lowest and the highest country-specific ASR is almost 60% of the Nordic country ASR.

Also within single countries, the overall ASR may not represent the regional ASRs and the amount of internal variability (Fig. 1) is well described by *r*/*R* (Table 1).

The smallest *r*/*R* value (9.6%) was observed in Iceland, where the small numbers of observed cases led to a non-negligible uncertainty in the regional estimates whose wide 95% CI overlapped the national one. A minor amount of variability ($r/R = 13.9\%$) was present in Norway, where the Northern region (ASR = 468.8) had a lower value and the Western region (539.8) had a higher ASR than the summary one. Almost the same *r*/*R* was present in Denmark (14.4%), where North Jutland (474.8) showed an ASR lower than the national value and Zealand (547.5) showed a greater one. Finland showed a greater inter-regional variability (31.9%), with Kuopio (345.5) and Turku (348.5) below and Helsinki (427.2) and Tampere (474.6) above the national mean. Finally, the slightly higher internal variability was found in Sweden (RR = 37.8%) where three regions, Northern

(364.5), Uppsala-Örebo (385.8) and Western (407.1), were below the national ASR and Stockholm–Gotland (526.3) higher than the country one.

Conclusion

This epidemiological exercise underlines that ASRs, which clearly provide the level of cancer incidence in a specific area and time for geographical and time comparisons, do not provide any information on possible internal variability. In fact, the SE, which usually accompanies ASR, refers only to the precision of the estimate and does not reflect the possible heterogeneity in cancer incidence in the area.

Therefore, the ASR of a CR, although correct from the computational point of view, and informative for geographical and time comparisons, could represent the incidence level only in some subareas or even in none.

If a CR also provides ASR for subareas, *r*/*R* is not necessary because the information on possible geographical heterogeneity is available. In contrast, if a CR only publishes a summary ASR, as happens for many CRs in Cancer incidence in five continents (Ferlay *et al.*, 2014), which is the most well-known and authoritative publication in the field, *r*/*R* is invaluable to have a clear impression of the variability behind the ASR.

When incidence data are available for different geographical layers, it is possible to add to the ASR a summary measure about the underlying variability. The Nordic countries dataset provided the invaluable chance of evaluating three subgeographical levels: supranational, national and regional.

We propose to compute the range between the highest and the lowest underlying ASRs to divide it by the ASR (*r*/*R*) and to express the result as a percentage.

The index *r*/*R* has been chosen among other more formal statistics (e.g. extreme quotient) (Gumbel and Keeney, 1950) because it only relies on ASRs and provides a direct measure of the effect of internal heterogeneity (range between maximum and minimum ASR in subareas) on the overall summary ASR.

In our example, on the basis of long-standing high-quality Nordic countries incidence data (Ferlay *et al.*, 2014), the *r*/*R* for the Nordic countries was quite high (57.1%), suggesting that the national ASRs could vary notably. In fact, the overall ASR for Nordic countries did not correspond with any of the national ASRs, out of Greenland's one (Fig. 1).

Also at a national level, when regional estimates are available, it is possible to add to the national ASR the *r*/*R* based on regional ASRs to express how well the national ASR represents the regional ones. In the dataset analysed, we showed that country ASR may reflect more (Iceland, Denmark and Norway) or less accurately (Finland and Sweden) the incidence of cancer in the different regions within a country.

The comparison between ASRs using the 95% CI overlap is simple and intuitive (Schenker and Gentleman, 2001) and showed major differences in ASRs between and within areas.

This study was based only on one incidence year. To check the reliability of *r*/*R*, we repeated the exercise also for the year 2012. The *r*/*R* in 2012 were similar to that in 2014 (data not shown) for almost all the countries, with the exception of Sweden, for which *r*/*R* showed in 2012 a smaller heterogeneity (*r*/*R* = 16.6%) than in 2014 (37.8%). The reason for this strong change was the change in the incidence ASR in the Stockholm–Gotland region from 2012 (420.5 cases/100 000) to 2014 (526.3). This change was the effect of a study on prostate cancer carried out in the county between 2012 and 2014 (Grönberg *et al.*, 2015). The ASR for all causes except skin and prostate cancer were 268.3 and 266.7, respectively. This example confirms that *r*/*R* reflects the true variability within an area.

Heterogeneity was identified among countries (areas between around 25 000 and 4 800 000 resident men) and among regions of several hundred thousand inhabitants, except for Iceland, where the population is smaller than in any of the other countries with regional information available.

It is possible to identify slight differences in cancer incidence between two geographical areas if the number of cases (population) is huge. Then, the ASRs are precise and the 95% CI is narrow. Thus, it is easier to detect a slight difference between two large (populated) regions than between two small ones. For example, between the ASRs of Kuopio and Oulu (highly populated), there is the same difference as that between the two Icelandic regions (poorly populated), but only the first two do not have overlapping CIs.

In general, the unavailability of a unique population-unit for subareas (countries, regions, provinces, counties, etc.) makes comparisons across areas difficult.

With the increase in the number of subareas, the variability among them is expected to increase and consequently the *r*/*R*. The aim of *r*/*R* is exactly to offer summary and straightforward information on possible outliers. In case *r*/*R* is small ($\sim < 10$–15%) it is immediately clear that all the ASR for each of the subareas are concentrated in a quite narrow range and if it is large ($>30\%$) it underlines that at least one of them is rather different from the overall ASR.

The *r*/*R* is a measure intended as a macro indicator of major heterogeneity among quite large subareas (e.g. regions in a country). For small areas and cluster analysis, other methods have to be chosen (Colonna and Sauleau, 2013).

CRs should start to provide also general information on internal cancer incidence geographical variability in

addition to standardised incidence rates. This would make the information more complete and clear for readers, avoiding misinterpretations. When incidence for subareas is available, *r*/*R*, which is very simple to compute, could be presented together with the general ASR as a first attempt to raise the issue.

The interpretation of incidence ASR requires the combined reading of ASR, SE and *r*/*R*: the ASR shows the level of incidence, the SE shows the precision of the ASR and *r*/*R* shows the amount of internal geographic variability. The *r*/*R* will be smaller if the ASR for subareas are quite similar to each other (more or less precisely estimated) or greater if they are rather different. This is the original and useful contribution provided by the *r*/*R*.

Acknowledgements

Conflicts of interest

There are no conflicts of interest.

References

Boyle P, Parkin DM (1991). Statistical methods for registries. . In: Jensen OM, Parkin DM, MacLennan R, Muir CS, Skeet, RG, editors. *Cancer registration: principles and principles* . Lyon, France: IARC Scientific Publication (No. 95). pp. 126–157.

Colonna M, Sauleau EA (2013). How to interpret and choose a Bayesian spatial model and a Poisson regression model in the context of describing small area cancer risks variations. *Rev Epidemiol Sante Publique* **61**:559–567.

Engholm G, Ferlay J, Christensen N, Kejs AMT, Hertzum-Larsen R, Johannesen TB, , *et al.* 2016 NORDCAN: cancer incidence, mortality, prevalence and survival in the nordic countries, version 7.3 (08.07.2016). Association of the Nordic Cancer Registries. Danish Cancer Society. Available at: *http://www.ancr.nu*. [Accessed 19 December 2016].

Ferlay J, Bray F, Steliarova-Foucher E, Forman D (2014). *Cancer incidence in five continents, CI5plus IARC CancerBase No 9*. Lyon, France: International Agency for Research on Cancer.

Grönberg H, Adolfsson J, Aly M, Nordström T, Wiklund P, Brandberg Y, *et al.* (2015). Prostate cancer screening in men aged 50–69 years (STHLM3): a prospective population-based diagnostic study. *Lancet Oncol* **16**: 1667–1676.

Gumbel EJ, Keeney RD (1950). The extremal quotient. *Ann Math Stat* **21**: 523–538.

Schenker N, Gentleman JF (2001). On judging the significance of differences by examining the overlap between confidence intervals. *Am Stat* **55**:182–186.

The thickness of melanomas has decreased in central Italy, but only for thin melanomas, while thick melanomas are as thick as in the past

Emanuele Crocetti[a], Adele Caldarella[a], Alessandra Chiarugi[b], Paolo Nardini[b] and Marco Zappa[a]

The objective of this study was to evaluate the time trend of melanoma thickness in a population-based case series. All invasive ($n=2862$) and in-situ ($n=605$) cutaneous melanoma incident cases diagnosed in 1985–2004 were retrieved from the Tuscany Cancer Registry, central Italy. Standardized (European population) incidence rates were computed for four periods: 1985–1989, 1990–1994, 1995–1999, 2000–2004, and for Breslow thickness classes (≤ 1, 1.01–2.00, >2 mm). The annual percent change (APC) of the standardized rates was computed. Thickness was evaluated on the basis of sex, age, morphology type, site and period of time. Median thickness was evaluated by means of a nonparametric K-sample test. The incidence rate of melanoma rose significantly for both invasive (APC = +5.1%) and in-situ lesions (APC = +11.1). The sex distribution of patients with invasive melanoma did not change over time (mean male/female ratio 0.95). The mean age at diagnosis did not change (57.2 years; SD = 17.2 years). From 1985–1989 to 2000–2004 the median value of thickness decreased from 1.68 to 0.8 mm ($P<0.001$). Within the Breslow categories the median value of thickness decreased significantly for thin melanomas (≤ 1 mm) but not for intermediate (1.01–2.00) or for thick melanomas (>2 mm). Among the most common melanoma types, the median thickness decreased for superficial spreading melanomas but not for nodular melanomas. Over time, the incidence of melanoma has increased notably and the median thickness has decreased. However, median thickness has decreased only among thin melanomas, whereas it has not changed for thick melanomas, most of which are of the nodular type. *Melanoma Res* 20:422–426 © 2010 Wolters Kluwer Health | Lippincott Williams & Wilkins.

Melanoma Research 2010, **20**:422–426

Keywords: cancer registry, epidemiology, incidence, melanoma, population-based, thickness, time trend

[a]Clinical and Descriptive Epidemiology Unit and [b]Melanoma Early Diagnosis Service, Institute for Cancer Study and Prevention ISPO, Florence, Italy

Correspondence to Emanuele Crocetti, Via di San Salvi 12, 50135, Florence, Italy
Tel: +39 55 6268320; fax: +39 55 679954;
e-mail: e.crocetti@ispo.toscana.it

Received 11 January 2010 **Accepted** 27 June 2010

Introduction

The incidence of melanoma has increased notably over the last decades in all western countries [1,2] including Italy [3]. A major cause for this increase has been early diagnosis. In addition, most of the diagnosed lesions are now in-situ [4] or thin melanomas [5–7], especially of the superficial spreading melanoma type (SSM) [1,8]; therefore, their prognosis is very good.

In contrast, the incidence of thick melanomas has not decreased [4,7,9–11] nor increased [1].

Thickness is the most relevant prognostic factor. Moreover, a large proportion of thick melanomas is of the nodular type [11–13]. Against the background of stable or even rising rates for thick melanomas, melanoma mortality has not significantly decreased [1].

In recent years, health-care professionals have focused on how to cope with thick deadly melanomas, but the problem is still unsolved [14].

The aim of this paper is to evaluate the incidence trend of cutaneous melanomas with special reference to thickness in an Italian population-based case series.

Materials and methods

We retrieved all the invasive ($n = 2862$) and in-situ ($n = 605$) cutaneous melanoma incident cases diagnosed in 1985–2004 from the archives of the Tuscany Cancer Registry (RTT). RTT is a population-based cancer registry active in the provinces of Florence and Prato (approximately 1 161 000 residents in the 2001 census), central Italy [15].

We arranged Breslow thickness in three classes, thin (≤ 1 mm), intermediate (1.01–2.00), and thick (>2 mm), as indicated by Balch *et al.* [16]. As for thick tumours, we also analysed thick (≥ 3 or ≥ 4 mm) lesions.

We used the median to measure thickness because of the skewness of its distribution.

Thickness categories were evaluated according to the following variables:

(1) Sex.
(2) Age (0–49, ≥ 50).
(3) Morphological type (the International Classification of Diseases for Oncology morphology codes):

0960-8931 © 2010 Wolters Kluwer Health | Lippincott Williams & Wilkins DOI: 10.1097/CMR.0b013e32833d9d36

SSM(ICDO-M = 8743), nodular melanoma (NM; ICDO-M = 8741), lentigo maligna melanoma (LMM, ICDO-M = 8742), not otherwise specified (NOS, ICDO-M = 8720), 'other types' (ICDO-M = 8722, 8730, 8740, 8744, 8770, 8771, 8772).

(4) Site: head and neck, trunk, upper limb, lower limb, others and unspecified.

(5) Period of time: 1985–1989, 1990–1994, 1995–1999, 2000–2004.

Statistical methods

Incidence rates were age-standardized through the direct method using the European standard population.

We computed the annual percent change (APC) of the standardized rates using the weighted least squares method, on the basis of single incidence year. We performed the above-mentioned computations by means of the SEER*Stat 6.3.6 software (*www.seer.cancer.gov/seerstat*). We compared proportions by means of the χ^2 test, or the Fisher's exact test, when we expected less than five observations.

As regards the median, we used the Stata command 'median' that performs a nonparametric K-sample test that evaluates the null hypothesis that those samples were drawn from populations with the same median (*www.stata.com*). In the case of two samples, the χ^2 statistic test is calculated with and without a continuity correction.

Results

From 1985 to 2004, 3467 patients residing in the RTT area had a diagnosis of malignant melanoma (2862 invasive and 605 *in situ*). The standardized incidence rate of invasive malignant melanoma rose from 6.4 per 100 000 in 1985–1989 to 13.6 in 2000–2004, with a mean annual pace of + 5.0% [95% confidence intervals (CIs), + 4.0/ + 8.2], Table 1. The growth of incidence was statistically significant for both men (APC = + 5.3, 95% CI, + 4.2/ + 6.5) and women (APC = + 4.9, 95% CI, + 3.5/ + 6.3).

In the analysed period, the sex ratio was stable, with a slight predominance of women. The median age at diagnosis did not change over time, being 58.1 years (range 21.2–100.6 years).

With regard to invasive melanomas, a statistically significant growing incidence was detected for thin (APC = + 9.5; 95% CI, + 7.1/ + 11.9), for intermediate (APC = + 3.1; 95% CI, + 1.1/ + 5.2) and for thick melanomas (APC = + 2.1; 95% CI, + 0.4/ + 3.7), and it was almost statistically significant for those melanomas without the information on Breslow thickness (APC = + 1.9; 95% CI, − 0.3/ + 4.1; Table 1.

