# Cause of death for patients with breast cancer: discordance between death certificates and medical files, and impact on survival estimates

Hava Izci[1*] , Tim Tambuyzer[2], Jessica Vandeven[2], Jérôme Xicluna[2], Hans Wildiers[1,3], Kevin Punie[1,3], Nynke Willers[4], Eva Oldenburger[1,5], Els Van Nieuwenhuysen[1,6], Patrick Berteloot[6], Ann Smeets[1,7], Ines Nevelsteen[1,7], Anne Deblander[6], Harlinde De Schutter[2], Patrick Neven[1,4], Geert Silversmit[2†] and Freija Verdoodt[2†]

## Abstract

**Background:** Registration and coding of cause of death is prone to error since determining the exact underlying condition leading directly to death is challenging. In this study, causes of death from the death certificates were compared to patients' medical files interpreted by experts at University Hospitals Leuven (UHL), to assess concordance between sources and its impact on cancer survival assessment.

**Methods:** Breast cancer patients treated at UHL (2009–2014) (follow-up until December 31st 2016) were included in this study. Cause of death was obtained from death certificates and expert-reviewed medical files at UHL. Agreement was calculated using Cohen's kappa coefficient. Cause-specific survival (CSS) was calculated using the Kaplan-Meier method and the relative survival probability (RS) using the Ederer II and Pohar Perme method.

**Results:** A total of 2862 patients, of whom 354 died, were included. We found an agreement of 84.7% (kappa-value of 0.69 (95% C.I.: 0.62–0.77)) between death certificates and medical files. Death certificates had 10.7% false positive and 4.5% false negative rates. However, five-year CSS and RS measures were comparable for both sources.

**Conclusion:** For breast cancer patients included in our study, fair agreement of cause of death was seen between death certificates and medical files with similar CSS and RS estimations.

**Keywords:** Breast cancer, Cause of death, Death certificates, Misclassification, Cause-specific survival, Relative survival

## Background

Accurate information on the principal cause of death, resulting in reliable cause-specific survival estimates, could be important for dependable estimation of

disease-specific mortality, support treatment decision making and the allocation of health care resources.

In Belgium, the principal or underlying cause of death is derived from the death certificate, which is completed by a physician and captured in a coding system i.e., the Tenth Revision of the International Classification of Diseases (ICD-10), by registrars [1]. Accuracy of the principal cause of death in death certificates compared to a second, verified clinical source has been investigated in other countries, for patients with breast cancer [2–6]

* Correspondence: hava.izci@kuleuven.be
†Geert Silversmit and Freija Verdoodt contributed equally to this work.
1Department of Oncology, KU Leuven - University of Leuven, Herestraat 49 box 7003-06, B-3000 Leuven, Belgium
Full list of author information is available at the end of the article

and other cancers [7–16]. Due to discordant reporting of cause of death, some of these studies recommend improved reporting of cause of death classification [17–20]. Since breast cancer has a favorable prognosis compared to some other cancers types, reporting the cause of death may be more difficult, particularly in patients with comorbidities. In Belgium, an in-depth review of the principal causes of death with an investigation of different survival measures has not been done yet.

Both relative survival (RS) and cause-specific survival (CSS) are approaches applied to estimate net survival of cancer patients. Cancer registries often use RS, as it does not need cause of death information, which is not always available at the population level. Both methods have weaknesses that could introduce bias in the net survival estimation. CSS requires accurate cause of death information from death certificates, whereas RS requires a disease-free reference group. Some studies have recommended the use of RS over CSS in breast cancer patients, because it is less susceptible to misclassification errors [21–27]. In this study, we compare the principal cause of death between death certificates and expert-reviewed medical files in a cohort of breast cancer patients at University Hospitals Leuven (UHL), a large-scale tertiary hospital with a specialized breast center in Belgium. Additionally, we explore its impact on CSS, and compare this with relative survival-based approaches to estimate net survival.

## Methods

### Study population

All female patients with a first invasive breast cancer diagnosis between January 1st, 2009 and December 31st, 2014, and treated in UHL were included in this study. Tumor and patient characteristics such as age, year of diagnosis, histological grade and tumor stage (TNM classification system, 6th and 7th edition) were obtained at the Belgian Cancer Registry (BCR) [28].

