Author’s response to reviews

Title: Characteristics of people living in Italy after a cancer diagnosis in 2010 and projections to 2020

Authors:

Stefano Guzzinati (stefano.guzzinati@regione.veneto.it)
Saverio Virdone (svirdone@cro.it)
Roberta De Angelis (roberta.deangelis@iss.it)
Chiara Panato (chiara.panato@cro.it)
Carlotta Buzzoni (c.buzzoni@ispo.toscana.it)
Riccardo Capocaccia (Riccardo.Capocaccia@istitutotumori.mi.it)
Silvia Francisci (silvia.francisci@iss.it)
Anna Gigli (anna.gigli@irpps.cnr.it)
Manuel Zorzi (manuel.zorzi@regione.veneto.it)
Giovanna Tagliabue (Giovanna.Tagliabue@istitutotumori.mi.it)
Diego Serraino (serrainod@cro.it)
Fabio Falcini (fabio.falcini@irst.emr.it)
Claudia Casella (claudia.casella@hsanmartino.it)
Antonio Russo (agrusso@ats-milano.it)
Fabrizio Stracci (fabs@unipg.it)
Bianca Caruso (rtnasentina@ausl.mo.it)
Maria Michiara (michiara@ao.pr.it)
Anna Caiazzo (a.caiazzo@aslsalerno.it)
Marine Castaing (marinecastaing@hotmail.com)
Stefano Ferretti (stefano.ferretti@unife.it)
Dear Prof. Amy Downing,

Thank you for the opportunity to improve our manuscript, and to the Reviewers for the appreciation of our work. We have now completed the revision, taking into consideration all their valuable comments, and included a point-by-point response. Please note, that the Abstract has been expanded and all revisions to the text have been made in red prints for your convenience.

Here enclosed is our revised article, we are resubmitting for a second evaluation of publication in the BMC Cancer journal.

The article has neither been published nor is under consideration of publication elsewhere. All authors have seen and approved the final version for this submission. The Authors declared no conflicts of interests. We still have to point out that our colleague Adriano Giacomin, contributing to all phases of the study, has recently passed away.

All correspondence regarding this submission should be addressed to: Dr. Luigino Dal Maso, Cancer Epidemiology Unit, CRO Aviano National Cancer Institute IRCCS, Via Franco Gallini 2, 33081 Aviano, Italy. Tel: +39-0434-659354; E-mail: epidemiology@cro.it

or

Stefano Guzzinati, Veneto Tumor Registry, Veneto Region, Padova, Passaggio Gaudenzio 1, 35131 Padova, Italy, Tel (+39) 049 877 8130; fax: (+39) 049 877 8147. E-mail: registro.tumori@regione.veneto.it

Thank you for the attention and best regards,

Sincerely,

Dr. Luigino Dal Maso
(on behalf of coauthors)

Point-by-point response to Reviewers’ reports:
Samantha Mason (Reviewer 1): General: A really interesting piece of work which provides foresight that long term care of cancer patients is likely to need to change as a result of increasing numbers of people living with and beyond a diagnosis in 2020.

RE: Thank you for appreciating of our work.

Mainly a few minor issues with sentence structure, and potentially some additions to tables to improve the clarity of what is being presented.

Background:

Line 119: "whereas and" could just say "whereas complete prevalence"

RE: Yes, the sentence has been corrected.

Methods:

Line 140: Not sure why the word contributing is included.

RE: the word ‘contributing’ has been deleted.

Line 145 to 148: I think there should be two separate sentences; one about what tumours, cancer types and combinations were included, and one that says which were excluded. Including the exclusion of non-malignant skin cancers in line 145 is a bit confusing as it initially reads like non-malignant tumours and 34 cancer types or their combinations were not included.

RE. The sentence has been rephrased. Please, see text.

Results:

Reporting of figures should be consistent. Line 189 reports that the percentage of men with colorectal cancer is 15.5% (as in the table) but the percentage of women with thyroid cancer (line 194) is reported in the text as 6%, despite being reported as 6.5% in table 2.

RE. We have corrected the inconsistencies throughout the paragraph.

