Determinants and interpretation of death certificate only proportions in the initial years of newly established cancer registries

Hermann Brenner *, Lina Jansen

Division of Clinical Epidemiology and Aging Research, German Cancer Research Center, INF 581, D-69120 Heidelberg, Germany

Abstract   **Background**: The proportion of cases notified by death certificate only (DCO) is a commonly used criterion to judge completeness of cancer registration even though it is affected by additional factors, particularly during initial years of newly established registries.
Methods: Based on cancer registry data from the United States, we provide model calculations to demonstrate the magnitude and time course of the impact of the following mechanisms on DCO proportions of "young" registries: registration of cancer deaths from patients diagnosed prior to the registration period and delayed registration by death certificate of patients diagnosed but not reported after initiation of registration.
Results: DCO proportions of up to ⩾30% can be expected from deaths of previously diagnosed patients during the first year of registration. Although this proportion is expected to gradually diminish over subsequent years, DCO proportions may be dominated for several years by this source, which may still be relevant after 10 or more years of cancer registration for cancers with relatively large proportions of late deaths. Otherwise, however, underreporting during patients' lifetime is expected to become the predominant source of DCO proportions in the long run.
Conclusions: Our results may guide interpretation of DCO proportions of relatively "young" cancer registries.
© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Population-based analyses of cancer incidence and survival are important tools for monitoring progress against cancer. Their validity strongly depends on high-quality and completeness of incidence and follow-up data of population-based cancer registries.[1–4] A commonly used data quality indicator is the proportion of cases notified by death certificate only (DCO).[2,5,6] Such cases may arise from various sources, including incomplete registration during lifetime of patients dying from cancer. DCO proportions are therefore often used as a criterion for inclusion of cancer registries in collaborative studies on cancer incidence and survival.[7–9]

---

* *Corresponding author*: Tel.: +49 6221 421301; fax: +49 6221 421302.

E-mail address: h.brenner@dkfz.de (H. Brenner).

However, for newly established cancer registries, interpretation of DCO proportions is complicated by the following mechanisms: On one hand, DCO cases may reflect deaths of cancer patients diagnosed prior to the registration period. The proportion of DCO cases arising from this source is expected to diminish over time and is unrelated to completeness of cancer registration. On the other hand, DCO cases pertaining to patients diagnosed during the registration period but not registered during lifetime may be very few during the first year of registration but are expected to increase over the initial years of registration through accumulation of the number and follow-up time of unregistered patients.

We aim to illustrate the expected contributions of both sources of DCO cases and their development over time during the initial years of registration of a new cancer registry in order to enhance interpretation of DCO proportions of relatively "young" cancer registries.

2. Methods

2.1. Database

Our illustrations are based on model calculations using data from the United States Surveillance, Epidemiology, and End Results (SEER) Program. The SEER-9 database issued in April 2012 includes data on incident cancer cases in 1973–2009 from population-based cancer registries in Connecticut, New Mexico, Utah, Iowa, Hawaii, Atlanta (from 1975 on), Detroit, Seattle-Puget Sound (from 1974 on) and San Francisco–Oakland, which together cover a population of around 30 million people.[10] Geographical areas are selected for inclusion in the SEER Program based on their ability to operate and maintain a high-quality population-based cancer reporting system and their epidemiologically significant population subgroups. The SEER population is comparable to the general United States population, although it is more urban and has a higher proportion of foreign-born persons than the latter, and for certain cancer sites there is underrepresentation of US cancer mortality experience.[11] We chose the SEER-9 database for illustration due to its long time series and as the true proportion of missed cases and DCO proportions are very low.[12]

2.2. Statistical analyses

In a first set of analyses, we assessed the expected DCO proportions arising from patients diagnosed prior to but dying from their cancer during the registration period. For this purpose we carried out model calculations in which we derived expected DCO proportions due to such cases in each calendar year from 2000 to 2009 if registration had started in 2000 only. Hypothetical numbers of incident cases in each calendar year were derived as the sum of true incident cases (T) and apparent incident cases contributed by DCO cases resulting from deaths due to the various cancers of patients diagnosed prior to 2000 (A). Given that the SEER-9 database includes cancer cases diagnosed since 1973 and their cause specific deaths at very high levels of completeness, such cases can be very reliably identified from the SEER-9 database. DCO proportions expected from this source were then derived as $A / (T + A)$ for each single calendar year from 2000 to 2009. The analyses were done separately for the 25 most common cancer sites in 2000–2009.

In a second set of analyses, we assessed the expected development of "true" DCO proportions from 2000 to 2009 if the SEER-9 registries had started registration in 2000 only and if various proportions of cases had been missed during lifetime. The following scenarios were assessed: (i) 20% of cases unregistered during lifetime in each calendar year, (ii) 10% of cases unregistered during lifetime in each calendar year and (iii) proportion of cases unregistered during lifetime gradually decreasing by 2 percent units per year from 20% in 2000 to 2% in 2009. Scenarios (i) and (ii) were chosen to demonstrate the impact of various extents of underreporting. Scenario (iii) was chosen to simulate a situation of decreasing underreporting over time which would typically be expected in the build-up and consolidation phase of a newly established registry. Expected DCO proportions resulting from various proportions of underreporting of cases during lifetime, denoted U, were determined by calendar year from a weighted analysis in which all cases diagnosed in 2000–2009 were included in the analysis with a weight of ($1-U$) as incident cases in the year of diagnosis, and cases dying from their cancer were additionally included with a weight of U as DCO case in their year of death.

In a third step, DCO proportions expected from both sources (pre 2000 diagnoses and underreporting) were jointly assessed by combining the first set of analyses with scenario (iii) of the second set of analyses.

All analyses were done by the SAS statistical software package version 9.2.

3. Results

Table 1 shows the numbers of newly registered cases in 2000–2009 and the expected DCO proportions for each calendar year for the 25 most common cancer diagnoses. The latter pertain to the hypothetical situation that registration had started in 2000 only, and deaths due to cancer of cancer patients diagnosed before 2000 were registered as DCO cases. Expected DCO proportions due to cancers diagnosed before 2000 were highest in 2000, and gradually decreased in subsequent calendar years. For cancers with very poor prognosis, such as cancers of the oesophagus, pancreas or lung, expected DCO proportions were very high (>20%) in 2000, but