Cancer Epidemiology 37 (2013) 20–28

Contents lists available at SciVerse ScienceDirect

Cancer Epidemiology

The International Journal of Cancer Epidemiology, Detection, and Prevention

journal homepage: www.cancerepidemiology.net

The impact of National Death Index linkages on population-based cancer survival rates in the United States

Christopher J. Johnson [a,*], Hannah K. Weir [b], Aliza K. Fink [c], Robert R. German [b], Jack L. Finch [d], Randi K. Rycroft [d], Daixin Yin [e]

the Accuracy of Cancer Mortality Study Group[1]

[a] Cancer Data Registry of Idaho, 615 North 7th Street, Boise, ID 83701, United States
[b] Division of Cancer Prevention and Control, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention (CDC), Atlanta, GA, United States
[c] ICF International, 530 Gaither Road, Rockville, MD 20850, United States
[d] Colorado Central Cancer Registry, Colorado Department of Public Health and Environment, Denver, CO, United States
[e] California Cancer Registry, Sacramento, CA, United States

ARTICLE INFO

Article history:
Received 14 January 2012
Received in revised form 10 August 2012
Accepted 12 August 2012
Available online 7 September 2012

Keywords:
Death certificate accuracy
National Death Index
Cancer survival

ABSTRACT

Background: In order to ensure accurate survival estimates, population-based cancer registries must ascertain all, or nearly all, patients diagnosed with cancer in their catchment area, and obtain complete follow-up information on all deaths that occurred among registered cancer patients. In the US, linkage with state death records may not be sufficient to ascertain all deaths. Since 1979, all state vital statistics offices have reported their death certificate information to the National Death Index (NDI). *Objective*: This study was designed to measure the impact of linkage with the NDI on population-based relative and cancer cause-specific survival rates in the US. *Methods*: Central cancer registry records for patients diagnosed 1993–1995 from California, Colorado, and Idaho were linked with death certificate information (deaths 1993–2004) from their individual state vital statistics offices and with the NDI. Two databases were created: one contained incident records with deceased patients linked only to state death records and the second database contained incident records with deceased patients linked to both state death records and the NDI. Survival estimates and 95% confidence intervals from each database were compared by state and primary site category. *Results*: At 60 months follow-up, 42.1–48.1% of incident records linked with state death records and an additional 0.7–3.4% of records linked with the NDI. Survival point estimates from the analysis without NDI were not contained within the corresponding 95% CIs from the NDI augmented analysis for all sites combined and colorectal, pancreas, lung and bronchus, breast, prostate, non-Hodgkin lymphoma, and Kaposi sarcoma cases in all 3 states using relative survival methods. Additional combinations of state and primary site had significant survival estimate differences, which differed by method (relative versus cause-specific survival). *Conclusion*: To ensure accurate population-based cancer survival rates, linkage with the National Death Index to ascertain out of state and late registered deaths is a necessary process for US central cancer registries.

Abbreviations: ACM, Accuracy of Cancer Mortality Statistics Based on Death Certificates; CDC, Centers for Disease Control and Prevention; ICD, International Classification of Diseases; KS, Kaposi sarcoma; NAACCR, North American Association of Central Cancer Registries; NCHS, National Center for Health Statistics; NDI, National Death Index; NPCR, National Program of Cancer Registries; SEER, Surveillance, Epidemiology and End Results; UCD, underlying cause of death; US, United States of America; USCS, United States Cancer Statistics; VSO, vital statistics office (state).

* Corresponding author. Tel.: +1 208 489 1380; fax: +1 208 344 0180.
E-mail addresses: cjohnson@teamiha.org (C.J. Johnson), hbw4@cdc.gov (H.K. Weir), Aliza.Fink@icfi.com (A.K. Fink), BobGerman2010@gmail.com (R.R. German), Jack.Finch@state.co.us (J.L. Finch), Randi.Rycroft@state.co.us (R.K. Rycroft), Dx35y@yahoo.com (D. Yin).

[1] See Appendix.

1877-7821/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.canep.2012.08.007

1. Introduction

Cancer registries play a critical role in monitoring effective cancer control activities by providing population-based incidence and survival data. In addition, information on cancer survival can be used to evaluate the overall effectiveness of healthcare delivery to cancer patients [1,2].

Net survival estimates the probability of surviving cancer in the absence of other causes of death and provides a means for tracking survival over time and across populations with different life expectancies. The two methods for estimating net survival are relative survival (i.e., the ratio of the observed survival in the cancer patient cohort to the expected survival from a comparable group in the general population) and cause-specific survival (i.e., probability of death from a specific cause where deaths from all others causes are treated as censored observations).

In order to ensure valid relative and cause-specific survival estimates, population-based cancer registries must first ascertain all, or nearly all, patients diagnosed with cancer in their catchment area [3], and second, obtain complete follow-up information on all deaths that occurred among registered cancer patients [4,5].