### Data sources

In order to calculate and compare the cause-specific survival (CSS) and relative survival (RS), different data sources were used. The vital status of patients, available up to December 31st, 2016, was obtained from linkage with the Belgian Crossroads Bank for Social Security (CBSS) using the patients' social security number [29]. Patients with unknown vital status or deceased patients with missing cause of death information were excluded from analyses. For CSS, cause of death information was extracted from two sources: the death certificates obtained from the regional authorities in Flanders, Brussels and Wallonia, and the medical files obtained from UHL. For RS, the population life tables from Statistics Belgium were used [30].

When someone dies, cause of death and associated conditions are described in the death certificate by a certified physician. These death certificates are collected physically and electronically by regional authorities: the 'Agentschap Zorg en Gezondheid' for Flanders [31], 'Observatoire de la Santé et du Social de Bruxelles-Capitale' for Brussels [32], and 'Agence pour une Vie de Qualité' for Wallonia [33]. International coding and classification rules are applied to the certificates in accordance with ICD-10 [1]. Principal cause of death is derived from the chain of events that resulted in death. Coding principal cause of death is automatically done for a subset (40%) of death certificates using international coding software (IRIS software) [34, 35], while 49% percent are coded semi-automatically, and 11% are manually reviewed by an encoder based on the wording or phrasing of the physician to determine principal cause of death [31, 36].

The death certificate is designed to state the chain of events leading to death, thus, the immediate, intermediate, underlying and associated cause of death. The immediate cause is the cause that has led directly to the passing of the patient, which can be caused by or coincide with the intermediate and underlying cause of death. The associated causes of death are important contributing factors to the death. For instance, if a patient with breast cancer develops brain metastases, but dies from a brain hemorrhage, the immediate cause of death would be the hemorrhage, the intermediate cause the brain metastases and the underlying cause breast cancer. The latter would be defined as the principal cause of death. A possible associated cause of death could be, for example, pre-existing atherosclerosis of blood vessels in the brain.

Besides death certificates, information from medical files can be a second source of information about cause of death that can be used to calculate CSS. Information from expert-reviewed medical files was considered the gold standard in this study. A physician (P.N.) from UHL checked cause of death information twice for all deceased breast cancer patients from the available medical files. In case of a discordant or unknown cause of death, the case was flagged and an expert panel consisting of 7 physicians from UHL (H.W., K.P., N.W., E.O., E.V.N., P.B. and A.D.) performed a blinded verification of principal cause of death. If there was a discordance between the two, another member of the expert panel randomly performed a second blinded verification to select the principal cause of death. Finally, the principal cause of death was determined by majority decision.

All experts from the panel individually consulted the electronic medical files of patients that were assigned to them. When necessary, the experts additionally followed-up with the physician of the last patient

contact, and consulted E-health (i.e., a platform for protected health information exchange among health care providers) [37]. The principal cause of death was defined as 'breast cancer', if this disease initiated the chain of events leading up to death and if the patient did not have an accident or injury that resulted in death. Breast cancer can initiate the chain of events leading up to the death in case of presence of breast cancer metastases, irrespective of the survival period of the patient. Death by the treatment of breast cancer is not considered as death due to breast cancer.

### Definition of survival measures
Different survival measures were calculated, i.e., CSS and RS [38, 39]. CSS only considers deaths due to breast cancer as an event, with cause of death being obtained from either death certificates or medical files. RS is defined as the ratio of overall survival (with death of any cause as an event) from the breast cancer patient cohort and expected survival of a comparable cohort from the general population, matched on sex, age, diagnosis year and region. Net survival, which encompasses the survival that would be observed if the only possible cause of death was the cancer under study, can be estimated with CSS or RS-based approaches. Survival time for patients was calculated from the incidence date to date of death or until last known date alive. Follow-up in death certificates was available up to end of 2016.

### Statistical analyses
Agreement between principal cause of death from death certificates and medical files was investigated by calculating the Cohen's Kappa coefficient (κ-value) [40]. Concordance was investigated further by correlating κ-value with tumor and patient characteristics as age, diagnosis year, histological grade, tumor stage (combined pathological and clinical stage, TNM classification system) and tumor multiplicity. For all tumor and patient characteristics, subgroups were created. The κ-value was calculated for every subgroup separately. The Spearman's correlation coefficient (ρ) was then calculated to measure the association strength between the subgroup and κ-value (*p*-value cutoff at 0.05) [41]. Subgroups in which stage or grade were unknown were excluded from the subgroup calculation. All analyses were performed with SAS 9.4 (SAS Institute, Cary, NC, USA) within the SAS Enterprise Guide software (version 7.15 of the SAS System for Windows).