Six hundred and five in-situ melanomas were diagnosed with a statistically significant growing trend (APC = + 11.1; 95% CI, + 8.1/ + 14.3); standardized incidence rates for in-situ melanomas rose from 0.8 per 100 000 in 1985–1989 to 4.0 in 2000–2004. The incidence trend for in-situ melanomas was statistically significant for both men (APC =+ 12.7; 95% CI, + 8.7/ + 16.9) and women (APC= + 9.4; 95% CI, + 6.1/ + 12.8). In-situ invasive melanomas did not show any statistically significant change, either in sex ratio (percentage of women 54.1) or in the median age at diagnosis (57.8 years).

The mean age at diagnosis for in-situ melanomas was lower than for invasive melanomas (55.5 vs. 57.2 years; $P = 0.02$). However, among invasive melanomas (≤ 1 mm) the mean age at diagnosis (52.7 years) was lower than for in-situ

Table 1 Tuscany Cancer Registry: invasive melanoma, absolute numbers, proportion of females, standardized (European population) incidence rates, annual percent change of standardised rates

	1985–1989	1990–1994	1995–1999	2000–2004	n/p
Invasive melanoma, *n*	442	565	835	1020	2862
% females	54.5	54.0	55.3	50.2	*P*=0.13
Incidence rate	6.4	8.0	11.4	13.6	*P*<0.01 for APC
Breslow thickness					
>0 to ≤ 1 mm, *n* (%)	92 (20.8)	182 (32.2)	375 (44.9)	476 (46.7)	1125
Incidence rate	1.4	2.7	5.5	6.7	*P*<0.001 for APC
1.01–2.00 mm, *n* (%)	76 (17.2)	93 (16.5)	112 (13.4)	129 (12.8)	410
Incidence rate	1.1	1.3	1.5	1.9	*P*=0.006 for APC
>2 mm, *n* (%)	117 (26.5)	134 (23.7)	153 (18.3)	200 (19.6)	604
Incidence rate	1.6	1.8	1.8	2.2	*P*=0.016 for APC
Unknown, *n* (%)	157 (35.5)	156 (27.6)	194 (23.4)	215 (21.0)	723
Incidence rate	2.2	2.2	2.5	2.8	*P*=0.07 for APC
In-situ melanoma, *n*	55	72	182	296	605
% females	61.8	63.9	52.2	51.4	*P*=0.15
Incidence rate	0.8	1.0	2.7	4.0	*P*<0.01 for APC
Morphology type					
SSM	2.9	4.5	7.3	9.0	<0.001 for APC
NM	0.8	0.8	0.8	0.8	0.96 for APC
LM	0.2	0.2	0.2	0.2	0.14 for APC
Other	0.3	0.3	0.4	0.4	0.15 for APC
N.o.s.	2.2	2.2	2.7	3.2	0.006 for APC

Annual percent change of standardized rates (APC) are computed on single years of diagnosis. Probability (*P*) for APC to be equal to 0 or for the probability that each sample has the same proportion of observations.
LM, lentigo melanoma; N.o.s., not otherwise specified; NM, nodular melanoma; SSM, superficial spreading melanoma.

melanomas ($P = 0.001$), whereas the mean age of 60.2 years for melanomas (> 1 mm) was higher than for both in-situ melanomas ($P < 0.001$) and those with a thickness of 1 mm or less ($P = 0.001$).

According to types of morphology, the increasing incidence trend was statistically significant for SSMs ($n =$ 1655, APC = + 7.0; 95% CI, + 5.2/ + 8.8) and for the group 'other types' ($n = 155$, APC = + 6.7%; 95% CI, + 2.6/ + 11.0), whereas for NMs ($n = 238$), LMM ($n = 79$) and for NOS melanomas ($n = 775$), the trends did not reach any statistical significance.

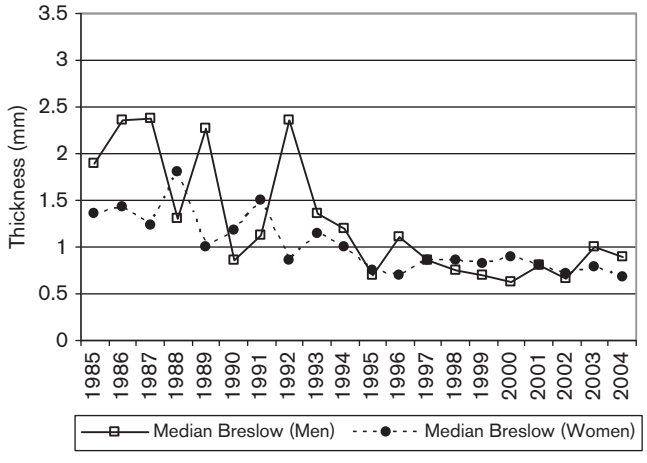
From 1985–1989 to 2000–2004 the median value of thickness, for invasive melanomas, decreased from 1.68 to 0.8 mm ($P < 0.001$). The median thickness decreased statistically significantly for both men (from 2.1 to 0.8 mm) and women (from 1.3 to 0.75 mm) Fig. 1.

The decrease over time in median thickness was not present in all Breslow categories. In fact, the median value of thickness decreased statistically significantly for thin melanomas (≤ 1 mm) but not for intermediate (1.01–2.00 mm) or thick melanomas (> 2 mm) Table 2.

Median thickness did not decrease among thick melanomas in other Breslow categories either (≥ 3 mm, ≥ 4 mm; data not shown).

The decrease in median thickness was statistically significant only for SSM (from 1.20 to 0.68 mm) and for other types (from 2.75 to 1.10), whereas it did not reach statistical significance for the other morphological types of melanoma (Table 2). In the most recent period (2000–2004), SSM represented the largest proportion of melanomas (62.6%; 639/1020), followed by melanomas NOS (17.5%; $n = 179$), the group 'other' (10.7%; $n = 109$), nodular melanomas, (6.2%; $n = 62$) and LMMs (2.9%; $n = 30$). Among the 208 thick melanomas diagnosed in 2000–2004, 45.7% were SSM and 24.5% of nodular type.

Fig. 1



Tuscany Cancer Registry. Invasive melanoma: median Breslow thickness for men and women by calendar years.

Table 2 Tuscany Cancer Registry

	1985–1989	1990–1994	1995–1999	2000–2004	*P*
Overall	1.68	1.2	0.8	0.8	<0.001
Males	2.1	1.3	0.8	0.8	<0.001
Females	1.3	1.15	0.8	0.75	<0.001
Breslow's thickness					
≤ 1 mm	0.66	0.57	0.55	0.50	0.001
1.01–2.00 mm	1.50	1.45	1.40	1.30	0.239
>2 mm	3.4	3.5	3.7	3.8	0.418
Morphology type					
SSM	1.20	0.85	0.70	0.68	<0.001
NM	3.33	3.39	3.30	4.00	0.517
LM	1.80	0.81	0.67	0.79	0.398
Other	2.75	3.20	1.10	1.10	0.003
N.o.s.	2.25	2.35	1.79	1.80	0.675

Invasive melanoma: median thickness by period of diagnosis for males and females, for Breslow's thickness categories, for morphology type (LM, lentigo melanoma; N.o.s., not otherwise specified; NM, nodular melanoma; SSM, superficial spreading melanoma) and for site. *P* shows the probability that the medians in different groups are medians of samples drawn from the same population.

With regard to age, the increase in incidence was present and statistically significant for both younger (0–49 years) men (APC = + 5.0, 95% CI, + 2.7; + 7.2) and women (APC = + 6.2, 95% CI, + 4.0; + 8.5) and for older (50 + years) men (APC = + 5.6, 95% CI, + 4.1; + 7.0) and women (APC = + 3.5, 95% CI, + 2.2; + 4.8).

With regard to skin sub-sites, by age group and sex (Table 3), melanomas of the head and neck did not show any statistically significant change either in incidence or in Breslow thickness in both sexes and age groups. The incidence of melanoma of the trunk increased significantly in both sexes and age groups, and the median thickness showed a statistically significant change towards a decrease, with the exception of younger women. The incidence of melanomas of the upper limbs increased in all age groups and both sexes, whereas their median thickness decreased statistically significantly only among younger subjects. With regard to melanomas of the lower limbs, there was a significant increase in incidence and a statistically significant decrease in median thickness among women. The group of melanomas of NOS sites increased over time, whereas their median thickness was stable.

Discussion

The increasing incidence of invasive melanomas observed from 1985 to 2004 in central Italy was mainly supported by increasingly thinner lesions, especially of the SSM type. Therefore, the overall median thickness of melanomas has decreased over time, being in recent years 0.8 mm. This result was supported by the decrease in the median thickness of thinner melanomas (≤ 1 mm; 0.5 mm during 2000–2004). However, the shift in thickness observed for thin melanomas does not affect thick melanomas, which in Italy are now as thick as they

Table 3 Tuscany Cancer Registry

	Males			Females		
	≤ 49 years	50+ years	All ages	≤ 49 years	50+ years	All ages
Head and neck						
Number	32	132	164	28	125	153
APC 1985–2004	−2.3	+1.6	+0.4	−1.9	+0.3	−0.3
95% CI	(−5.8; +1.4)	(−1.6; +4.9)	(−2.6; +3.5)	(−5.3; +1.7)	(−3.5; +4.3)	(−3.8; +3.4)
Thickness 2000–2004	0.55	2.7	2.3	0.7	1.19	0.9
P thickness change 1985–2004	0.10	0.64	0.41	0.54	0.20	0.17
Trunk						
Number	179	472	651	200	184	384
APC 1985–2004	+4.7	+4.8	+4.8	+6.9	+4.5	+6.1
95% CI	(+1.3; +8.2)	(+2.7; +7.1)	(+2.9; +6.7)	(+3.1; +10.8)	(+1.1; +8.1)	(+2.8; +9.5)
Thickness 2000–2004	0.86	0.7	0.8	0.7	0.7	0.7
P thickness change 1985–2004	0.005	<0.001	<0.001	0.69	0.01	0.02
Upper limb						
Number	54	102	156	79	128	207
APC 1985–2004	+4.5	+7.5	+7.6	+5.7	+5.1	+5.3
95% CI	(+0.9; +8.3)	(+4.0; +11.2)	(+4.7; +10.5)	(+1.2; +10.4)	(+1.8; +8.7)	(+2.7; +8.0
Thickness 2000–2004	0.64	0.85	0.65	0.6	0.9	0.68
P thickness change 1985–2004	0.054	0.20	0.03	0.07	0.60	0.12
Lower limb						
Number	71	132	203	204	423	627
APC 1985–2004	+8.3	+4.4	+5.3	+4.3	+1.8	+3.0
95% CI	(+3.8; +13.0)	(+0.6; +8.3)	(+2.5; +8.1)	(+1.2; +7.5)	(0; +3.6)	(+1.1; +4.9)
Thickness 2000–2004	0.8	1.8	1.1	0.71	1.3	1.1
P thickness change 1985–2004	0.10	0.40	0.09	0.06	0.04	0.08
N.o.s						
Number	58	110	168	56	93	149
APC 1985–2004	+7.4	+8.9	+7.9	+12.2	+11.1	+14.2
95% CI	(+2.0; +13.0)	(+3.7; +14.3)	(+3.9; +12.1)	(+3.8; +21.6)	(+5.4; +17.2)	(+7.6; +21.1)
Thickness 2000–2004	0.53	0.5	0.5	0.52	0.72	0.52
P thickness change 1985–2004	0.51	0.23	0.16	0.47	0.68	0.56

Melanoma, number of cases according to sex, age (0–49, 50+years and all ages), annual percent change (APC) of standardized incidence rates 1985–2004 with corresponding 95% Confidence Intervals (95% CI), median thickness during 2000–2004, probability (*P*) that the median thicknesses for the periods 1985–1989, 1990–1994, 1995–1999 and 2000–2004 belong to populations with the same medians according to a non parametric K sample test.
N.o.s., not otherwise specified.

were in the past. In fact, the median thickness of intermediate (1.01–2.00 mm) and thick melanomas (> 2 mm) did not decrease over the analysed period.

Moreover, although the proportion of intermediate and thick melanomas has reduced over time, their incidence rates show a statistically significant increasing trend [1]. In addition, their absolute number has increased [7,14].

In agreement with the literature, we observed a strong increasing trend for in-situ melanomas in central Italy [4], and almost half of the invasive melanomas are now thin. Such lesions have a very good prognosis, in so much as the UK recently proposed a change in guideline recommendations for those less than 0.5 thick, suggesting less frequent follow-up [17]. The mean age at diagnosis was lower for ≤ 1 mm melanomas than from in-situ melanomas. This would indicate that in-situ melanomas are not a precursor lesion of melanoma, but have a different pathway than invasive melanomas.

As observed in other reports, we did not detect any decrease in the thickness of NMs at diagnosis [7]. We documented that in this population-based series a significant percentage (24.5%) of thick melanomas are NMs, reaching around 30% when melanomas NOS are excluded [16].

There was an increase in incidence for all skin sites, with the exception of head and neck, for which the rates were stable over time. The trends for skin sites were similar for younger (0–49 years) and older (50 + years) patients of both sexes. A statistically significant change (decrease) in median thickness was present for melanomas of the trunk in men and older women, for upper limbs only in younger patients, and for lower limbs only in women.

The epidemiology of melanoma is further divided into different groups of lesions [18]. A first group includes those melanomas easily detectable by enhanced early diagnosis, increasingly thinner, mainly of a SSM type, with a very good prognosis and presumably with some amount of over-diagnosis [19]. There is a second group of melanomas, more aggressive, thick at diagnosis, and often of a nodular type [20]. The relationship between these two types of lesions is still unknown. Therefore, it is crucial to investigate further the biological history of thick melanomas.

Acknowledgement

The authors acknowledge Ente Cassa di Risparmio di Firenze.