Line 202: Sentence "Conversely, 59% of cervical cancer patients had their diagnosis ≥15 years before, they were 35% for stomach cancer, 31% for endometrial, but only 4% for prostate and 13% for lung cancer patients" could be rewritten as it is not straightforward to understand. Suggest "Conversely, the percentage of prevalent cases diagnosed ≥ 15 years before was 59% for cervical cancer, 35% for stomach cancer and 31% for endometrial cancer, but only 4% for prostate and 13% for lung cancer patients"?

RE. The sentence has been changed, as required.
Line 209: Should the sentence should read, "after a prostate cancer diagnosis"?

RE. Yes, corrected as suggested.

Line 211: The figure should be 212,863 rather than 212.863.

RE. Yes, corrected.

Line 216: I can't find anything in table 4 that references 22%. I assumed this was 22% of the Italian population in 2020? There is a brief implication in the method that population figures were forecast as well as prevalence figures. It may be worth including population figures for 2020 in a table so that they can be referred to. If the 22% refers to something else, the sentence should make this clear.

RE. The sentence has been clarified adding the information presented in the last column of Table 4 (21,657/100,000). The same specification has been made two rows down.

Discussion:

The discussion is very well presented and easy to follow.

Line 233: Should this say, "long-term survivors diagnosed ≥ 20 years before"?

RE. Yes, the sentence has been changed.

Line 234: The figure should be 519,356 rather than 519.356.

RE. Yes, corrected as suggested.

Tables: You should recheck the figures. In tables 1, 2 and 4 I found at least one instance where the sum of all the prevalent cases for the age groups was not the same as the total for all ages for a cancer type.

RE. Yes, thank you for the checks. Numbers and proportion in the tables, for the age groups and overall were rounded. We have checked that sum of the total (e.g., Table 1, total men 1,194,033) may differ from the sum of cases by age group only because of rounding (e.g., 1,194,034=4,844+84,172+87,091+198,505 + 363,932 + 357,051 + 98,439). We hope no other discrepancy has remained in the tables.

S Mallone (Reviewer 2): Study overview

Title
"Characteristics of people living in Italy after a cancer diagnosis in 2010 and projections to 2020"

Background

Improved survival and population ageing imply a progressive increase in tumour prevalence. Tumour prevalence regardless the time since diagnosis -i.e. complete prevalence- is less frequently estimated than limited duration prevalence. Estimates and projections of complete tumour prevalence are necessary to help clinicians and health care planners in improving long-term care of patients and appropriately allocating health care resources.

Please, revise the quality of the text.

RE. Yes, all the Background section has been revised. Please, see text.

In particular:

row 116 'previous five years' with respect to?

RE. Cancer diagnosis, as stated, we think in the sentence ... including only patients living after a cancer diagnosis made in the previous five years.

row 119 'wherease and', is it 'wherease'?

RE. Yes, corrected.

rows 120-123 'this indicator' does the sentence refer to complete prevalence? It's not clear

RE. 'this indicator' was used to avoid repetition. The sentence has been changed, for clarity.

The aims reported are:

* "to provide an updated description of the number of people living in Italy in 2010 (please clarify updated with respect to…; in 2010 or at 1st january 2010?) after a cancer diagnosis, for all cancers combined and for a selection of cancer types by sex, age, and time since diagnosis".

RE. The sentence has been clarified, as required.

* "projections of cancer prevalence in Italy are presented up to the year 2020".

Methods
Please reshape these paragraphs in order to let the reader better understand the method/data used and its complexity. A possible example could be the following

Study design and data sources

This is a descriptive analysis of individual data collected during the period 1976-2009 by Italian population-based cancer registries (CR) identified, according to the Italian legislation, as collectors of personal data for surveillance purpose without explicit individual consent. The approval of a research ethic committee is for this reason not required.

Case ascertainment and follow-up of vital status provided by the CRs (is missing in the text). Please add a very brief description or a proper reference.

RE. The first paragraph has been completely rephrased.

All tumours except non melanoma skin cancer (more than one primary tumour are considered) and a selected list of 34 tumour/combination of tumours (only the 1st tumour is considered) defined on the basis of the ICD-10 classification, are included.

Benign-uncertain-in situ urinary bladder cancers are included also.

The ICD-O-3 morphology codes were used to define specific subtypes. Please briefly explain which subtypes.

RE. The paragraph has been revised, see also answers to Reviewer 1.