CSS was calculated based on the principal cause of death from death certificates and medical files. CSS considered the survival time from date of diagnosis until the date of death from breast cancer (outcome of interest), death due to other causes (censored) or until last known date alive (censored). CSS estimation were performed with the Kaplan-Meier method in SAS [42]. Next, RS was calculated and compared with CSS. RS was calculated by the Ederer II method in SAS and R (R Core Team, 2017) [43, 44], and the more recent Pohar Perme method in R [45]. The SAS code uses broad pre-specified time intervals in the actuarial approach (mostly 1-year broad intervals), whereas the R code uses data driven time intervals (at each event and censoring time).

## Results
A total of 2862 breast cancer patients of which 354 died and for which cause of death information was available in both data sources, were included in the analyses (Table 1). The median follow-up period was 54.6 months. Blinded review for principal cause of death was performed by the expert panel in 70 cases, of which 8 patients needed a second blind verification. Concordance in principal cause of death between both sources showed a 4.5% false negative proportion (*n* = 16), and 10.7% false positive proportion (*n* = 38) (Table 2). False negatives were patients who were misclassified as having died from another cause than breast cancer, and false positives were patients who were misclassified as having died from breast cancer. The κ-value was 0.69 (95% C.I.: 0.62–0.77) [46].

For false negatives, the most common cause of death in death certificates was primary cancer at the site of metastasis instead of breast cancer (*n* = 6 or 37.5%). None of the false negatives reported breast cancer in the listed immediate, intermediate, underlying or associated causes of death in the death certificate. Three out of 16 false negatives (18.8%) reported the ICD-10 code for ill-defined and unknown cause of death (ICD-10 code R99.0) as the principal cause of death. Other causes of misclassification were registration of another disease, another primary cancer or comorbidities from the patients' history reported as principal cause of death (*n* = 7 or 43.7%). Seventeen out of 38 false positives, had their principal cause of death from medical files (i.e., not breast cancer) listed as intermediate or immediate cause of death in their death certificates. Some of these patients died from an acute unrelated death (stroke or cardiac arrest), that was reported as death from breast cancer (*n* = 6) in the death certificate.

Next, κ-value was calculated according to subgroups (Table 3). The agreement of principal cause of death between both sources had a weak inverse correlation with increasing age, stage and diagnosis year (n.s., *p* > 0.05). The Spearman's correlation coefficients were – 0.7, – 0.8 and – 0.26 for increasing age, stage and diagnosis year respectively, thus correlation was lower in older age subgroups, higher stage and patients with a more recent year of diagnosis, however the *p*-value was not

**Table 1** General patient and tumor characteristics (*n* = 2,862)

| Category | Subgroups | N of patients | (%) |
|---|---|---|---|
| **Age at time of diagnosis** | <40 years | 188 | 6,6 |
| | 40-49 | 562 | 19,6 |
| | 50-59 | 778 | 27,2 |
| | 60-69 | 709 | 24,8 |
| | 70-79 | 423 | 14,8 |
| | ≥80 | 202 | 7,1 |
| **Year of diagnosis** | 2009 | 489 | 17,1 |
| | 2010 | 448 | 15,7 |
| | 2011 | 468 | 16,4 |
| | 2012 | 483 | 16,9 |
| | 2013 | 500 | 17,5 |
| | 2014 | 474 | 16,6 |
| **Stage at diagnosis** | Stage 0 | 47 | 1,6 |
| | Stage I | 1063 | 37,1 |
| | Stage II | 1174 | 41,0 |
| | Stage III | 357 | 12,5 |
| | Stage IV | 158 | 5,5 |
| | Stage unknown | 63 | 2,2 |
| **Tumor Grade** | Grade 1 | 319 | 11,1 |
| | Grade 2 | 1378 | 48,1 |
| | Grade 3 | 1126 | 39,3 |
| | Grade unknown | 39 | 1,4 |
| **Multiple tumors** | No | 2733 | 95,5 |
| | 2+ | 129 | 4,5 |
| **Treatment[a]** | Breast Conserving Surgery | 1374 | 48,0 |
| | Mastectomy | 1311 | 45,8 |
| | Radiotherapy | 2354 | 82,3 |
| | Chemotherapy | 1296 | 45,3 |
| | Hormonal therapy | 2344 | 81,9 |
| **Morphology** | Ductal | 2118 | 74,0 |
| | Lobular | 337 | 11,8 |
| | Other | 407 | 14,2 |

[a]Some patients receive multiple different treatment interventions within each subgroup (surgery or systemic therapy) during their follow-up period