References

1 Linos E, Swetter SM, Cockburn MG, Colditz GA, Clarke CA. Increasing burden of melanoma in the United States. *J Invest Dermatol* 2009; **129**:1666–1674.
2 De Vries E, Coebergh JM. Cutaneous malignant melanoma incidence in Europe. *Eur J Cancer* 2004; **40**:2355–2366.
3 Crocetti E, Capocaccia R, Casella C, Guzzinati S, Ferretti S, Rosso S, *et al.*; Network of the Italian Cancer Registries (AIRT). Population-based incidence and mortality cancer trends (1986–1997) from the network of Italian cancer registries. *Eur J Cancer Prev* 2004; **13**:287–295.
4 Criscione VD, Weinstock MA. Melanoma thickness trends in the United Stases, 1988–2006. *J Invest Dermatol* 2010; **130**:793–799.
5 Garbe C, Leiter U. Melanoma epidemiology and trends. *Clin Dermatol* 2009; **27**:3–9.
6 Baumert J, Schmidt M, Giehl KA, Volkenandt M, Plewig G, Wendtner C, Schmid-Wedntner MH. Time trends in tumour thickness vary in subgroups: analysis of 6475 patients by age, tumour site and melanoma type. *Melanoma Res* 2009; **19**:24–30.
7 Lipsker DM, Hedelin G, Heid E, Grosshans EM, Cribier BJ. Striking increase of thin melanomas contrasts with stable incidence of thick melanoma. *Arch Dermatol* 1999; **135**:1451–1456.
8 Warycha MA, Christos PJ, Mazumdar M, Darvishian F, Shapiro RL, Berman RS, *et al.* Changes in the presentation of nodular and superficial spreading melanomas over 35 years. *Cancer* 2008; **113**:3341–3348.
9 Richardson A, Flechter L, Sneyd M, Cox B, Reeder AI. The incidence and thickness of cutaneous malignant melanoma in New Zealand 1994–2004. *N Z Med J* 2008; **121**:18–26.
10 Crocetti E, Carli P. Unexpected reduction of mortality rates from melanoma in males living in central Italy. *Eur J Cancer* 2003; **39**:818–821.
11 Tejera-Vaquerizo A, Mendiola-Fernandez M, Fernandez-Orland A, Herrera-Ceballos E. Thick melanoma: the problem continues. *J Eur Acad Dermatol Venerol* 2008; **22**:575–579.
12 Geller AC, Elwood M, Swetter SM, Brooks DR, Aitken J, Youl PH, *et al.* Factors related to the presentation of thin and thick melanoma from a population-based cancer registry in Queensland Australia. *Cancer* 2009; **115**:1318–1327.
13 Crocetti E, Mangone L, Lo Scocco G, Carli P. Prognostic variables and prognostic groups for malignant melanoma** The information from Cox and classification and regression trees analysis: an Italian population-based study. *Melanoma Res* 2006; **16**:429–433.
14 Murray CS, Stockton DL, Doherty VR. Thick melanoma: the challenge persists. *Br J Dermatol* 2005; **152**:104–109.
15 Paci E, Crocetti E, Benvenuti A, Buzzoni C, Caldarella A, Giovannetti L, *et al.* Cancer incidence in florence & prato (1998–2002). In: Curado MP, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M, Boyle P, editors. *Cancer incidence in five continents,* Vol. IX IARC Scientific Publications No. 160. Lyon: IARC; 2007.
16 Balch CM, Gershenwald JE, Soong S, Thompson JF, Atkins MB, Byrd DR, *et al.* Final version of 2009 AJCC melanoma staging and classification. *JCO* 2009; **36**:6199–6209.
17 Einwachter-Thompson J, MacKie RM. An evidence base for reconsidering current follow-up guidelines for patients with cutaneous melanoma less than 0.5 mm thick at diagnosis. *Br J Dermatol* 2008; **159**:337–341.
18 Ipsker D, Engel F, Cribel B, Velten M, Hedelin G. Trends in melanoma epidemiology suggest three types of melanoma. *Br J Dermatol* 2007; **157**:338–343.
19 Welch HG, Woloshin S, Schwartz LM. Skin biopsy rates and incidence of melanoma: population based ecological study. *BMJ* 2005; **331**:481.
20 Betti R, Martino P, Vergani R, Gualandri L, Crosti C. Nodular melanomas: analysis of the casistic and relationship with thick melanomas and diagnostic delay. *J Dermatol* 2009; **35**:643–650.

The thickness of melanomas has decreased in central Italy, but only for thin melanomas, while thick melanomas are as thick as in the past

Author: Emanuele Crocetti, Adele Caldarella, Alessandra Chiarugi, et al
Publication: Melanoma Research
Publisher: Wolters Kluwer Health, Inc.
Date: Oct 1, 2010

License Not Required

Wolters Kluwer policy permits only the final peer-reviewed manuscript of the article to be reused in a thesis. You are free to use the final peer-reviewed manuscript in your print thesis at this time, and in your electronic thesis 12 months after the article's publication date. The manuscript may only appear in your electronic thesis if it will be password protected. Please see our Author Guidelines here: https://cdn-tp2.mozu.com/16833-m1/cms/files/Author-Document.pdf?_mzts=636410951730000000.

BACK | CLOSE WINDOW

Melanoma incidence in central Italy will go on increasing also in the near future: A registry-based, age–period–cohort analysis

Emanuele Crocetti[a], Paolo Carli[b] and Guido Miccinesi[a]

The aim of the study was to evaluate malignant melanoma incident trends in central Italy by means of an age–period–cohort approach. A total of 1977 malignant melanoma (15–84 years) incidents in the area of the Tuscany Cancer Registry between 1987 and 2001 were analysed. Poisson regression has been used to estimate age, cohort and period effect. A nonlinear regression model was used to estimate the expected number of new cases in the period 2002–2006. Incidence rates increased in all age, period and cohort groups. The model that best fitted the data included age and 'drift'. The linear effect ('drift') showed, in each age group, an increase of the risk of malignant melanoma diagnosis of about 36.6% every 5 years of period or cohort. For the period 2002–2006, 1112 new cases were predicted with a standardized rate (age 15–84 years) of 19.2 × 100.000. In the Tuscany Cancer Registry area, no clues for malignant melanoma incidence rates levelling off were documented. Growing rates and number of malignant melanoma are expected in the near future. ***European Journal of Cancer Prevention*** **16:50–54 © 2007 Lippincott Williams & Wilkins.**

European Journal of Cancer Prevention 2007, **16**:50–54

Keywords: age–period–cohort, epidemiology, estimate, incidence, malignant melanoma, population-based, trend

[a]Clinical and Descriptive Epidemiology Unit, CSPO and [b]Department of Dermatology, University of Florence, Florence, Italy

Correspondence to Dr Emanuele Crocetti, MD, UO Epidemiologia Clinica e Descrittiva-CSPO, Via di San Salvi 12, 50135 Florence, Italy
Tel: +39 0556268320; fax: +39 055679954; e-mail: e.crocetti@cspo.it

Received 8 November 2005 **Accepted** 1 December 2005

Introduction

Malignant melanoma (MM) incidence has increased over the last decades in almost all Western countries (Lens and Dawes, 2004). In central Italy also incidence trends showed a significant increasing trend in both men and women (Crocetti and Carli, 2003). More recently, levelling of or even decreasing trends in incidence have been identified among young participants in Denmark, in the Netherlands, in Switzerland and in the United Kingdom (De Vries *et al.*, 2003). A decrease in incidence among young cohorts indicates that there will be a future overall decline in rates as soon as these participants contribute to older age groups.

The aim of the present paper was to explore incidence trends for cutaneous MM in central Italy focusing on the age, period and cohort effect.

Materials and methods

Incidence data were retrieved from the Tuscany Cancer Registry (RTT), a population-based cancer registry active in the provinces of Florence and Prato (about 1 160 000 residents in the 2001 census), central Italy, since 1985. The description of the criteria for collection, and registration followed by the Registry, has been presented elsewhere (Paci *et al.*, 2002).

During 1987–2001, 2071 incident cutaneous melanomas were registered in the RTT; in the present analysis, we selected age range 15–84 years and 1977 incident cases were included.

Number of cases and person-years were aggregated in 5-year age groups (from 15–19 to 80–84 years), 5-year periods (1987–1991, 1992–1996, 1997–2001). Five-year cohorts of birth were computed according to age and period. In the figure, cohorts are labelled according to the central year of the cohort of birth.

The incidence rates were standardized using direct method with the European standard population as the reference.

Poisson regression has been used to estimate age, cohort and period effect. We may consider cohort effects as influences that affect rates in a specified generation or birth cohort throughout life, whereas period effects affect rates equally across all age groups at a specified period (Clayton and Schifflers, 1987a). When there is a regular temporal trend, which cannot be ascribed to either period or cohort influences, it is called 'drift'. Only when we observe irregular or sudden changes, we can ascribe the observed temporal trend to either period or cohort influences. The goodness-of-fit of different models was assessed by the deviance. The closer the deviance with the degree of freedom, the better the fit of the model. Differences between deviances allowed us to compare

0959-8278

nested models (e.g. age alone vs. age + cohort) (Clayton and Schifflers, 1987b).

The expected number of new cases for the years 2002–2006 was estimated according to a nonlinear regression model, proposed by Dyba *et al.* (1997), including the effect of age and age-specific temporal trends. Friendly STATA macros for the application of such method for short-term prediction are available on the Web site of the European Network of Cancer Registries (http://www.encr.com.fr/stata-macros.htm). This model assumes that the absolute change in incidence rate over time for a given age group is often larger when the baseline rate is larger. The age-specific absolute change of incidence is proportional to the corresponding age-specific baseline rate, whereas for a given time period the relative change in incidence is the same for all age groups. Owing to these characteristics this model preserves in the period of prediction the age pattern of incidence existing in the data, and the age-specific predicted rates are substantially more precise than those for a linear model (Dyba *et al.*, 1997).

Several methods have been used for making predictions of the future cancer burden; the one used in the present analysis showed good performance when compared with other methods (Moller *et al.*, 2003).

We used the annual sex-specific and 5-year age-specific population for the years 2002–2006 based on the official data and future estimates of the Regional Office of the National Institute of Statistics (years 2002–2003, 2008) and the estimates based on the Waring method for 2004–2006 (Shryock *et al.*, 1976).

Total number of expected cases and prediction intervals are presented. Predictions of incidence for each age group are also presented.

Results

4Table 1 shows age-specific and age-standardized melanoma incidence rates from the Tuscany Cancer Registry area for three periods (1987–1991, 1992–1996 and 1997–2001).

Incidence rates increased in all age groups. The age-specific incident rates increased in participants born in more recent years (Fig. 1), and they also increased from the first to the third analysed period (1987–1991, 1992–1996, 1997–2001) (Fig. 2).

The model approach used to disentangle age, period and cohort effect showed that the model that best fitted the data was the one including age and 'drift' (deviance of the 'age + cohort' model – deviance of the 'age + drift' model = 15.7; number of parameters of the 'age + cohort' model – number of parameters of the 'age-drift' model = 14; *P*-value of 15.7 on a χ^2distribution with 14 degrees of freedom = 0.392; for the age + period model vs. age + drift model; *P*-value = 0.808).

Table 1 **Malignant cutaneous melanoma**

	Period		
	1987–1991	1992–1996	1997–2001
Age (years)			
15–19	0.8	1	2.4
20–24	2	5.7	6.1
25–29	3.9	7	10.2
30–34	5.7	7.6	13
35–39	5.1	8.4	12.9
40–44	8.9	11.1	15.2
45–49	9.5	13.1	18.4
50–54	10.3	16.3	23.1
55–59	14.5	17.9	26
60–64	14.8	19.1	19.4
65–69	16.2	18	23.5
70–74	16.2	19.4	25.1
75–79	15	25.9	31.1
80–84	14.2	19.7	28.9
No. of cases	449	636	892
Standardized Rates[a]	8.3	11.6	15.9

Number of incident cases, age-specific (males and females) and age-standardized (European population) incidence rates in the Tuscany Cancer Registry, according to time period.
[a]Standardized rates are computed in the population 15–84 years.

The linear temporal effect ('drift') showed that in each age-group, there was an increase of the risk of MM diagnosis of about 36.6% every 5 years of period or cohort (mean annual increase 6.4%). The model including age and drift resulted the best also for explaining incidence data for men and women, when analysed separately (data not shown).

The predicted standardized incidence rate (age 15–84 years) for the period 2002–2006 was 19.2; the predicted number of new cases was 1090 (95% prediction intervals 984–1196).

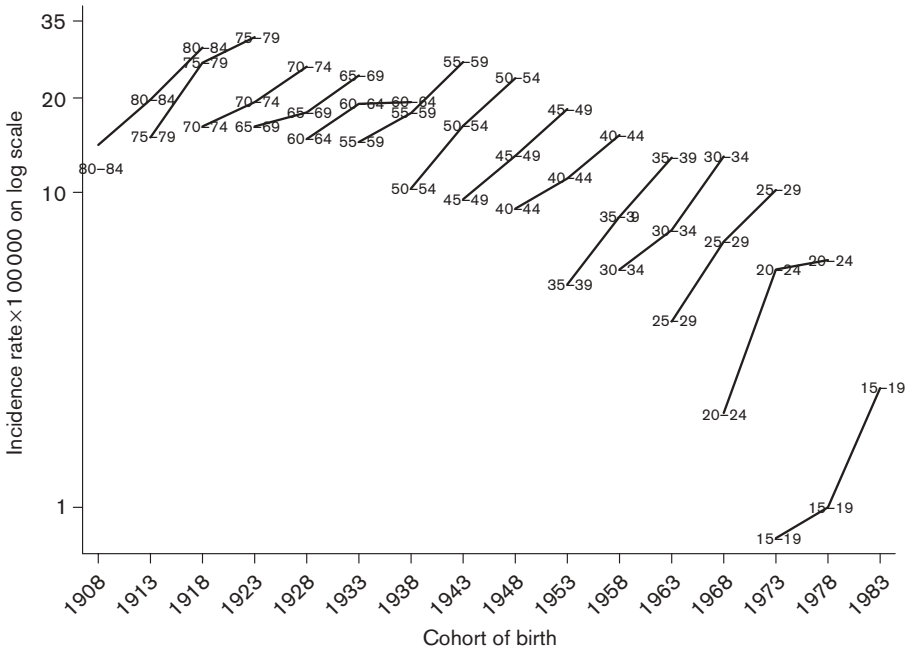
Figure 3 reports age-specific incidence predictions for 2002–2006, together with their prediction intervals.

Discussion

The results of the present study indicate that MM incidence rates in the population covered by the Tuscany Cancer Registry, central Italy, are still increasing. This trend is present in all age groups, and also younger participants, both men and women, do not show any sign of levelling of.