Ascertaining deaths can be particularly challenging to cancer registry staff as the resources required to conduct follow-up increases as the number of registered patients increases. Over time, these numbers will continue to increase due to the maturity of the cancer registry, a growing and aging population, and improved survival [6].

To assist in the ascertainment of cancer patients who may have been missed at the time of their diagnosis, or who may have been diagnosed with cancer only at the time of their death, cancer registries routinely link their incidence data with death certificate data from their jurisdictional vital records offices [7]. During this linkage process, known as death clearance, the cancer registry database can be updated with cause and date of death information among deceased incident cases.

Over the past several decades, advances in early detection and treatment have resulted in increased survival time for cancer patients [8]. Patients may move from one jurisdiction to another within the same country between the time of their diagnosis and their death and, depending on practices regarding the recording of deaths in the country, the cancer registry reporting the incident case may not learn of the patient's death. In the United States (US), deaths are recorded in the state where the decedent expired and shared, if different, with the state of residence at the time of death. This may or may not be the state in which a person resided when they were diagnosed with cancer. Another limitation may result from the exchange of information between state vital statistics offices (VSO). For example, a cancer patient may seek treatment out of state and subsequently die. The VSO where the death occurred may be slow to report the death to the VSO in the state where the patient resided, or the VSO may place restrictions on the use of death certificate data in a way that precludes or impedes the use of the death records in linkages with cancer registry records [9]. Therefore, linkage with state death records may not be sufficient to ascertain all deaths that occurred among cancer patients registered in statewide or metropolitan-area based cancer registries.

The US Centers for Disease Control and Prevention (CDC)'s National Center for Health Statistics (NCHS) maintains the National Death Index (NDI). Since 1979, VSOs in all 50 states, Puerto Rico, and the Virgin Islands have reported their death certificate information to the NDI [10,11]. Cancer registries are encouraged to link their incidence records with both state death records and the NDI for the purpose of ascertaining deaths and updating date and cause of death information [9]. Record linkage with the NDI has been used with both adult and pediatric cohort studies [12–17].

The present study uses secondary data from the Accuracy of Cancer Mortality Statistics Based on Death Certificates (ACM) study [18]. The main objective of the ACM study was to characterize the concordance between cancer cause of death information from death certificates and primary cancer site at diagnosis recorded in US cancer registries [19]. Cases included in the ACM study had follow-up for at least 9 years, and data collected for the ACM study provided a unique opportunity to investigate the impact of NDI linkages on survival estimates by comparing survival rates based on deaths ascertained solely by state death record linkages with rates based on linkage to state death records and the NDI.

2. Materials and methods

The ACM study has been described in detail elsewhere [18]. Briefly, population-based central cancer registries in California, Colorado, and Idaho were selected because they met study eligibility criteria including but not limited to the following: the registry was statewide and population-based; cancer incidence data were complete and high quality (e.g., met publication criteria [20]); and the registry performed routine death clearance with state death records. In addition, these registries agreed to send their incidence data to the NDI to ascertain deaths that were not recorded in their state vital records offices. If a death was ascertained via the state process, the record was not sent for NDI linkage.

In the current study, we investigated the impact of NDI linkage on 5-year cancer survival rates based on incident cases diagnosed between January 1, 1993 through December 31, 1995 and deaths that occurred up to 5 years after diagnosis.

2.1. Analysis

Data from the three statewide cancer registries were combined into one dataset using SAS (Version 9.2, Cary, NC), and two datasets were subsequently created and processed using SEER*Prep (Version 2.4.5, Information Management Services, Inc., Silver Spring, MD). A field demarcating the source of follow-up information at the state cancer registry was used to identify deaths ascertained through NDI linkages. To maintain consistency between the datasets, all patients with vital status alive were censored at the end of the study period, December 31, 2004. The first dataset included deaths ascertained through state processes and NDI linkages (NDI augmented file). In order to keep sample sizes consistent in both datasets, the second dataset (NDI censored file) was created with all NDI deaths censored at the end of the study period (vital status alive as of December 31, 2004) as if the NDI linkage had not been performed. A small proportion of deaths was ascertained via follow-up by hospitals and other sources and lacked cause of death information (1.6% in California, 1.0% in Colorado, and 0.7% in Idaho). These are included in the totals for state source of death ascertainment in Table 1 and included in the survival analyses as is the general practice in the states.

Analysis was performed using the survival functions in SEER*Stat (Version 7.0.5, Information Management Services, Inc., Silver Spring, MD). Cancer site-specific 5-year survival rates were generated according to SEER primary site recodes that group cases by major site/histology categories and are commonly used in the reporting of cancer statistics [21]. Ninety-five percent confidence intervals (CIs) for the survival functions were based on the log–log transformation. Parallel survival analyses were conducted on the two datasets in order to compare site-specific survival rates with NDI (augmented) and without NDI (censored) linkage results. Thus, the comparison is between death ascertainment solely by state processes versus state processes supplemented by the NDI.