**Table 2** Concordance/discordance table for the principle cause of death between medical files reviewed by board of experts (gold standard), and death certificates (*n* = 354 deaths)

| | | Medical files | | Total |
|---|---|---|---|---|
| | | Other Causes | Breast cancer | |
| Death certificates | Other causes | 136 (38.4%) | 16 (4.5%) | 152 (42.9%) |
| | Breast cancer | 38 (10.7%) | 164 (46.3%) | 202 (57.1%) |
| Total | | 174 (49.2%) | 180 (50.9%) | 354 (100%) |

Izci *et al. Archives of Public Health*        (2021) 79:111

Page 5 of 8

**Table 3** Agreement analyses (kappa statistic) for different patient subgroups (grouped based on patient and tumor characteristics), comparing principal cause of death from death certificates and medical files

| Category | Subgroups | Number of patients alive | Number of deceased patients (FU 31-12-2016) | Kappa statistic | Spearman's correlation coefficient and *p*-value |
|---|---|---|---|---|---|
| Age | <40 years | 177 | 11 | ND[a] | Rho= -0.7 |
| | 40-49 | 525 | 37 | 0.64 | *P*-value= 0.23 |
| | 50-59 | 736 | 42 | 0.83 | |
| | 60-69 | 649 | 60 | 0.73 | |
| | 70-79 | 334 | 89 | 0.54 | |
| | ≥80 | 87 | 115 | 0.50 | |
| Year of diagnosis | 2009 | 393 | 96 | 0.77 | Rho= -0.26 |
| | 2010 | 379 | 69 | 0.59 | *P*-value= 0.67 |
| | 2011 | 398 | 70 | 0.71 | |
| | 2012 | 429 | 54 | 0.67 | |
| | 2013 | 472 | 28 | 0.78 | |
| | 2014 | 437 | 37 | 0.54 | |
| Stage at diagnosis | Stage 0 | 46 | 1 | ND[a] | Rho= -0.8 |
| | Stage I | 1019 | 44 | 0.76 | *P*-value= 0.33 |
| | Stage II | 1053 | 121 | 0.69 | |
| | Stage III | 288 | 69 | 0.75 | |
| | Stage IV | 56 | 102 | 0.35 | |
| | Stage unknown[b] | 46 | 17 | 0.43 | |
| Grade of tumor | Grade 1 | 301 | 18 | ND[a] | ND[c] |
| | Grade 2 | 1223 | 155 | 0.58 | |
| | Grade 3 | 960 | 166 | 0.81 | |
| | Grade unknown[b] | 24 | 15 | 0.86 | |
| Multiple tumors | Yes | 94 | 35 | 0.66 | ND[c] |
| | No | 2414 | 319 | 0.69 | |

[a]Not defined due to events within subgroup being too small
[b] subgroups with unknown stage or grade were not included in Spearman analyses
[c]not defined due to subgroup without ranked variables

significant (n.s., *p* > 0.05). The agreement was classified as 'fair' in the subgroup with stage IV at diagnosis [47].

To investigate the impact of misclassification of cause of death on survival measurements, 5-year CSS was calculated based on both sources separately. CSS calculated from principal cause of death obtained from medical files resulted in slightly higher 5-year CSS estimates (93.1% (95% C.I.: 91.9–94.1)), compared to principal cause of death obtained from death certificates (92.3% (95% C.I.: 91.2–93.4)) (Table 4).

**Table 4** 5-year cause-specific survival (CSS) using primary cause of death information from medical files and death certificates (Follow-up until December 31st, 2016)

| | 5-year breast cancer specific survival, % (95% CI) | | |
|---|---|---|---|
| | Inclusion of known COD from medical files and death certificates (354 deaths) | Inclusion of known COD from death certificates (373 deaths) | Inclusion of known COD from medical files (402 deaths) |
| CSS (death certificates) | 92.3 (91.2, 93.4) (Events[a]: 183/2862) | 92.7 (91.6, 93.7) (Events[a]: 188/2882) | |
| CSS (medical files) | 93.1 (91.9, 94.1) (Events[a]: 162/2862) | | 93.1 (91.9, 94.1) (Events[a]: 165/2726) |