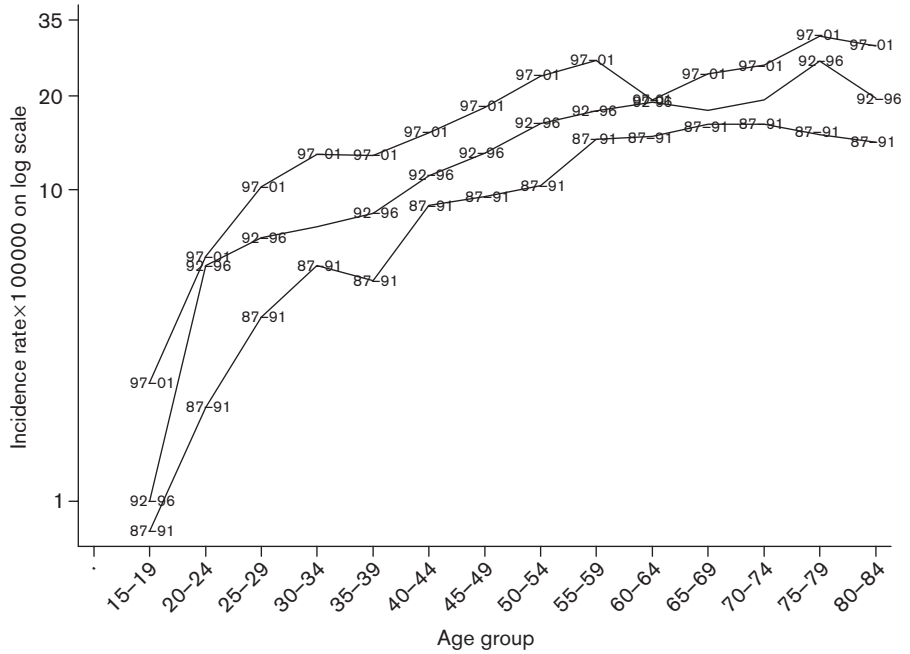
The present figure contrasts with that recently found in some northern European populations. According to data extracted from the EUROCIM database 165 cancer registries, a deceleration in incidence trends occurred

Fig. 1



Tuscany Cancer Registry, malignant melanoma 1987–2001. Age-specific incident rates according to the cohort of birth, men + women.

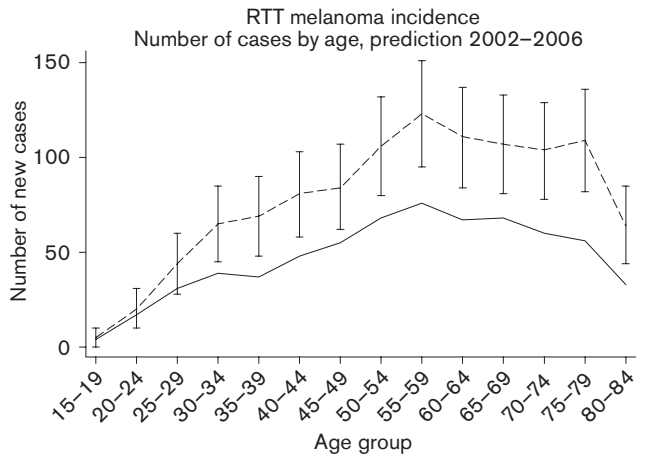
Fig. 2



Tuscany Cancer Registry, malignant melanoma 1987–2001. Age-specific incident rates according to three different periods (1987–1991, 1992–1996, 1997–2001), men + women.

recently in Northern European countries among persons younger than 70 years; whereas in eastern and southern Europe incidence rates were still increasing (De Vries *et al.*, 2003). Therefore, we are probably observing a shift between northern and southern European populations in terms of future scenarios about

Fig. 3



Melanoma of both sexes, Tuscany Cancer Registry area. Five-year mean number of observed incident cases in 1987–2001 (solid line) and incident cases predicted for 2002–2006 (dashed line) with approximate 95% prediction intervals; by age.

melanoma 'epidemic'. Earlier detection and a growing public awareness about risk associated with excessive sun exposure are the most plausible explanations for the deceleration found in northern Europe (De Vries *et al.*, 2003).

Little is, however, known on the factors still holding up the increasing incidence trends in Mediterranean people.

In this study, models including the effects of age, period and birth cohort were used to adequately analyse the rising trends.

As known, cohort effect reflects exposures that affect rates in a specified cohort equally throughout life, whereas period effect affects rates equally across all age groups at a specific period (Clayton and Schifflers, 1987a). An age effect (increasing risk according to ageing) and a linear effect are causing the increasing incidence.

Owing to the linear dependence between the linear part of cohort and period (and age) however, the interpretation of regular incidence trend is not possible. Indeed, only when we observe irregular changes we must consider age-period or age-cohort models. On the contrary, this so-called 'drift' effect represents a situation equally well described by two models, age and period (linear) or age and cohort (linear) and cannot be ascribed to either period or cohort effect (Clayton and Schifflers, 1987a).

Although statistics does not allow one to disentangle period from cohort effect, some suggestions may come from the epidemiological knowledge.

The major environmental risk factor for MM is the intermittent exposure to ultraviolet radiation (UV) (Elwood and Jopson, 1997), especially during childhood (Naldi *et al.*, 2000). Such exposure has become popular after the Second World War, and in Italy, particularly from the second-half of the 50s to the early 60s as a consequence of a rapid economic growth.

The present data are, however, not fully explained by the hypothesis of increasing risk for UV exposure during childhood only. In fact, the increase in incidence was evidenced in all birth cohorts, also in the oldest ones, in participants who during the 1950s–1960s were middle-aged. Our results are consistent with the hypothesis that UV exposure increases the MM risk independently from the age of exposure (Elwood and Jopson, 1997). Recent data suggest that cutaneous melanomas may arise through two pathways, one associated with melanocyte proliferation and the other with chronic exposure to sunlight (Whiteman *et al.*, 2003). Australian patients with head and neck melanomas – lentigo maligna melanoma excluded – compared with patients with melanomas of the trunk, were statistically significantly less likely to have more than 60 nevi [odds ratio (OR) = 0.34, 95% confidence interval (CI) = 0.15–0.79] but were statistically significantly more likely to have more than 20 solar keratoses (OR = 3.61, 95% CI = 1.42–9.17); moreover, they were more prone to a past history of excised solar skin lesions (OR = 1.87, 95% CI = 0.89–3.92) (Whiteman *et al.*, 2003). This finding has been confirmed in Italian patients (Carli and Palli, 2003). This means that not only exposure in early life, generally intermittent, but also cumulative lifelong exposure may contribute to melanoma development. Recent data from Crete show that in the relatively dark-skinned population, sun exposure indices represent the most important risk markers for cutaneous melanoma, which contrast with data from fair-skinned Caucasian populations in which melanocytic naevi are the main risk factors (Lasithiotakis *et al.*, 2004). Therefore, the change in lifestyle with increasing exposure to UV owing to the growing popularity of sunbathing and tanning seem to have affected all age-groups as a period effect presumably occurred during late 1950s–early 1960s.

On the other hand, more frequent excision of pigmented skin lesions may contribute to explain the increasing number of melanomas diagnosed overtime in the Tuscany Cancer Registry area by means of a period effect. In the Tuscany area, the awareness among population and the development of early diagnosis activity has increased over the last decades; a preventive campaign addressed both to family doctors (general practitioners) and to general population for the surveillance of pigmented skin lesions is active in the RTT area since the late 1980s (Carli *et al.*, 2002). A Pigmented Lesion Clinic working in the Dermatology Department of University of Florence was

also implemented for rapid referral of participants with self-detected or GP-detected suspicious lesion (Carli *et al.*, 2002). A clue for the effect of early diagnosis was the growing rates of 'thin' lesions (≤ 1 mm) that showed a mean annual increase of about 16.1% from 1985 to 1997 (Crocetti and Carli, 2003). This situation probably explains the role, if any, of the period-effect in sustaining the incidence rates.

In conclusion, in the Tuscany Cancer Registry area MM rates will probably go on increasing in the next future. According to prediction model, the standardized incidence rate (age 15–84 years) for the period 2002–2006 will be 19.2 cases per 100 000, approaching that observed in northern European populations. Although the increasing trend was explained by an age-drift model, the two major explaining factors – changes in lifestyle with increasing exposure to UV and increased early diagnosis – seem to have acted more as period than as cohort effects.

References

Carli P, Palli D (2003). Re: melanocytic nevi, solar keratoses, and divergent pathways to cutaneous melanoma. *J Natl Cancer Inst* **95**:1801.

Carli P, De Giorgi V, Nardini P, Mannone F, Palli D, Giannotti B (2002). Melanoma detection rate and concordance between self-skin examination and clinical evaluation in patients attending a pigmented lesion clinic in Italy. *Br J Dermatol* **146**:261–266.

Clayton D, Schifflers E (1987a). Models for temporal variation in cancer rates. I: age-period and age-cohort models. *Stat Med* **6**:449–467.

Clayton D, Schifflers E (1987b). Models for temporal variation in cancer rates. II: age-period-cohort models. *Stat Med* **6**:469–481.

Crocetti E, Carli P (2003). Unexpected reduction of mortality rates from melanoma in males living in central Italy. *Eur J Cancer* **39**:818–821.

De Vries E, Bray FI, Coebergh JWW, Parkin DM (2003). Changing epidemiology of malignant cutaneous melanoma in Europe 1953–1997: rising trends in incidence and mortality but recent stabilizations in Western Europe and decreases in Scandinavia. *Int J Cancer* **107**:119–126.

Dyba T, Hakulinen T, Paivarinta L (1997). A simple non-linear model in incidence prediction. *Stat Med* **16**:2297–2309.

Elwood JM, Jopson J (1997). Melanoma and sun exposure: an overview of published studies. *Int J Cancer* **73**:198–203.

Lasithiotakis K, Kruger-Krasagakis S, Ioannidou D, Pediaditis I, Tosca A (2004). Epidemiological differences for cutaneous melanoma in a relatively dark-skinned Caucasian population with chronic sun exposure. *Eur J Cancer* **40**:2502–2507.

Lens MB, Dawes M (2004). Global perspectives of contemporary epidemiological trends of cutaneous malignant melanoma. *Br J Dermatol* **150**: 179–185.

Moller B, Fekjaer H, Hakulinen T, Sigvaldason H, Storm HH, Talback M, Haldorsen T (2003). Prediction of cancer incidence in the Nordic countries: empirical comparison of different approaches. *Stat Med* **22**:2751–2766.

Naldi L, Lorenzo Imberti G, Parazzini F, Gallus S, La Vecchia C (2000). Pigmentary traits, modalities of sun reaction, history of sunburns, and melanocytic nevi as risk factors for cutaneous malignant melanoma in the Italian population: results of a collaborative case–control study. *Cancer* **88**:2703–2710.

Paci E, Crocetti E, Miccinesi G, Benvenuti A, Intrieri T, Sacchettini C, Giovannetti L (2002). Tuscany cancer registry. In: Parkin DM, Whelan SL, Ferlay J, Teppo L, Thomas DB, editors. *Cancer incidence in five continents*. Vol VIII. Lyon: IARC Scientific Publications no. 155; 362–363.

Shryock HS, Siegel JS, Stockwell EG (1976). *The methods and materials of demography*. New York: Academic Press.

Whiteman DC, Watt P, Purdie DM, Hughes MC, Hayward NK, Green AC (2003). Melanocytic nevi, solar keratoses, and divergent pathways to cutaneous melanoma. *J Natl Cancer Inst* **95**:806–812.

Wolters Kluwer

Melanoma incidence in central Italy will go on increasing also in the near future: A registry-based, age–period–cohort analysis

Author: Emanuele Crocetti, Paolo Carli, and Guido Miccinesi

Publication: European Journal of Cancer Prevention

Publisher: Wolters Kluwer Health, Inc.

Date: Feb 1, 2007

Prognostic variables and prognostic groups for malignant melanoma. The information from Cox and Classification And Regression Trees analysis: an Italian population-based study

Emanuele Crocetti[a], Lucia Mangone[c], Giovanni Lo Scocco[d] and Paolo Carli[b]

The common way to analyse the prognostic role of selected variables in cutaneous melanoma patients is by means of Cox proportional hazard model. The prognostic effect of the simultaneous presence of more than one independent variable in the same patient is, however, difficult to establish. This hampers the possibility of tailoring a survival expectance for a selected patient as well as to communicate it to the patient himself/herself. The objectives of the study were to compare information on cutaneous melanoma prognosis from multivariate Cox proportional hazard model and from Classification And Regression Trees analysis. Classification And Regression Trees analysis is an automatic method that splits data by means of a binary recursive process creating a 'tree' of groups with different profiles according to the analysed outcome, for example, the risk of death. This approach automatically produces data that is easily interpreted by clinicians. A total of 1403 invasive cutaneous melanoma patients, 1110 from the Tuscan Cancer Registry and 293 from the Reggio Emilia Cancer Registry, Italy, were included. Cases were incident during 1996–2001 and followed up at the end of 2003. Cox proportional hazard model and Classification And Regression Trees analysis were applied to the following variables: age, sex, Breslow thickness, Clark level, registry, subsite and morphologic type. The Classification And Regression Trees analysis identified 10 categories with statistically different survival; this results were summarized into six classes of different risks based on Breslow thickness, age and sex. The best prognostic group (5-year observed survival, 98.1%) included those subjected with Breslow less than 0.94 mm and age 19–44 years. The same thickness but an older age (50–69 years) was associated with a statistically significant different prognosis (5-year observed survival, 92.8%). The Cox proportional hazard model found sex, age, Breslow thickness, Clark and morphologic type to have a significant independent prognostic value. In conclusion, compared with the conventional approach based on Cox hazard model, Classification And Regression Trees analysis produces data closer to the clinical need of defining the prognostic profile of a specific patient. This may help the clinician both in the communication of risk and in the follow-up strategy. *Melanoma Res* 16:429–433 © 2006 Lippincott Williams & Wilkins.

Melanoma Research 2006, **16**:429–433

Keywords: Classification And Regression Trees, Cox, malignant melanoma, prognosis, survival

[a]Tuscany Cancer Institute, Tuscany Cancer Registry, Clinical and Descriptive Epidemiology Unit, Centre for the Study and Cancer Prevention, [b]Dermatology Department, University of Florence, Florence, [c]Reggio Emilia Cancer Registry, Epidemiology Unit, Public Health Department, Reggio Emilia and [d]Dermatology Unit, Prato Hospital, Prato, Italy

Correspondence and requests for reprints to Dr Emanuele Crocetti, Via di San Salvi 12, Firenze 50135, Italy
Tel: +390556268320; fax: +39055679954; e-mail: e.crocetti@cspo.it

Received 5 December 2005 **Accepted** 1 May 2006

Introduction

For malignant cutaneous melanoma (CM), the identification of prognostic factors has a crucial value for both dermatologists and patients. According to new evidence on effective prognostic factors, the CM staging system has changed during recent years [1–4].

Usually, more factors act together, for example, thickness, age, site, Clark's level of invasion, ulceration, sex, etc. [2], and therefore their cumulative effect should be evaluated by means of multivariate models. The use of multivariate Cox proportional hazard model for survival analysis is common in the scientific literature [5]. For example, in the PubMed library, more than 900 papers matched a search for the terms 'Cox AND survival AND cancer', among those published during the year 2005 (overall 671 163). By means of this statistical approach however, the clinical relevance of the simultaneous presence in a patient of more than one independent prognostic factor cannot be easily established. This is a pity, as the possibility for the clinician of tailoring a prognostic profile for each patient taking into account the role of independent survival predictors simultaneously present would be clinically relevant both to plan follow-up and to communicate the personal risk to the patient him/herself.