Statistical methods

The clinical and demographic characteristics of the persons registered with a diagnosis of cancers in the Italian CRs are used to:

1. estimate how many of them are still alive at 1st January 2010 regardless the time since diagnosis -i.e. complete prevalence count- by type of tumour, sex and age group;
2. estimate the prevalence proportion in Italy at 2010 for each type/group of tumour, by sex, age;
3. obtain estimates of the complete prevalence (count and proportion) at 1st January 2015 and 2020;
4. disentangle the complete prevalence estimates (count and proportion) by time since diagnosis and describe the changing over time of these estimates.
RE. Thanks, a paragraph has been added at the beginning of the Statistical Methods section to provide the reader a better understanding of the method/data used and its complexity.

1. Complete prevalence count is calculated on the basis of population-based cancer registries (CRs) data through the completeness index method

* the maximum observed prevalence (7- to 34-year prevalence) is calculated for each of the 27 CRs with at least 7 years of cancer registration as of 31/12/2009 and follow-up of vital status of December 31, 2013; with seerstat?

RE. Yes, as stated in the first paragraph of Statistical Methods.

* a set of completeness indices by tumour type, sex, age group, time since diagnosis (7- to 34- as for the maximum observed prevalence?) are obtained by statistical modelling incidence/survival data drawn from 8 population-based Italian CRs with at least 18 years of cancer registration before 2010 (18,19); with COMPREV (17)

* the above indices are used as correction factors of the maximum observed prevalence in each CR (17);

2. For each type of tumour, the number of prevalent cases in Italy at 2010 (at 1st January 2010?) are obtained by multiplying the corrected prevalence proportion, stratified by sex and age, by the corresponding Italian population (at 1st January 2010?). What about the adjustment of the prevalence proportion?

RE. We expanded the last sentence of Statistical Methods section to clarify the issue.

Completeness Indices - Statistical modelling of CRs data

Relative survival and incidence functions were estimated by means of parametric models within the period 1985-2011 for survival and 1985-2009 for incidence.

The survival model was a parametric cure model assuming that a proportion of individuals with cancer were bound to die (fatal cases) with a survival following a Weibull distribution, while the remaining proportion (cured fraction) had the same mortality rate as that of the general population with the same age and gender stratification (14,20). The parameters of the survival model were estimated by cancer type, sex, and age class (0-14, 15-44, 45-54, 55-64, 65-74, 75+ years). A period effect was included on the hazard of dying of cancer. Incidence data were categorised according to cancer type, sex, five-year age group, and birth cohort (<1899, 1900-1904,…, 2005-2009). With SAS? Please, add a ref.

RE. Yes, we already described that we used a SAS procedure for incidence modelling. We have added the specification that SAS was also used for survival estimates according to the model described. Ref 21 provides additional details of all assumptions, methods, and software used.
A sixth degree polynomial age-cohort model of crude incidence rates was fitted through the SAS logistic procedure for each cancer type and sex (21). Refs. 20 and 21 are different as regards the CRs data/period used. Please clarify.

RE. Ref. 21, cited here, describes in detail the method used also in the present paper, and in some other studies. CR data and period used were updated in the present study (see Appendix 1), including more years of observation and follow-up, in comparison with ref. 21 and with previous studies on the same topic.

3. and 4. Projections at 2015, 2020 are obtained assuming that complete prevalence proportions follow a linear function based on the trend estimated for the last 3 calendar years (2007-2009). Please briefly explain how this function/trend is estimated or add a proper reference.

RE. A sentence has been added to clarify the alternative possibilities explored (e.g., log-linear models) and two references have also been added.

Results

In general, I would suggest the authors to put more emphasis on the projections at 2020 rather than on the 2010 estimates and to divide the paragraph as follows.

RE. We have appreciated the suggestion. Two headings for distinct paragraphs have been added. We have provided more details (i.e., more cancer types and age categories) for 2010 estimates than for 2020 projections, since they were based on more ‘robust’ data. In other words, 2010 estimates were largely based on observation, while 2020 projections were largely based on statistical models and assumption. Nevertheless, we are aware of the possible impact on public health of projections.

Prevalence estimates at 2010.

In Italy, in 2010, there are more than 2.6 million of persons (4.6% of the Italian population) alive with a previous diagnosis of tumour (appendix 2, tables 1, 2).