[a]Events are breast cancer-related deaths within 5 years after diagnosis

Izci *et al. Archives of Public Health*        (2021) 79:111

Page 6 of 8

**Table 5** 5-year net survival estimates (relative survival (RS) and cause-specific survival (CSS)) using different methods[a]

|  | 5-year net survival, % (95% CI) |
|---|---|
| RS - Pohar Perme (in R) | 93.2 (91.5, 94.8) |
| RS - Ederer II (in R) | 93.7 (92.3, 95.2) |
| RS - Ederer II (in SAS) | 92.8 (91.3, 94.2) |
| CSS (Medical files) (in SAS) | 93.1 (91.9, 94.1) |
| CSS (Death certificates) (in SAS) | 92.3 (91.2, 93.4) |

[a]Patients with COD available in both medical files and death certificates were selected (354 deaths). Years of diagnosis 2009-2014, follow-up of vital status up to December 31st, 2016

Finally, different net survival approaches were used in order to compare these estimates (Table 5). A small difference could be seen in survival estimates from RS calculated with Pohar Perme and Ederer II method and CSS as calculated with the Kaplan-Meier method.

## Discussion

This study evaluated accuracy of death certificates by validation of causes of death against a medical file review by a board of experts. Additionally, we investigated the impact of misclassification of cause of death on CSS. We found fair agreement between causes of death reported in death certificates and medical files, although this kappa-value interpretation has been defined slightly differently in publications over the years and should be interpreted relative to the setting [46–49]. Further, CSS with cause of death information obtained from medical files was slightly higher as a result of less deaths due to breast cancer, compared to survival using causes of death from death certificates, but was generally similar. Expert review was useful to identify and solve difficult cases where cause of death was unclear or difficult to determine.

First, we investigated discordant causes of death between death certificates and medical files. Among the false negatives (4.5% of cases), misattribution of breast cancer-specific death in death certificates was linked to the presence of comorbidities, metastases (from the primary breast cancer), or unspecified causes. We found more false positives (10.7% of cases) or over-reporting of breast cancer deaths than underreporting. Our results are consistent with literature for breast cancer that state more false positive cases of breast cancer-related deaths in comparison to false negatives [4, 5], although earlier studies from 1980s reported underestimation of breast cancer as principal cause of death in death certificates [2, 3].

Subsequently, we looked into trends of misclassification in specific subgroups based on patient and tumor characteristics. For age, diagnosis year and stage at diagnosis, Spearman's correlation coefficient could be calculated as a measure for the strength of relationship between the agreement factor and subgroups. Although not statistically significant, a weak inverse correlation was seen for age, stage and diagnosis year. A previous study in Geneva, Switzerland by Schaffar et al. [5] found more misclassification in older adults and patients with advanced disease. Older patients with cancer are more likely to have multiple comorbidities, which could lead to an increased risk of misclassifying the principal cause of death. Besides that, patients with metastases at diagnosis are more likely to have misclassification of cause of death since their site of metastases might be reported as the primary cause of death.

Several studies have investigated and validated the reporting and misclassification of causes of death in breast cancer patients, since the quality of death certification has been questioned [2–6]. Previous studies obtained discordance rates of 8.8% [5], 9.0% [2, 3] and 10.0% [6]. Our study showed a discordance rate of 15.2% between death certificates and expert-reviewed medical files, which was higher than previous studies. Coding of cause of death in death certificates according to international ICD-10 guidelines is semi-automatic, which helps to unify all codes with rule definitions, but details of the chain of causes of death can get lost in this coding system [36]. In addition, certification errors by the clinician responsible for assigning the causes of death in the death certificate, for example due to incomplete information, could lead to misclassification, as the IRIS system is dependent on the quality and information mentioned on the death certificates.

Identifying cause of death in medical files could be difficult in breast cancer patients with comorbidities, as it may be unclear if the patient has died from cancer, comorbidities or complications related to the cancer treatment. Breast cancer in particular is less lethal than some other cancers or comorbidities. This makes it more challenging to identify the cause of death correctly, since patients are more likely to die from non-cancer related causes. Other cancers with more lethal outcome, such as lung cancer, have shown higher overestimation of death due to cancer than breast cancers [16].

We validated death certificates by using an expert board that actively checked different data sources to evaluate medical history of the patient and designate an accurate principal cause of death. Review of medical files is routinely done for all patients in the Geneva cancer registry [5], since it is useful to have exact registration of causes of death for patients and obtain exact cause of death information. These specialized registrars are trained to carry out yearly follow-up of the registry with the aim to calculate CSS with this cause of death information. Unfortunately, a manual review of medical files is often not possible in the real-life setting, given labor intensity and costs. Guidelines for registrars and

Izci *et al. Archives of Public Health*          (2021) 79:111

Page 7 of 8

physicians have been developed according to ICD-10 in order to improve reliability of cause of death reporting. Periodic reviews of (a sample of) cause of death data and implementation of these guidelines would be beneficial in the future, as this could help to have more accurate disease-specific survival data and respond to epidemiological trends. When limited resources are available, such reviews could be restricted to patients with more discordance, such as patients with older age and higher disease stage.