During the last two decades, a tree-building technique has been slowly entering the literature. It is the so-called Classification and Regression Tree (CART) analysis [6]. This is an automatic method that splits data by means of a binary recursive process creating a 'tree' of groups with different profiles according to the analysed outcome, for

0960-8931

example, the risk of death. Although this approach automatically produces data that is easily interpreted by clinicians, it does not seem to be very popular yet.

The aim of the present study was to compare malignant melanoma survival information from both Cox and CART approaches in a multicentric, population-based Italian data set.

Material and methods

In the present study, invasive CM incident in the period 1996–2001 in the area of the Tuscany Cancer Registry (RTT) and of the Reggio Emilia Cancer Registry (RECR) were included.

Both RTT and RECR are located in central Italy and are included in the Italian Network of Cancer Registries (AIRT, www.registri-tumori.it). Further details on their organization and data management are available for RTT at http://www.cspoweb.it/rtt.asp and for RECR at http://www.ausl.re.it/Home/DocumentViewer.aspx?ID = 773&-TIPODOC = IAP.

Overall, 1403 malignant melanoma were analysed (1110 from RTT and 293 from RECR).

Follow-up was carried out up to 31 December 2003 or death, whichever came first. The mean follow-up time was 4.2 years, ranging from 0 to 8 years (median follow-up, 4.0 years).

The CART analysis was used to identify different prognostic groups. This is an automatic method that builds up a 'tree' by means of a binary recursive partitioning of data. Full data are evaluated for all the possible binary splits. It means that the method evaluates the possibility of splitting each variable into two groups if there is a statistically significant difference according to the analysed outcome (e.g. risk of death). The process repeats building up a tree until all groups are unsplittable [6].

The variables included in the CART analysis were as follows: sex (males, females), 5-year age classes, registry (RTT, RECR), Breslow's thickness (continue), Clark level (2–5, missing), site (head and neck, trunk, upper arm, lower arm, not specified), morphology [superficial spreading (SSM), nodular, others, not otherwise specified. We chose not to analyse separately lentigo maligna melanoma owing to their small number, $n = 31$].

Information on lymph node involvement was not available.

CART uses the martingale residuals of a Cox model to calculate (approximate) χ^2 values for all the possible cut points on all the CART covariates. The significant value for cut point definition was 0.05. The minimum size of the group was defined as 30 participants.

Once CART had identified different groups, Kaplan–Meier survival curves were computed and compared by means of log-rank test [7]. Only statistically significant groups were considered for the definition of the final groups of patients with significantly different risks of dying.

Moreover, we analysed the prognostic role of the same variables also in a more conventional approach by means of the Cox conditional hazard model. Within each variable, the statistical significance of the difference between values was evaluated; the original variable definition was modified according to the comparison, in particular the site was been recoded in head and neck, other specified, not specified.

The variables that showed a significant effect in the univariable analysis were included in a multivariable Cox proportional hazard model to evaluate their independent effect. By means of a step-forward approach, the effect of each variable in improving the fit of the model was evaluated with the likelihood ratio test.

The product-limit estimate according to Kaplan–Meier was used for computing overall and variable-specific observed survival probabilities [8].

The analysis was performed by Stata 8 College Station, Texas (www.stata.com).

Results

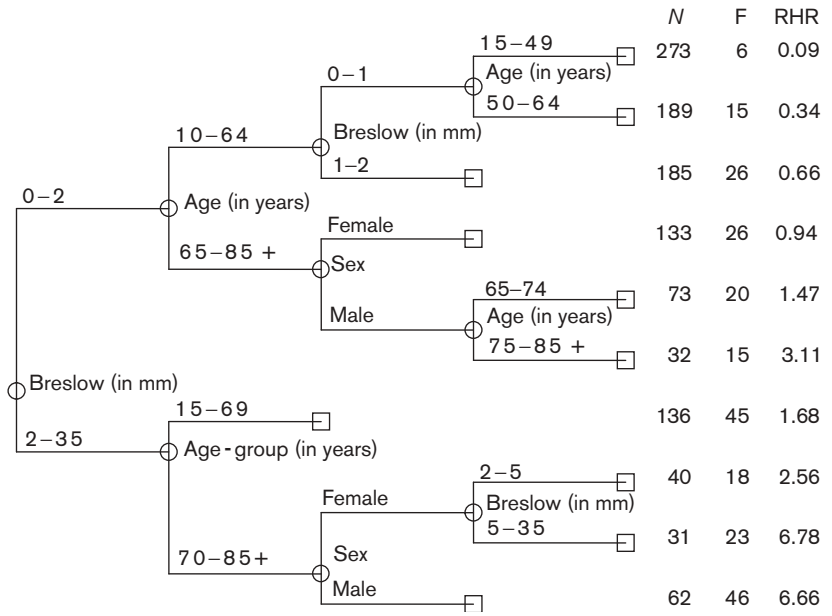
The analysis was based on 1403 patients with newly diagnosed, incident CM; during the follow-up (mean time, 50.5 months) 343 of these patients died.

The overall observed survival was 93.1% at 1-year, 82.7% at 3-year and 74.7% at 5-year period.

In the CART analysis, 1154 patients with complete information for sex, age, Breslow's thickness, Clark's levels, registry, subsite and morphology were included and all those variables were evaluated. The main relevant split involved Breslow's thickness identifying lesions thinner and thicker than 2.25 mm. The second split of the regression tree involved age classes. The third split was based on sex, and again on Breslow thickness, as the fourth one (Fig. 1).

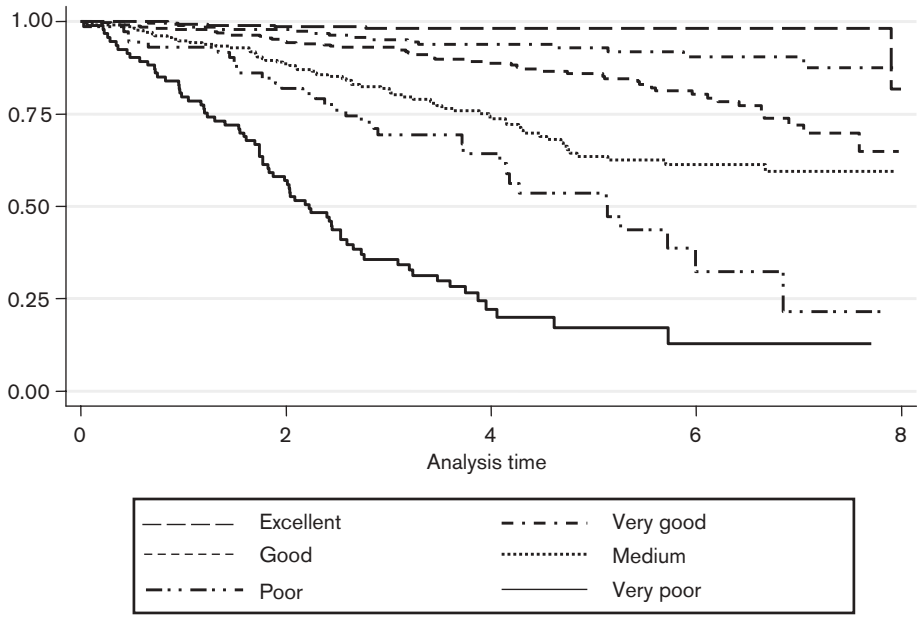
The CART analysis identified 10 prognostic groups according to Breslow thickness, sex and age. For each of these groups, Kaplan–Meier survival curves were computed and compared by means of log-rank test. When

Fig. 1



Classification And Regression Trees with the following variables: sex, registry, subsite, age, Breslow, Clark and morphology. Split if (adjusted) $P<0.05$. RHR, relative hazard ratio.

Fig. 2



Kaplan–Meier survival rates of the six prognostic groups defined by Classification And Regression Trees analysis (group definition is given in Table 1).

there was no statistical difference in survival between groups, they were summed. The result produced six groups of patients with significantly different risks of dying (Fig. 2).

In Table 1 the CART results are used to show a straightforward way to define these six prognostic groups according to Breslow, age class and sex. The best prognostic group (5-year observed survival, 98.1%)

Table 1 Malignant melanoma, 1996–2001

Group of prognosis		Prognostic factors			5-year observed survival
	No.	Breslow (mm)	Age (years)	Sex	%
Excellent	273	<0.94	19–49	Female/Male	98.1
Very good	189	<0.94	50–64	Female/Male	92.8
Good	318	0.93–2.24	10–64	Female/Male	85.9
		<2.25		Female	
Medium	209	<2.25	65–74	Male	63.5
		≥ 2.25	15–69	Female/Male	
Poor	72	<2.25	75+	Male	53.5
		2.25–4.99	70+	Female	
Very poor	93	≥ 5.00	70+	Female	17.1
		≥ 2.25	70+	Male	

Prognostic groups according to Classification And Regression Trees analysis. Each group is defined according to the characteristics of the included patients, for example, male and female patients with Breslow lower than 0.94 mm, 19–49 years old belong to the group with excellent prognosis.

Table 2 Malignant melanoma

Variable	No.	Crude HR	95% CI	Adjusted HR	95% CI
Sex					
Women	762	1		1	
Men	644	1.40	1.14–1.73	1.30	1.05–1.61
Age (annual)					
<60	784	1		1	
60+	622	3.73	2.95–4.70	2.92	2.30–3.71
Registry					
Firenze	1113	1			
Reggio Emilia	293	0.96	0.73–1.24		
Breslow (mm)					
<1	631	1		1	
≥ 1	524	3.93	2.95–5.23	1.97	1.27–3.07
Missing	251	5.39	3.95–7.37	1.76	0.96–3.23
Clark					
2	396	1		1	
3	361	1.60	1.06–2.41	1.09	0.67–1.76
4	337	4.04	2.79–5.84	1.45	0.84–2.51
5	48	11.08	6.85–17.92	2.89	1.51–5.53
Missing	264	5.72	3.95–8.29	1.79	0.93–3.45
Site					
Head and neck	136	1			
Other specified	1085	0.43	0.32–0.57		
Not specified	185	0.91	0.63–1.32		
Morphology					
Superficial spreading	903	1		1	
Nodular	133	3.72	2.76–5.03	1.59	1.14–2.22
n.o.s.	225	4.14	3.21–5.33	2.23	1.61–3.08
Other	142	2.07	1.44–2.97	1.19	0.81–1.73

Crude and adjusted risk of dying (hazard ratio -HR) for several prognostic variables with corresponding 95% confidence intervals (95% CI).
The multivariable model includes sex, age, Breslow's thickness, Clark's level and morphology subtype.
n.o.s., not otherwise modified.

included male and female patients with Breslow thickness less than 0.94 mm and age 19–44 years. The same thickness but an older age (50–69 years) was associated with a statistically significant different prognosis (5-year observed survival 92.8%). The prognosis was good (85.9% at 5 years) for both men and women, with Breslow thickness between 0.93 and 2.24 mm and age 10–64 years or for females with thickness less than 2.25 mm. The next group (5-year survival, 63.5%) included men aged 65–74 years with Breslow < 2.25 mm and both men and women aged 15–69 years with Breslow ≥ 2.25 mm. The prognosis was poor (5-year survival, 53.5%) for men with Breslow thickness < 2.25 mm and age 70+ years and for women with Breslow thickness 2.25–4.99 mm and age 70+ years. The group with the worst prognosis (5-year survival, 17.1%) included participants of 70 or more years of age, women with Breslow ≥ 5.00 mm, or men with Breslow ≥ 2.25 mm.

Data on traditional survival analysis based on univariate approach followed by multivariate Cox regression models are shown in Table 2. The survival was significantly higher among women than among men, among younger patients (< 60 years) than among older ones, for melanoma with Breslow thickness less than 1 mm than for ≥ 1 mm, for smaller Clark's levels, for specified subsites (or not-specified ones) than for head and neck and for superficial spreading melanoma than for not specified ones. No statistically significant differences were found between the two registries.

Discussion

The population-based series of CM analysed in this study yielded results that were perfectly in line with that expected – according to the literature – in terms of overall 5-year survival (75%) and significance of major variables of risk of death. Indeed, the conventional Cox hazard analysis showed that Breslow thickness, sex, Clark level, morphology and age resulted in independent prognostic variables [9,10].

In this study, a CART analysis of survival was performed in order to easily investigate what happens when more than one independent prognostic variable is present in a patient. CART analysis enables clinicians to identify subgroups of risk – characterized by the simultaneous presence of more than one variable – that differ significantly among themselves in terms of survival. Indeed, by means of this approach, the identification of different subgroups of risk is possible, with easier risk communication and, eventually, a more patient-tailored follow-up strategy.

In the present study, the CART analysis constructed a pruned tree of 10 groups. Breslow thickness defined the first and the third split [9,11]. Moreover, the cut-point automatically identified by the software in the third split (0.94 mm) almost overlapped the well-accepted limit of 1 mm used to separate thin from thick melanomas. The original 10 groups were reduced, according to Kaplan–Meier comparisons, into six homogeneous groups with statistically significant different risks of dying. The attribution of a patient to the corresponding group seems very simple and it is based on the age, sex and Breslow. Using this method, it is possible to define the prognosis of a patient in a qualitative (from excellent to very poor) and quantitative (from 98.1 to 17.1% survival at 5 years) way. For example, facing a patient with a melanoma thinner than 0.94 mm, the prognosis significantly changes according to age with better 5-year survival (98.1%) if the patient is younger than 44 years than that expected if older than 50 (92.8%, $P < 0.01$). It is interesting that the effect of sex as independent prognostic factor seems to act for some subgroup, only when balanced with the effect of major predictors of risk such as Breslow thickness and age. Indeed, the sex was not able to significantly modify the risk of dying facing the subgroups at better prognosis (defined in this study excellent and very good (5-year observed survival, more than 92%); on the contrary, sex was statistically relevant facing the other groups at less favourable prognosis.

The results from the present study may be influenced by the number of cases included and by the underlying tumours characteristics; however, the three prognostic variables identified in the present study were the same as that in a study including 12728 German patients [9].

The use of the results from Cox analysis seems not so easy. In fact, Cox model produces for each variable coefficients (β) that show the relationship between the specific value of the variable and the outcome. The overall risk should take into account all the computed coefficients, according to the following formula: exp(coefficient [e.g. sex] + coefficient [e.g. age] + coefficient [e.g. Breslow] + coefficient [e.g. Clark] + etc.) in comparison with all the variables in the reference categories.