Prevalence proportions increase with age depending on sex: 4%, 8%, 11% and 2%, 6%, 13% at age 45-54, 55-64, 65-74 for females and males respectively (table 2). The most frequent tumours in terms of prevalence are prostate (25% of prevalent cases at 1st Jan 2010), bladder (16%), colorectal (16%) tumours for males and breast (42%), colorectal (12%), and endometrial (7%) tumours for females (tables 1, 2).

RE. We tried to simplify the Results section, as required.

Overall, 59% and 21% (60% and 20% in the paper) of the persons alive at 2010 in Italy with a tumour received a diagnosis at least 5 and 15 years before respectively (figure 1). Prevalence distribution at 2010 by time since diagnosis depends on tumour type: 7% and 39%
persons are still alive with a diagnosis of cervical and lung tumours received 2 years before. Considering the same type of tumour, the percentage of persons diagnosed at least 15 years before are 59% and 13% respectively (figure 1).

Please explain, why these prevalence estimates are a bit different respect to the ones reported in ref. 21.

RE. We have corrected 21% while 60% is correct as is (the sum of proportions for periods >5 is 59% due to rounding). We have also simplified the sentence.

Prevalence projections at 2020.

The number of persons alive with a tumour diagnosis will increase by 37% in 2020, in Italy (from 2.6 to 3.6 million of persons).

RE. Numbers were maintained at the beginning, to emphasize absolute amount of prevalence in Italy.

Female breast (0.8 millions of persons) and prostate (0.5 millions of persons) tumours will be the most frequent type of cancer in terms of prevalence. The largest increases are foreseen for prostate (+85%) and for thyroid cancers (+79%); the smallest for stomach, ovary, larynx in females. Stomach in females (-10%) and cervical tumours (-13%) show a decline in prevalence. Please try to discuss this result.

RE. The comment on this result was already present in the Discussion, and it has now been expanded (rows 257-261).

22% and 0.7% of the population aged ≥75 and <45 respectively will have had a previous cancer diagnosis (table 4), being thyroid tumour the most frequent in terms of prevalence.

For all tumours, the largest changing in prevalence proportions from 2006 to 2020 will be observed for persons diagnosed with cancer at least 20 years before the calendar year considered (Figure 2).

RE. Some of the suggestions have been included even if a detailed version of the Results section has been maintained since also Reviewer 1 asked for some additional clarification about the presented numbers.

Discussion

Please, reshape this paragraph discussing -very briefly- the 2010 estimates then the projections/variation over time; clarify the reasons why the 2010 results are different from the ones showed in ref. 21.
RE: A brief sentence has been added to clarify the issue at the beginning of Strengths and Weaknesses paragraph. To an international reader, we believed, that it is reasonable to assume that more recent estimates came from more updated information (available comparing Appendix 1 and ref. 21).

Row 237: '...among prevalent cancer cases at? at 2010…'

RE: The required specification has been added.

Row 265: we adjusted estimated proportions in cancer registry areas for the age distribution of the whole Italian population. Please clarify and add this information in the paragraph 'Statistical methods'.

RE: A sentence has been added at the end of Statistical methods to clarify the issue.

Row 280-284 '…the hypothesis that complete prevalence at 2020 can be predicted by a linear function of calendar year as regressor variable is supported by empirical evidence, consistently showing an approximate linear trend in recent years [5,8,21]. Please clarify the way refs. 5,8 support this issues. Ref.21?

RE. By observing Figure 1, obtained from published results by De Moor (ref.5, left) and by Hermann (ref. 6, right), one can see an approximate medium-term linear trend, at least for all cancer types combined, but also, exploring previous studies from US, for more frequent cancer types.

Maddams [4] provided long term projections (to 2040) and deeply discussed effect on estimates of different scenarios for incidence and survival. In particular, "Implausibly high projections of prostate cancer incidence were obtained using the log linear regression model…’ The long term projection, however, are beyond the scope of present study.

A sentence has been added and the ref. 8 deleted (irrelevant here, if not compared with previous papers of the same group, Colonna EJC 2008).

Notably, the use of a longer period (5 calendar years) to estimate linear slope did not materially modify the estimates (ref?).

RE. We have a set of validation tables on this issue, available upon request. Ref could be a Personal communication but we assume that, if interested, a reader can ask us clarification or additional details.