Consequently, we wanted to see what the impact of misclassification of cause of death would be on survival. The survival results from these approaches were very similar. For breast cancer patients included in the study, RS measure that does not require cause of death information was comparable to CSS measures. A recent publication by Wissing et al. [17] recommends reporting and interpreting the CSS, RS and overall survival measure altogether to complement each other. Detailed description of the procedure and data sources to identify cause of death when reporting these measures is also recommended in the future.

The limitations of this study were that for a few cases, cause of death information was not available in any of the available medical files and could not be investigated further by consulting external sources. We did, however, have the chance to use extensive medical files from UHL, a large-scale tertiary Hospital in Belgium with a specialized breast center. This allowed strict adherence to guidelines and adequate clinical follow-up for patients. It would also be interesting to investigate the classification rate for causes of death in breast cancer patients in a secondary hospital in the future, to compare these results.

## Conclusions

For patients with breast cancer, we observed a fair agreement of cause of death classification between death certificates and verified medical files in UHL. Attribution of cause of death to comorbidities was the most common reason for discordant reporting of breast cancer-specific death. CSS calculated with cause of death information from death certificates following ICD-10 rules showed similar CSS compared to medical files. Results for CSS and RS were similar, as well. Although there are clear guidelines for registration of cause of death, periodic reviews of the implementation of these rules and continuous training of registrars and physicians may be needed in order to obtain accurate cause of death data, and measure survival based on these data. Registries should ideally combine information from different sources and review discordant cases.

## Declarations

### Ethics approval and consent to participate
This retrospective chart review study involving human participants was in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The Ethics Committee (IRB) of University Hospitals Leuven approved this study.

### Consent for publication
Consent for publication was received for individual patient's data included in the study.

### Competing interests
None of the authors have conflicts of interest related to this topic.

**Author details**
[1]Department of Oncology, KU Leuven - University of Leuven, Herestraat 49 box 7003-06, B-3000 Leuven, Belgium. [2]Belgian Cancer Registry, Research Department, Brussels, Belgium. [3]Department of General Medical Oncology, University Hospitals Leuven, B-3000 Leuven, Belgium. [4]Department of Gynecological Oncology, University Hospitals Leuven, B-3000 Leuven, Belgium. [5]Department of Radiation Oncology, University Hospitals Leuven, B-3000 Leuven, Belgium. [6]Department of Gynecology and Obstetrics, University Hospitals Leuven, B-3000 Leuven, Belgium. [7]Department of Surgical Oncology, University Hospitals Leuven, B-3000 Leuven, Belgium.

**References**
1. World Health Organization. ICD-10: international statistical classification of diseases and related health problems: tenth revision, 2nd ed. World Health Organization: 2004.
2. Brinkley D, Haybittle JL, Alderson MR. Death certification in cancer of the breast. Br Med J. 1984;289(6443):465–7. https://doi.org/10.1136/bmj.289.6443.465.
3. Rutqvist LE. Validity of certified causes of death in breast carcinoma patients. Acta Oncol (Madr). 1985;24:385–90.
4. Goldoni CA, Bonora K, Ciatto S, Giovannetti L, Patriarca S, Sapino A, et al. Misclassification of breast cancer as cause of death in a service screening area. Cancer Causes Control. 2009;20(5):533–8. https://doi.org/10.1007/s10552-008-9261-3.
5. Schaffar R, Rapiti E, Rachet B, Woods L. Accuracy of cause of death data routinely recorded in a population-based cancer registry: impact on cause-

Izci *et al. Archives of Public Health*       (2021) 79:111

Page 8 of 8

specific survival and validation using the Geneva cancer registry. BMC Cancer. 2013;13(1). https://doi.org/10.1186/1471-2407-13-609.

6.   Brenner DR, Tammemägi MC, Bull SB, Pinnaduwaje D, Andrulis IL. Using cancer registry data: agreement in cause-of-death data between the Ontario Cancer registry and a longitudinal study of breast cancer patients. Chronic Dis Can. 2009;30(1):15–8. https://doi.org/10.24095/hpcdp.30.1.04.

7.   Percy C, Stanek E, Gloeckler L. Accuracy of cancer death certificates and its effect on cancer mortality statistics. Am J Public Health. 1981;71(3):242–50. https://doi.org/10.2105/AJPH.71.3.242.