Although the CART approach has more closeness to the clinical reasoning process than regression models, it is not as common as expected. The main reason for the low dissemination of CART seems to have been just the low diffusion itself that had hampered both the growth of statistical interest and knowledge of it and therefore the availability of specific procedures within more diffused software packages. Although it has been criticized [12], its application has also been suggested [13], and the method has been developed over time [14], also to overcome the limits of traditional models [15]. As regards the latter point, one of the advantages of CART analysis is just to be a nonparametric method that works with quantitative and qualitative data without any assumptions on their distributions [16].

The present method has several appeals for clinicians: its automatic process, the easily interpretable results and the growing availability of software packages [16]. The number of papers that use this method to evaluate cancer prognosis is growing, and this will stimulate discussion and knowledge on this possibly very useful clinical tool.

In conclusion, the present study contributes to demonstrate the possible usefulness, in addition to traditional approach, of the CART analysis for defining the risk profile of a melanoma patient in a clinical set.

References

1 Balch CM, Buzaid AC, Atkins MB, Cascinelli N, Coit DG, Fleming ID, *et al.* A new American Joint Committee on Cancer staging system for cutaneous melanoma. *Cancer* 2000; **88**:1484–1491.

2 Balch CM, Buzaid AC, Soong S-J, Atkins MB, Cascinelli N, Coit DG, *et al.* Final version of the American Joint Committee on cancer staging system for cutaneous melanoma. *J Clin Oncol* 2001; **19**:3635–3648.

3 Balch CM, Buzaid AC, Soong S-J, Atkins MB, Cascinelli N, Coit DG, *et al.* New TNM melanoma staging system: linking biology and natural history to clinical outcomes. *Semin Surg Oncol* 2003; **21**:43–52.

4 Balch CM, Soong S-J, Atkins MB, Buzaid AC, Cascinelli N, Coit DG, *et al.* An evidence-based staging system for cutaneous melanoma. *CA Cancer J Clin* 2004; **54**:131–149.

5 Cox DR. Regression models and life tables. *J R Stat Soc* 1972; **34**: 187–220.

6 Breiman L, Friedman JH, Olshen RA, Stone CJ. *Classification and regression trees*. New York: Chapman & Hall (Wadsworth, Inc); 1984.

7 Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemother Rep* 1966; **50**:163–170.

8 Kaplan EL, Meier P. Non-parametric estimation from incomplete observations. *J Am Stat Assoc* 1958; **53**:457–481.

9 Leiter U, Buettner PG, Eigentler TK, Garbe C. Postic factors of thin cutaneous melanoma: an analysis of the central malignant melanoma registry of German dermatological society. *J Clin Oncol* 2004; **22**:3651–3653.

10 Garbe C, Buttner P, Bertz J, Burg G, d'Hoedt B, Drepper H, *et al.* Primary cutaneous melanoma. Identification of prognostic groups and estimation of individual prognosis for 5093 patients. *Cancer* 1995; **75**:2484–2491.

11 Garbe C, Ellwanger U, Tronnier M, Brocker EB, Orfanos CE. The new American Joint Committee on cancer staging system for cutaneous melanoma. A critical analysis based on data of the German central malignant melanoma registry. *Cancer* 2002; **94**:2305–2307.

12 Marshall RJ. The use of classification and regression trees in clinical epidemiology. *J Clin Epidemiol* 2001; **54**:603–609.

13 Lemon SC, Roy J, Clark MA, Friedmann PD, Rakowski W. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. *Ann Behav Med* 2003; **26**:172–181.

14 Friedman JH, Meulman JJ. Multiplice additive regression trees with application in epidemiology. *Stat Med* 2003; **22**:1365–1381.

15 Xu R, Adak S. Survival analysis with time-varying regression effects using a tree-based approach. *Biometrics* 2002; **58**:305–315.

16 Lewis RJ. An introduction to classification and regression tree (CART) analysis. Presented at the 2000 Annual Meeting of the Society for Academic Emergency Medicine in San Francisco California (www.saem.org/download/lewis1.pdf).

Prognostic variables and prognostic groups for malignant melanoma. The information from Cox and Classification And Regression Trees analysis: an Italian population-based study

Author: Emanuele Crocetti, Lucia Mangone, Giovanni Scocco, et al

Publication: Melanoma Research

Publisher: Wolters Kluwer Health, Inc.

Date: Oct 1, 2006

Indicators of the standard of care for melanoma: Tuscany data

Emanuele Crocetti[a], Adele Caldarella[a], Daniela Massi[c], Claudio Sacchettini[a], Gianni Amunni[b,d] and Lorenzo Borgognoni[e]

Formal indicators for the evaluation of the quality of melanoma care are needed. We identified 13 process indicators, which encompassed early diagnosis, pathology reporting and surgical treatment. We evaluated the adherence to these indicators using a population-based series on incident skin melanomas (only primary melanomas) for the year 2004 (687 cases) and for the first semester of 2008 (539 cases). We compared the indicators for these 2 years. The melanoma incidence increased between 2004 and 2008. There were statistically significant increases in the percentage of thin (≤ 1 mm) melanomas (from 50.7 to 61.3%) and in the number of pathology reports that mentioned ulceration (from 61.4 to 84.6%) and margin status (from 76.8 to 84.3%). The percentage of patients staged by sentinel lymph node biopsy was stable (63%) and was higher for patients younger than 75 years of age (74%). The number of nodes almost invariably exceeded the proposed site-specific cutoff reference and, in 2008, the number of nodes removed was always reported for lymphadenectomy. From 2004 to 2008, surgical and pathological waiting times increased. Collection and analysis of these indicators would enable continuous evaluation of the quality of melanoma care in Tuscany and provide sources for a comparative study between Italy and elsewhere. ***Melanoma Res*** **23:283–289 © 2013 Wolters Kluwer Health | Lippincott Williams & Wilkins.**

Melanoma Research 2013, **23**:283–289

Keywords: epidemiology, practice variations, quality indicators, skin melanoma

[a]Clinical and Descriptive Epidemiology Unit, Institute for Cancer Study and Prevention (ISPO), [b]Institute for Cancer Study and Prevention (ISPO), [c]Department of Critical Care Medicine and Surgery, Division of Pathological Anatomy, [d]Department of Gynaecology, Perinatology and Human Reproduction, University of Florence and [e]Plastic and Reconstructive Surgery Unit, Regional Melanoma Referral Centre, S.M. Annunziata Hospital, Tumour Institute of Tuscany (ITT), Florence, Italy

Correspondence to Emanuele Crocetti, MD, Clinical and Descriptive Epidemiology Unit, Institute for Cancer Study and Prevention – ISPO, Via delle Oblate 2, 50141 Florence, Italy
Tel: +39 055 797 2508; fax: +39 055 797 2535;
e-mail: e.crocetti@ispo.toscana.it

Received 13 September 2012 **Accepted** 23 April 2013

Introduction

The incidence of skin melanoma has greatly increased in Italy during recent years [1], as in many other western countries. In Tuscany (central Italy), cancer care is coordinated by the Tumour Institute of Tuscany (ITT), a network-based Institute that organizes and supervises public health services in the community with the aim of providing optimized cancer treatment. The ITT also conducts studies on tumours (*http://www.ittumori.it*). In particular, in 2007, it published clinical recommendations on standardizing diagnostic and therapeutic pathways for skin melanoma for specialists and hospitals throughout the region. On the basis of the ITT's recommendations, the standard and effective treatment for cutaneous melanoma is now surgery-based [2]. As most skin melanomas are diagnosed early (when, in most cases, the lesion is thin) [3], complete pathological examination and timely surgery represent the best option for the majority of patients. To reduce variability in cancer care procedures and improve the general quality of care, the ITT works on the basis of indicators. Specific indicators for breast, lung, colorectal and ovarian cancer care are available [4].

The role of indicators in evaluating clinical care is increasing [5,6]. However, to our knowledge, in Italy, there are no formal indicators for care quality for melanoma. In the USA, the four measures available in the National Quality Measures Clearinghouse address only a few points in the clinical course (diagnosis, imaging and follow-up, *http://www.qualitymeasures.ahrq.gov/*), although further indicators, mainly for surgical treatment, have recently been evaluated [7].

This paper aims at evaluating the clinical care of melanoma in Tuscany through formal indicators and gauges whether ITT recommendations have succeeded in instituting and improving standardized patterns of care.

Materials and methods

Setting and study population

We first evaluated primary skin melanoma cases occurring in Tuscany between 2004 and 2008, excluding melanomas with unknown primary sites. We retrieved melanoma incidence data for 2004 from the archive of the Tuscan Regional Cancer Registry (RTRT), a population-based cancer registry that collects data from several information sources, including hospital discharge notes, pathology reports, and death certificates [8]. The RTRT was not, however, active in the Tuscan province of Arezzo in 2004, accounting for 330 123 out of 3 566 071 Tuscan inhabitants.

We identified melanoma cases occurring in 2008 from the regional pathology archive (*http://web.rete.toscana.it/attinew/*

0960-8931 © 2013 Wolters Kluwer Health | Lippincott Williams & Wilkins DOI: 10.1097/CMR.0b013e328362916a

?LO = 00000001a6b7c8d900000003000000340000 3f7a397d5a0f000000000012b800000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000&MItypeObj = application/pdf), which includes all pathology reports, positive or uncertain, for cancers, as defined by SNOMED (vers. 3) morphology codes M-8**** and M-9****. We also collated the incidence of the term 'melanoma' in each pathology report from the first semester of 2008 (from January to June). We included individuals residing in Tuscany and restricted our interest to those with primary melanomas, making sure that the individuals had no pathology reports nor hospital discharge notes, nor had they undergone any outpatient treatment for melanoma previously. When melanoma-related reports were not available (neither in the RTRT nor in the regional pathology archive system), as was the case for a few individuals, we searched through the reports in the regional pathology departments. In total, we analyzed 687 melanomas from 2004 (161, 23.4% *in situ*) and 539 from the first semester of 2008 (143, 26.5% *in situ*). For each patient, we collected detailed information on pathology characteristics, diagnostic methods, staging and type and dates of surgery.

The list of indicators

Working groups within the ITT studied specific cancers in depth, in order to better define and exchange experiences on specific topics and to make better clinical recommendations. The melanoma working group is a multidisciplinary panel of individuals and institutions, each belonging to the regional cancer network, including experts in melanoma (dermatologists, pathologists, surgeons, oncologists and epidemiologists), academic institutions and community hospitals. In 2007, the ITT published a set of recommendations on the diagnosis and treatment of melanoma based on the investigations of the melanoma working group, whose components agreed upon a consensus procedure based on scientific literature and on their experience [2].

The ITT Melanoma Working Group (see Acknowledgements) identified a preliminary set of indicators by analysing scientific guidelines on melanoma diagnosis and treatment, literature reviews [9–42] and panel discussions. A necessary condition while selecting the indicator was that only data from the Cancer Registry could be used, so that a homogeneous analysis could be carried out and objective results valid for the whole of Tuscany could be obtained. After meetings, discussions and email exchanges, the Group selected 13 indicators. The chosen indicators measured the appropriateness of the clinical care procedure on the basis of ITT recommendations [2] and encompassed the diagnosis, pathology reporting and surgical treatment of melanoma care.

Statistical analysis

We calculated percentages and medians for quality measures using the χ^2-test or median test to compare values between the years. The results were reported at a regional level but were also available at subregional levels for local use.

Results

Table 1 presents the selected quality indicators.

Table 2 compares the average regional values of each quality indicator for 2004 and 2008.

Percentage of incident cases $\leq$ 1 mm

The thickness of a melanoma is one of the most relevant prognostic factors. The ITT and other recommendations state that a pathology report must specify the thickness. The number of new melanoma cases increased from 2004 to 2008, the standardized incidences being 14.1 and 18.4 per 100 000, respectively. Almost all the pathology reports recorded information on Breslow's thickness (97.5% in 2004, 513/526 and 98.0% in 2008, 388/396). The percentage of thin ($\leq$ 1 mm) melanomas increased over time from 50.7 to 61.3% ($P = 0.001$). Moreover, there was a significant percentage of in-situ melanomas: 23.9% (161

Table 1 Quality indicators of skin melanoma care and areas of interest

	Area	Main indicators
1	Early diagnosis	Percentage of incident cases $\leq$ 1 mm
2	Pathology reporting	Percentage of pathology reports on incident invasive melanomas with ulceration mentioned (present/absent)
3	Pathology reporting	Percentage of pathology reports on excised incident invasive melanomas with the margin status (positive/negative)
4	Surgery	Percentage of patients staged by SLB
5	Surgery	Percentage of patients staged by SLB <75 years of age
6	Surgery–pathology	Percentage of patients with positive sentinel nodes
7	Pathology reporting	Percentage of lymphadenectomies with the number of removed lymph nodes reported
8	Surgery–pathology	Percentage of cervical lymphadenectomies with >15 lymph nodes removed
9	Surgery–pathology	Percentage of axillary lymphadenectomies with >10 lymph nodes removed
10	Surgery–pathology	Percentage of inguinal lymphadenectomies with >5 lymph nodes removed
11	Time for pathology	Waiting time between the first surgery and the pathology report (time for pathology of diagnosis) (days, median)
12	Time for surgery	Waiting time between the diagnostic pathology report and the second surgery (enlargement/SLB – time for surgery) (days, median)
13	Time for overall surgical treatment in positive SLB	Waiting time between the first melanoma diagnosis and the final lymphadenectomy in patients with a positive SLB (days, median)

SLB, sentinel lymph node biopsy.

Table 2 Average regional value of the quality indicators of skin melanoma care for 2004 and 2008, probability of equality according to testing (χ^2, rank sum)

	Main indicators	Regional average 2004	Regional average 2008	*P*
1	Percentage of incident cases $\leq$ 1 mm	50.7	61.3	0.001
2	Percentage of pathology reports on incident invasive melanomas with ulceration mentioned (present/absent)	61.4	84.6	<0.001
3	Percentage of pathology reports on excised incident invasive melanomas with the margin status (positive/negative)	76.8	84.3	0.005
4	Percentage of patients staged by SLB	63.6	63.3	0.95
5	Percentage of patients staged by SLB <75 years old	73.9	74.3	0.95
6	Percentage of patients with positive sentinel nodes	26.1	26.3	0.97
7	Percentage of lymphadenectomies with the number of removed lymph nodes reported	98.2	100	0.51
8	Percentage of cervical lymphadenectomies with >15 lymph nodes removed	100	100	1
9	Percentage of axillary lymphadenectomies with >10 lymph nodes removed	85.7	66.7	0.20
10	Percentage of inguinal lymphadenectomies with >5 lymph nodes removed	100	100	1
11	Waiting time between the first surgery and the pathology report (time for pathology of diagnosis) (days, median)	9	13	<0.001
12	Waiting time between the diagnostic pathology report and the second surgery (enlargement/SLB – time for surgery) (days, median)	23	26	0.084
13	Waiting time between the first melanoma diagnosis and the final lymphadenectomy in patients with a positive SLB (days, median)	72	82.5	0.71

SLB, sentinel lymph node biopsy.

out of 674 invasive and in-situ melanomas) in 2004 and 26.9% (143/531) in 2008.