8.   Mattsson B, Rutqvist LE. Some aspects on validity of breast cancer, pancreatic cancer and lung cancer registration in Swedish official statistics. Radiother Oncol. 1985;4(1):63–70. https://doi.org/10.1016/S0167-8140(85)80063-3.

9.   Pérez-Gómez B, Aragonés N, Pollán M, Suárez B, Lope V, Llácer A, et al. Accuracy of cancer death certificates in Spain: a summary of available information. Gac Sanit Elsevier. 2006;20:42–51. https://doi.org/10.1157/13101089.

10.   Engel LW, Strauchen JA, Chiazze L, Heid M. Accuracy of death certification in an autopsied population with specific attention to malignant neoplasms and vascular diseases. Am J Epidemiol. 1980;111(1):99–112. https://doi.org/10.1093/oxfordjournals.aje.a112879.

11.   German RR, Fink AK, Heron M, Stewart SL, Johnson CJ, Finch JL, et al. The accuracy of cancer mortality statistics based on death certificates in the United States. Cancer Epidemiol. 2011;35(2):126–31. https://doi.org/10.1016/j.canep.2010.09.005.

12.   Hoel DG, Ron E, Carter R, Mabuchi K. Influence of death certificate errors on cancer mortality trends. J Natl Cancer Inst. 1993;85(13):1063–8. https://doi.org/10.1093/jnci/85.13.1063.

13.   Gobbato F, Vecchiet F, Barbierato D, Melato M, Manconi R. Inaccuracy of death certificate diagnoses in malignancy: an analysis of 1,405 autopsied cases. Hum Pathol. 1982;13(11):1036–8. https://doi.org/10.1016/S0046-8177(82)80096-8.

14.   Yin D, Morris CR, Bates JH, German RR. Effect of misclassified underlying cause of death on survival estimates of colon and rectal cancer. J Natl Cancer Inst. 2011;103(14):1130–3. https://doi.org/10.1093/jnci/djr207.

15.   Rampatige R, Mikkelsen L, Hernandez B, Riley I, Lopez AD. Systematic review of statistics on causes of deaths in hospitals: strengthening the evidence for policy-makers. Bull World Health Organ. 2014;92(11):807–16. https://doi.org/10.2471/BLT.14.137935.

16.   Tan KS. Misclassification of the actual causes of death and its impact on analysis: a case study in non-small cell lung cancer. Lung Cancer. 2019;134: 16–24. https://doi.org/10.1016/j.lungcan.2019.05.016.

17.   Wissing MD, Greenwald ZR, Franco EL. Improving the reporting of cancer-specific mortality and survival in research using cancer registry data. Cancer Epidemiol Elsevier. 2019;59:232–5. https://doi.org/10.1016/j.canep.2019.02.004.

18.   Johansson LA, Westerling R, Rosenberg HM. Methodology of studies evaluating death certificate accuracy were flawed. J Clin Epidemiol. 2006; 59(2):125–31. https://doi.org/10.1016/j.jclinepi.2005.05.006.

19.   Johansson LA, Björkenstam C, Westerling R. Unexplained differences between hospital and mortality data indicated mistakes in death certification: an investigation of 1,094 deaths in Sweden during 1995. J Clin Epidemiol. 2009;62(11):1202–9. https://doi.org/10.1016/j.jclinepi.2009.01.010.

20.   Begg CB. Attribution of deaths following Cancer treatment. CancerSpectrum Knowl Environ. 2002;94:1044–5.

21.   Schaffar R, Rachet B, Belot A, Woods L. Cause-specific or relative survival setting to estimate population-based net survival from cancer? An empirical evaluation using women diagnosed with breast cancer in Geneva between 1981 and 1991 and followed for 20 years after diagnosis. Cancer Epidemiol. 2015;39(3):465–72. https://doi.org/10.1016/j.canep.2015.04.001.

22.   Schaffar R, Rachet B, Belot A, Woods LM. Estimation of net survival for cancer patients: relative survival setting more robust to some assumption violations than cause-specific setting, a sensitivity analysis on empirical data. Eur J Cancer. 2017;72:420–6.

23.   Dignam JJ, Huang L, Ries L, Reichman M, Mariotto A, Feuer E. Estimating breast cancer-specific and other-cause mortality in clinical trial and population-based cancer registry cohorts. Cancer. 2009;115(22):5272–83. https://doi.org/10.1002/cncr.24617.