Percentage of pathology reports on incident invasive melanomas mentioning ulceration (present or absent)

Ulceration has been included in the staging process since 2002 [42]. The ITT recommendations state that pathology reports must specify ulceration [2]. The percentage of relevant pathology reports mentioning ulceration statuses (present or absent) increased over time from 61.4% (404/526) in 2004 to 84.6% (334/396) in 2008 ($P < 0.001$). This increase concerned both thin (from 53.1 to 82.5%, $P < 0.001$) and thick melanomas (from 70.4 to 88.0%, $P < 0.001$).

Ulceration was present in a decreasing percentage of reports over time: 34.4% in 2004 (111/323) and 25.7% in 2008 (86/335) ($P = 0.005$).

Percentage of pathology reports on excised incident invasive melanomas with margin statuses (positive or negative)

The margin status confirms the completeness of melanoma excision. The ITT recommendations state that pathology reports must specify the margin status [2]. The percentage of the margin status (whether positive or negative) mentioned in the pathology reports increased significantly, from 76.8% (404/526) in 2004 to 84.3% (334/396) in 2008 ($P = 0.005$). On stratifying for thin and thick melanomas, an almost similar statistically significant increase was observed (thin 79.1 vs. 85.4%, $P = 0.064$; thick 74.3 vs. 82.7%, $P = 0.052$).

In 2008, 92% of the excisions had negative margins. Among the 24 cases with positive margins, 15 (62.5%) dealt with large melanomas of the face and acral sites, for which an incisional biopsy was planned.

Percentage of patients staged by sentinel lymph node biopsy

Sentinel node biopsy (SLB) enables melanoma pathological staging. The sentinel node (SN) status is a major prognostic factor. The ITT recommendations state that SLB has to be performed when the thickness is at least 1 mm, in cases of ulceration or when the Clark level is at least IV [2]. The percentages of SLB were similar in 2004 and 2008: 63.6% (161/253) and 63.3% (95/150), respectively.

Percentage of cases with sentinel lymph node biopsy in patients <75 years old

SLB is a technique underutilized in elderly patients [43]. As it is more invasive than excision, it may sometimes not be indicated in older patients because of the presence of comorbidities. Moreover, most SLB trials did not include patients above 75 years of age [17,44,45]. Among the total number of patients who underwent SLB, the majority were less than 75 years of age – 73.9% in 2004 (139/188) and 74.3% in 2008 (75/101).

Percentage of patients with a positive sentinel node

The SN status is a major prognostic factor; therefore, node/s should be correctly removed and pathologically examined [2]. Around one-quarter of SLBs were positive both in 2004 [26.1% (42/161)] and in 2008 [26.3% (25/95)].

Percentage of lymphadenectomies with the number of nodes reported

The ITT recommendations state that the pathological report must specify the number of nodes removed after lymphadenectomy [2]. The number of nodes with metastases is a basic point for staging [42]. The number of nodes was mentioned in almost all the relevant pathology reports in both years. The number of the nodes removed in cases of complete node dissection was

also reported in almost all the cases, both in 2004, 98.2% (54/55) and in 2008, 100% (24/24).

Indicators 8, 9 and 10 concern the number of lymph nodes removed during lymphadenectomy. The number of lymph nodes removed by lymphadenectomy could be considered a measure of the quality of surgical care.

Percentage of cervical lymphadenectomies with >15 lymph nodes removed

In cervical lymphadenectomy, a cutoff value of 15 lymph nodes has been adopted [46]. The number of nodes removed exceeded the cutoff value in all cases in both years. In 2008, the median number of excised nodes was 24.5.

Percentage of axillary lymphadenectomies with >10 lymph nodes removed

In axillary lymphadenectomy, a cutoff value of 10 lymph nodes has been adopted [46]. The number of nodes removed exceeded the cutoff value in 85.7% of cases in 2004 (24/28) and in 66.7% (6/9) cases in 2008. In 2008, the median number of excised nodes was 16.

Percentage of inguinal lymphadenectomies with >5 lymph nodes removed

In inguinal lymphadenectomy, a cutoff value of five lymph nodes has been adopted [46]. The number of nodes removed exceeded the cutoff value for all patients in both years. In 2008, the median number of excised nodes was 14.

Indicators 11, 12 and 13 concern the waiting time.

Waiting time between the first surgery and the corresponding pathology report

The time between the removal of a suspected skin lesion and the pathological diagnosis of melanoma may influence the timeliness of the necessary subsequent actions (enlargement, SLB, etc.). The median diagnostic time was 9 days in 2004 and 13 days in 2008 ($P < 0.001$). On analysing thin and thick melanomas separately, the median time increased for thin melanomas (from 9 to 14 days, $P < 0.001$) and was stable for thick ones (from 10 to 11, $P = 0.509$).

Waiting time between the diagnostic pathology report and second surgery (enlargement/sentinel node biopsy)

The ITT recommendations [2] suggest performing enlargement or SLB within 3 months from the biopsy. The median time was 23 days in 2004 and 26 days in 2008 ($P = 0.084$). There were no differences with regard to Breslow's thickness.

Waiting time between the first melanoma diagnosis and final lymphadenectomy in patients with positive sentinel node (time for overall treatment in sentinel node positive patients)

The total waiting time for the surgical treatment of a melanoma patient is the time from the first biopsy to the report of the lymphadenectomy (in the case of a positive SN). It was reported to be 72 days in 2004 and 82.5 days in 2008 ($P = 0.71$).

Discussion

This study provides a comprehensive population-based evaluation of the diagnostic and surgical treatment of melanoma in the central Italian region of Tuscany. Our proposed indicators are based on local cancer registry data and the official regional pathology databases. The method of computation of the indicators enables us to compare different years or geographical/administrative areas when the same information systems are available. The population-based setting theoretically provides unbiased information and shows average measures for clinical care in the population. In Tuscany, a multilevel strategy for early diagnosis of skin melanoma has been in use since the 1990s. The strategy includes specific initiatives for raising awareness in the population, namely, distribution of information packs for skin self-examination, organizing training courses for general practitioners (GP) and setting up skin units across the region. In addition, the multidisciplinary ITT Melanoma Working Group has developed specific clinical recommendations to optimize the clinical course throughout the region [2].

A hypothetical and optimal clinical course for skin melanoma should start when an individual discovers a changing skin lesion, the GP confirms the need for further evaluation, the skin clinic dermatologist performs the excisional biopsy and the pathologist reports the diagnosis, specifying all the main prognostic factors that address the subsequent appropriate surgery. Therefore, the aim of the ITT Melanoma Working Group in identifying possible melanoma indicators was to find parameters useful in evaluating the diagnostic phase, the completeness of the pathological reporting and the surgical treatment of the melanoma. The indicators analysed in the present study were selected by the ITT Melanoma Working Group on the basis of data available in the Cancer Registry in order to have unbiased, homogeneous and comparable information. On the basis of the above requirements, the ITT Melanoma Working Group analysed scientific guidelines on melanoma diagnosis and treatment and performed literature reviews [10–41] and panel discussions with the purpose of identifying possible indicators of an early diagnosis and appropriate melanoma treatment. After discussions, meetings, reports and email exchanges, the Melanoma Working Group selected 13 indicators on the basis of the Cancer Registry data (Table 1). All measures were process indicators that are among the most suitable management tools for measuring quality [9].

During the analysed period, there was a strong increase in melanoma diagnosis, in both the standardized and crude rates. As they are based on the number of cases, crude rates show the real workload of a healthcare system in

terms of the need for diagnosis and treatment. This is the rate that must be taken into account when the 2 years are compared. The increased incidence was mainly driven by the growing percentage of thin melanomas (≤ 1 mm), which in 2008 was 61.3% of all newly diagnosed invasive melanomas. This indicator attests to the efficacy of early diagnosis in the region, as a consequence of increased awareness in the population, the role of the GP and of the activity of the dermatologists in skin clinics (with the support of dermoscopy). In the area of the present study, a randomized clinical trial confirmed the usefulness of dermoscopy in improving melanoma diagnosis and reducing the number of pigmented skin lesions excised for diagnostic verification [47]. The possible use of dermoscopy in the diagnosis of pigmented skin lesions is included in the ITT melanoma recommendations [2,20,47].

Mortality for melanoma is still increasing in Tuscany [48]. No effective treatments for advanced melanomas were available in Italy at the time of the study. Encouraging results from new molecular drugs have emerged only recently [49–51]. Therefore, early diagnosis and appropriate surgical treatment are the primary recourse for the healthcare system. However, an increase in diagnostic drift and instances of reclassification of previous severely atypical dysplastic nevi as melanoma has been found [52]. The increased diagnoses of thin but not deadly melanomas [3] have a positive significance, although the possible role of overdiagnosis of indolent pigmented lesions should also be considered [53].

Melanoma prognosis is strictly dependent on characteristics that may be evaluated only by a pathologist, and, for this reason, there are indicators that address the completeness of the pathology reports. The quantity of information available in the pathology reports has increased over time. In 2008, almost all reports included Breslow's thickness (98%), around 85% included the presence or absence of ulceration, 84% included the margin status for excised melanomas and all reports included the number of nodes in patients of lymphadenectomy. The latter results were as good as those recently documented in the USA [7].

As regards the weaknesses of this study, for the patients diagnosed in 2008, we used only the information from pathology reports. Therefore, few cases may have been lost and the estimates may be slightly overestimated.

Moreover, not all the relevant prognostic factors were analysed; for example, the number of mitoses was not included. However, the analysed data refer to 2004 and 2008, and it was only in 2009 that AJCC melanoma staging [54] included the number of mitoses as an important prognostic factor. For this reason, recent melanoma guidelines, such as those in the UK [12], include the mitotic count among the requirements for a pathological report, and this parameter will be included in the updated ITT recommendations for melanoma that are to be published. They will also be included on future indicator analyses.

About two-thirds of eligible patients had SLB. The percentage for patients younger than 75 years (74%) shows that this procedure was performed mainly on nonelderly patients. Presumably, the presence of comorbidity in patients older than 75 years influenced the decision of the surgeon for SLB. Once performed, an SLN biopsy showed approximately the same rate of positivity in 2004 as that in 2008 – about 26%. The mean regional figure of 26% in Tuscany is higher than the percentage of 16% reported in the Multicenter Selective Lymphadenectomy Trial I [17] and the 19.9% reported in an Italian Multicentre Study [55], indicating appropriate surgical and pathological approaches. Pathology reports always specify the number of lymph nodes removed during lymphadenectomy in the entire region. The cutoff values obtained by Bilimoira in recent work on melanoma indicators in the USA [7] were adopted to evaluate the quality of the surgical procedures in stage III patients. Moreover, the same cutoff values made it possible to compare the results between the two studies. The number of nodes removed from sites in the cervical and inguinal regions always exceeded the levels of 15 and five nodes, respectively. As regards the axillar region, two-thirds of patients had 10 or more nodes removed in the first semester of 2008. However, this number is based on just nine patients. The values registered in Tuscany are higher than those reported in Bilimoira's study [7]. Moreover, it has been reported in a previous American study that less than 50% of patients with positive SLB in the USA underwent complete node dissection [46], whereas the percentage in Tuscany was 63%.

A crucial aspect for any public healthcare system, for both patients and healthcare policy makers, is the waiting time. Three indicators addressed waiting times in the diagnostic and surgical course of melanoma. The median time for receiving the report for a skin lesion suspected to be a melanoma was 13 days in 2008. The length of this period depended on the pathologist, his/her workload, and the availability of resources. It increased from 2004 by 44%; however, the crude incidence rate (number of melanomas) increased during the same period by about 33% and, therefore, the longer waiting time for melanoma diagnosis resulted mainly from the increased number of patients awaiting diagnosis. The waiting time after a pathology diagnosis depended on the surgeons' workload and resources. The waiting time, for wider excision or SLB, increased by 13% (median from 23 to 26 days) during the aforementioned period (during which there was an increase in melanoma incidence). In particular, there was an increase not only in the number of thin melanomas but also in the number of cases with a thickness greater than 1 mm (an increase of 4%). The total surgery-related waiting time, commencing with

the positive SLB report and ending with lymphadenectomy, increased by about 15% (from 72 days in 2004 to 82.5 days in 2008). The implementation of a modified EORTC protocol in Tuscany as the standard procedure for extensive pathological handling of SLB [56] presumably had an impact from 2004 to 2008, lengthening the time of reporting.

Conclusion

In this study, we analyzed a set of 13 quality indicators for evaluation of the diagnosis, surgical treatment and pathological definition of melanoma. These indicators, evaluated using a population-based dataset in Tuscany, showed an increase in the number of thin melanomas (a mark of early diagnosis), an improvement in the completeness of the pathology reports, and good surgical management. Some of the latter improvements may be related to the development of ITT recommendations for melanomas for the regional healthcare system in 2007, a product of the regional multidisciplinary Melanoma Group. These as well as other indicators, such as the mitotic rate [54,57], should be monitored to identify the potential improvements and shortcomings in melanoma care.