24.   Howlader N, Ries LAG, Mariotto AB, Reichman ME, Ruhl J, Cronin KA. Improved estimates of cancer-specific survival rates from population-based data. J Natl Cancer Inst. 2010;102(20):1584–98. https://doi.org/10.1093/jnci/djq366.

25.   Sarfati D, Blakely T, Pearce N. Measuring cancer survival in populations: relative survival vs cancer-specific survival. Int J Epidemiol. 2010;39(2):598–610. https://doi.org/10.1093/ije/dyp392.

26.   Skyrud KD, Bray F, Møller B. A comparison of relative and cause-specific survival by cancer site, age and time since diagnosis. Int J Cancer. 2014; 135(1):196–203. https://doi.org/10.1002/ijc.28645.

27.   Makkar N, Ostrom QT, Kruchko C, Barnholtz-Sloan JS. A comparison of relative survival and cause-specific survival methods to measure net survival in cancer populations. Cancer Med. 2018;7(9):4773–80. https://doi.org/10.1002/cam4.1706.

28.   Cancer Burden in Belgium 2004–2017, Belgian Cancer Registry, Brussels. 2020 [cited 2020 Feb 18]. Available from: https://kankerregister.org/Publications

29.   CBSS - Crossroads Bank for Social Security. [cited 2020 Feb 18]. Available from: https://www.ksz-bcss.fgov.be/nl/documents-list

30.   Statbel. StatBel - Life expectancy and life tables [Internet]. 2020 [cited 2020 Feb 18]. Available from: https://statbel.fgov.be/en/themes/population/mortality-life-expectancy-and-causes-death/life-expectancy-and-life-tables

31.   Agentschap Zorg en Gezondheid. 2017 [cited 2020 Feb 18]. Available from: http://www.zorg-en-gezondheid.be/cijfers/

32.   Observatoire de la Santé et du Social de Bruxelles-Capitale. Available from: https://www.ccc-ggc.brussels/nl/observatbru/indicatoren/mortaliteitsindicatoren

33.   Agence pour une Vie de Qualité (AViQ). Available from: http://sante.wallonie.be/?q=transfert-competences-sante/donnees-socio-sanitaires

34.   Johansson L, Pavillon G. IRIS: A language-independent coding system based on the NCHS system MMDS. WHO-FIC 2005/B.6.2;1-5.

35.   IRIS Software. 2020 [cited 2020 Feb 17]. Available from: https://www.dimdi.de/dynamic/en/classifications/iris-institute/downloads/

36.   Harteloh P. The implementation of an automated coding system for cause-of-death statistics. Informatics Heal Soc Care. 2020;45:1–14.

37.   WHO. E-Health. 2016 [cited 2020 Jun 16]. Available from: https://www.who.int/ehealth/en/

38.   Therneau TM, Grambsch PM. Expected survival. 2000, Expected Survival.

39.   Melnick EL. Modeling Survival Data. Lovric M, editor. Int Encycl Stat Sci. Berlin: Springer Berlin Heidelberg; 2011;841–4.

40.   Kirch W, editor. Kappa CoefficientKappa coefficient. Encycl Public Heal. Dordrecht: Springer Netherlands; 2008. p. 821–822.

41.   Spearman Rank Correlation Coefficient. Concise Encycl Stat. New York: Springer New York; 2008. p. 502–5.

42.   Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc. 1958;53(282):457–81. https://doi.org/10.1080/01621459.1958.10501452.

43.   R Team Core. A language and environment for statistical computing. Vienna: R Found Stat Comput; 2017.

44.   Ederer F, Axtell LM, Cutler SJ. The relative survival rate: a statistical methodology. Natl Cancer Inst Monogr. 1961;6:101-21.

45.   Perme MP, Stare J, Estève J. On estimation in relative survival. Biometrics. 2012;68(1):113-20. https://doi.org/10.1111/j.1541-0420.2011.01640.x. Epub 2011 Jun 20. PMID: 21689081.

46.   Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. Am J Ment Defic. 1981;86(2):127-37. PMID: 7315877.

47.   Mcginn T, Newman T, Keitz S, Leipzig R, Guyatt G. Tips for teachers of evidence-based medicine: 3. Understanding and calculating kappa (vol 171, pg 1369, 2004). Can Med Assoc J. 2005;173:18.

48.   Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74. PMID: 843571.

49.   Fleiss JL. Measuring nominal scale agreement among many raters. Psychol Bull. 1971;76(5):378–82. https://doi.org/10.1037/h0031619.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.