Acknowledgements

The authors acknowledge the Tumour Institute of Tuscany (ITT) melanoma Working Group: P. Apicella (SS. Cosma e Damiano Hospital, Pistoia), G. Bagnoni, †M. Ceccarini, R. Incensati, R. Sarnelli, G. Talini, V. Zucchi (Ospedale Civile, Livorno), P. Barachini, G. Bevilacqua, M.C. Cagno, G. Cervadoro, G. Cristiani, G. Fornaciari, B. Loggini, G. Marchetti, F. Minnei, A.G. Naccarato, V. Nardini, R. Pingitore, S. Pistolesi, A. Pucci, A. Romanini, A. Tognetti (S. Chiara – Azienda Ospedaliero-Universitaria Pisana, Pisa), S. Battistini, A. Cavazzana (Presidio Ospedaliero Apuane Massa e Carrara, Carrara), B. Begliomini, A. Giannini, S. Ketabchi, G. Lo Scocco, C. Salvini (Misericordia e Dolce Hospital, Prato), M. Biagioli, C. D'Aniello, M.C. De Nisi, M. Fimiani, M. Maio, C. Miracco, P. Tosi (S. Maria alle Scotte Universitary Hospital, Siena), M. Biancalani, L. Brandini, A. Calcinai (S. Giuseppe General Hospital, Empoli), L. Borgognoni, P. Brandani, L. Fioretto, G. Gerlini, A. Ribecco, S. Sestini, C. Urso (S.M. Annunziata Hospital, Florence), A. Caldarella, A. Chiarugi, E. Crocetti, L. Miligi, P. Nardini, E. Paci, D. Palli, M. Zappa (Institute for Cancer Study and Prevention, Florence), A. Carnevali, A. Castelli, D. Floridi, F. Mirri (S. Donato Hospital, Arezzo), R. Cecchi, S. Innocenti (Ospedale del Ceppo, Pistoia), S. Crispino (Local Health Authority 7, Siena), V. De Giorgi, N. Pimpinelli, (Azienda Ospedaliero Universitaria Careggi – Local Health Authority 10, Florence), M. Dini, P. Fabbri, O. Gallo, R. Gattai, F. Gensini, M. Genuardi, M. Innocenti, G. Mariotti, D. Massi, S. Moretti, M. Santucci (Azienda Ospedaliero Universitaria Careggi, Florence), G. Fanelli, M. Pantaloni (Azienda Ospedaliero-Universitaria Pisana, Pisa), G. Gianfranceschi, F. Marsili, L. Martini, P. Viacava (Local Health Authority 12, Viareggio Lido di Camaiore), A. Ghilardi, S. Lorenzi, P. Martini, P. Massei, E. Menni (Campo di Marte Hospital, Lucca), M.A. Pata, E. Reali (Local Health Authority 10, Florence), E. Pattarino (General Practice, Florence), L. Presenti, F. Zolfanelli (S. Giovanni di Dio Hospital, Florence), A. Rossi, R. Sirna (Misericordia Hospital, Grosseto), C. Sabò (Campo di Marte Hospital, Lucca), B. Stecca (Istituto Toscano Tumori, Florence).

The authors acknowledge the support of Ente Cassa di Risparmio di Firenze.

Conflicts of interest

There are no conflicts of interest.

References

1 Airtum working group. Italian cancer figures, report 2009. Cancer trend (1998–2005). *Epidemiol Prev* 2009; **33**:54–55.
2 ITT Melanoma WG. *Clinical recommendations for skin melanoma*. Florence (Italy), Tuscany Tumor Institute – ITT: Scientific Press; 2007.
3 Crocetti E, Caldarella A, Chiarugi A, Nardini P, Zappa M. The thickness of melanomas has decreased in central Italy, but only for thin melanomas, while thick melanomas are as thick as in the past. *Melanoma Res* 2010; **20**:422–426.
4 Caldarella A, Amunni G, Angiolini C, Crocetti E, Di Costanzo F, Di Leo A, *et al*. Feasibility of evaluating quality cancer care using registry data and electronic health records: a population-based study. *Int J Qual Health Care* 2012; **24**:411–418.
5 Mainz J. Defining and classifying clinical indicators for improvement. *Int J Qual Health Care* 2003; **15**:523–530.
6 Mainz J, Hansen AM, Palshof T, Bartels P. National quality measurement using clinical indicators: the Danish National Indicator Project. *J Surg Oncol* 2009; **99**:500–504.
7 Bilimoria KY, Raval MV, Bentrem DJ, Wayne JD, Balch CM, Ko CY. National assessment of melanoma care using formally developed quality indicators. *J Clin Oncol* 2009; **27**:5445–5551.
8 Paci E, Crocetti E, Miccinesi G, Benvenuti A, Intrieri T, Sacchettini C, Giovannetti L. Tuscany Cancer Registry. In: Curado MP, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M, Boyle P, editors. *Cancer incidence in five continents*. Lyon: IARC Scientific Publications2007; **Vol. IX**:p. 319.
9 Lilford LJ, Brown CA, Nicholl J. Use of process measures to monitor the quality of clinical practice. *BMJ* 2007; **335**:648–650.
10 Garbe C, Peris K, Hauschild A, Saiag P, Middleton M, Spatz A, *et al*. Diagnosis and treatment of melanoma: European consensus-based interdisciplinary guideline. *Eur J Cancer* 2010; **46**:270–283.
11 Garbe C, Peris K, Hauschild A, Saiag P, Middleton M, Spatz A, *et al*. Diagnosis and treatment of melanoma. European consensus-based interdisciplinary guideline – update 2012. *Eur J Cancer* 2012; **48**: 2375–2390.
12 Marsden JR, Newton-Bishop JA, Burrows L, Cook M, Corrie PG, Cox NH, *et al*. Revised UK guidelines for management of cutaneous melanoma 2010. *Br J Dermatol* 2010; **163**:238–256.
13 Italian medical oncologic association (AIOM). Guidelines: melanoma, 2010. Available at: *http://www.aiom.it/area+pubblica/area+medica/prodotti+scientifici/linee+guida/Archivio/1,333-333* [Accessed 15 April 2013].
14 Coit DG, Andtbacka R, Bichakjian CK, Dilawari RA, Dimaio D, Guild V, *et al*. NCCN melanoma panel. Melanoma. *J Natl Compr Canc Netw* 2009; **7**:250–275.
15 Dummer R, Guggenheim M, Arnold AW, Braun R, von Moor R. Updated Swiss guidelines for the treatment and follow up of cutaneous melanoma. *Swiss Med Wkly* 2011; **141**:w13320.
16 Dummer R, Hauschild A, Guggenheim M, Keilholz U, Pentheroudakis G. on behalf of the ESMO Guidelines Working Group. Cutaneous melanoma: ESMO clinical practice guidelines for diagnosis, treatment and follow up. *Ann Oncol* 2012; **7 (Suppl 1)**:vii86–vii91.

17 Morton DL, Thompson JF, Cochran AJ, Mozzillo N, Elashoff R, Essner R, *et al*. Sentinel-node biopsy or nodal observation in melanoma. *N Engl J Med* 2006; **355**:1307–1317.
18 Aitken JF, Janda M, Elwod M, Youl PH, Ring IT, Lowe JB. Clinical outcomes from skin screening clinics within a community-based melanoma screening program. *J Am Acad Dermatol* 2006; **54**:105–114.
19 Carli P, De Giorgi V, Betti R, Vergani R, Catricala C, Mariani G, *et al*. Relationship between cause of referral and diagnostic outcome in pigmented lesion clinics aimed to melanoma screening: a multicenter survey of the Italian Multidisciplinary group on Melanoma (GIPMe). *Melanoma Res* 2003; **13**:207–211.
20 Carli P, De Giorgi V, Giannotti B. Dermoscopy and early diagnosis of melanoma. The light and the dark. *Arch Dermatol* 2001; **137**:1641–1644.
21 Schiffner R, Schiffner-Rohe J, Landthaler M, Stolz W. Long-term dermoscopic follow-up of melanocyitc naevi: clinical outcome and patient compliance. *Br J Dermatol* 2003; **149**:79–86.
22 Wagner JD, Gordon MS, Chuang TY, Coleman JJ. Current therapy of cutaneous melanoma. *Plast Reconstr Surg* 2000; **105**:1774–1799.
23 Lens M. Current clinical overview of cutaneous melanoma. *Br J Nurs* 2008; **17**:300–305.
24 Garbe C, Eigentler TK. Diagnosis and treatment of cutaneous melanoma: state of the art 2006. *Melanoma Res* 2007; **17**:117–127.
25 Erickson JL, Velasco JM, Hieken TJ. Compliance with melanoma treatment guidelines in a community teaching hospital: time trends and other variables. *Ann Surg Oncol* 2008; **15**:1211–1217.
26 Blazer DG 3rd, Sondak VK, Sabel MS. Surgical therapy of cutaneous melanoma. *Semin Oncol* 2007; **34**:270–280.
27 Riker AI, Kirksey L, Thompson L, Morris A, Cruse CW. Current surgical management of melanoma. *Expert Rev Anticancer Ther* 2006; **6**: 1569–1583.
28 Lavie A, Desouches C, Casanova D, Bardot J, Grob JJ, Legré R, *et al*. Surgical management of cutaneous malignant melanoma. Review. *Ann Chir Plast Esthet* 2007; **52**:1–13.
29 Eddy DJ. Surgical treatment of melanoma. *Br J Dermatol* 2003; **149**:2–12.
30 Wright F, Spithoff K, Easson A, Murray C, Toye J, McCready D, *et al*. Melanoma Disease Site Group of Cancer Care Ontario's Program in Evidence-based Care. Primary excision margins and sentinel lymph node biopsy in clinically node-negative melanoma of the trunk or extremities. *Clin Oncol* 2011; **23**:572–578.
31 Mocellin S, Pasquali S, Nitti D. The impact of surgery on survival of patients with cutaneous melanoma: revisiting the role of primary tumor excision margins. *Ann Surg* 2011; **353**:238–243.
32 Sladden MJ, Balch C, Barzilai DA, Berg D, Freiman A, Handiside T, *et al*. Surgical excision margins for primary cutaneous melanoma. *Cochrane Database Syst Rev* 2009;4.
33 Lens MB, Nathlan P, Bataille V. Excision margins for primary cutaneous melanoma: updated pooled analysis of randomized controlled trials. *Arch Surg* 2007; **142**:885–891.
34 Haigh PI, DiFonzo LA, McCready DR. Optimal excision margins for primary cutaneous melanoma: a systematic review and meta-analysis. *Can J Surg* 2003; **46**:419–426.
35 Borgognoni L, Urso C, Vaggelli L, Brandani P, Gerlini G, Reali UM. Sentinel node biopsy procedures with an analysis of recurrence patterns and prognosis in melanoma patients: technical advantages using computer-assisted gamma probe with adjustable collimation. *Melanoma Res* 2004; **14**:311–319.
36 Amersi F, Morton DL. The role of sentinel lymph node biopsy in the management of melanoma. *Adv Surg* 2007; **41**:241–256.
37 Sondak VK, Zager JS, Messina JL. Sentinel lymph node biopsy as the standard of care for cutaneous melanoma: pro. *Clin Adv Hemat Oncol* 2007; **5**:483–486.
38 Cook MG, Green MA, Anderson B, Eggermont AM, Ruiter DJ, Spatz A, *et al*. The development of optimal pathological assessment of sentinel lymph nodes for melanoma. *J Pathol* 2003; **200**:314–319.
39 Cook MG, Di Palma S. Pathology of sentinel lymph nodes for melanoma. *J Clin Pathol* 2008; **61**:897–902.
40 Chakera AH, Hesse B, Burak Z, Ballinger JR, Britten A, Caracò C, *et al*. EANM-EORTC general recommendations for sentinel node diagnostics in melanoma. *Eur J Nucl Med Mol Imaging* 2009; **36**:1713–1742.
41 Lens MB, Dawes M, Goodacre T, Newton-Bishop JA. Elective lymph node dissection in patients with melanoma: systematic review and meta-analysis of randomized controlled trials. *Arch Surg* 2002; **137**:458–461.
42 Greene FL, Page DL, Fleming ID, Fritz AG, Balch CM, Haller DG, Morrow M*AJCC cancer staging manual*. 6th ed. Chicago: Springer; 2002.
43 Cormier JN, Xing Y, Ding M, Lee JE, Mansfield PF, Gershenwald JE, *et al*. Population-based assessment of surgical treatment trends for patients with melanoma in the era of sentinel lymph node biopsy. *J Clin Oncol* 2005; **23**:6054–6062.
44 McMasters KM, Noyes RD, Reintgen DS, Goydos JS, Beitsch PD, Davidson BS, *et al*. Lessons learned from the Sunbelt Melanoma Trial. *J Surg Oncol* 2004; **86**:212–223.
45 Reintgen DS, Jakub JW, Pendas S, Swor G, Giuliano R, Shivers S. The staging of malignant melanoma and the Florida Melanoma Trial. *Ann Surg Oncol* 2004; **11**:186S–191S.
46 Bilimoria KY, Balch CM, Bentrem DJ, Talamonti MS, Ko CY, Lange JR, *et al*. Complete lymph node dissection for sentinel node-positive melanoma: assessment of practice patterns in the United States. *Ann Surg Oncol* 2008; **15**:1566–1576.
47 Carli P, De Giorgi V, Chiarugi A, Nardini P, Weinstock MA, Crocetti E, *et al*. Addition of dermoscopy to conventional naked-eye examination in melanoma screening: a randomized study. *J Am Acad Dermatol* 2004; **50**:683–689.
48 Chellini E, Crocetti E, Carli P, Martini A, Giovannetti L. The melanoma epidemic debate: some evidence for a real phenomenon from Tuscany, Italy. *Melanoma Res* 2007; **17**:129–130.
49 Eggermont AM. Advances in systemic treatment of melanoma. *Ann Oncol* 2010; **21**:vii339–vii344.
50 Hodi FS, O'Day SJ, Mcdermott DF, Weber RW, Sosman JA, Haanen JB, *et al*. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med* 2010; **363**:711–723.
51 Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, Larkin J, *et al*. Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N Engl J Med* 2011; **364**:2507–2516.
52 Frangos JE, Duncan LM, Piris A, Nazarian RM, Mihm MC Jr, Hoang MP, *et al*. Increased diagnosis of thin superficial spreading melanomas: a 20-year study. *J Am Acad Dermatol* 2012; **67**:387–394.
53 Welch HC, Black WC. Overdiagnosis in cancer. *J Natl Cancer Inst* 2010; **102**:605–613.
54 Balch CM, Gershenwald E, Soong SJ, Thompson JF, Atkins MB, Byrd DR, *et al*. Final version of 2009 AJCC melanoma staging and classification. *J Clin Oncol* 2009; **27**:6199–6206.
55 Testori A, De Salvo GL, Montesco MC, Trifirò G, Mocellin S, Landi G, *et al*. Clinical considerations on sentinel node biopsy in melanoma from an Italian multicentric study on 1313 patents (SOLISM_IMI). *Ann Surg Oncol* 2009; **16**:2018–2027.
56 Cook MG, Green MA, Anderson B, Eggermont AM, Ruiter DJ, Spatz A, *et al*. The development of optimal pathological assessment of sentinel lymph nodes for melanoma. *J Pathol* 2003; **200**:314–319.
57 Thompson JF, Soong SJ, Balch CM, Gershenwald JE, Ding S, Coit DG, *et al*. Prognostic significance of mitotic rate in localized primary cutaneous melanoma: an analysis of patients in the multi-institutional American Joint Committee on Cancer melanoma staging database. *J Clin Oncol* 2011; **29**:2199–2205.

Indicators of the standard of care for melanoma: Tuscany data

Author: Emanuele Crocetti, Adele Caldarella, Daniela Massi, et al
Publication: Melanoma Research
Publisher: Wolters Kluwer Health, Inc.
Date: Aug 1, 2013